Transportøkonomisk institutt
Stiftelsen Norsk senter for samferdselsforskning

# Noen transportøkonomiske emner

Harald Minken

1936/2023

Tittel: Noen transportøkonomiske emner
Tittel engelsk: Some issues in transport economics
Forfatter: Harald Minken
Dato: 02.2023
TØI-rapport: 1936/2023
Antall sider: 228
ISSN elektronisk: 2535-5104
ISBN elektronisk: 978-82-480-1995-4
Finansieringskilder: Transportøkonomisk institutt
TØIs p.nr.: 4330 – Samfunnsøkonomiske metoder
Prosjektleder: Askill Harkjerr Halse
Kvalitetsansvarlig: Askill Harkjerr Halse
Fagfelt: Samfunnsøkonomiske analyser
Emneord: Samfunnsøkonomisk analyse, regneregler, handlingsregler, representativ konsument, gjennomsnittskonsument

## Kort sammendrag

Denne rapporten er en samling artikler om metodespørsmål i samfunnsøkonomisk analyse av transporttiltak. En av hensiktene med rapporten er å etablere regler for hvordan samfunns-økonomiske analyser i samferdselssektoren skal beregnes og rapporteres. Grunntrekkene i disse reglene er allerede godt etablert i Norge, men vi fremmer også forslag om å beregne virkninger vi ikke tar hensyn til i dag. En annen hensikt er å vise at skissemessige små modeller kan brukes til å trekke slutninger om hva som er ønskelig samferdselspolitikk. Disse modellene består i regelen bare av en aggregert etterspørselsfunksjon, en kostnadsfunksjon og en budsjettbetingelse.

## Summary

This report is a collection of papers on the methods of transport project appraisal. One of the purposes of the report is to establish rules for computing and reporting costs and benefits of all kinds of transport policies. The basics of these rules are already well established in Norway, but we also propose to calculate impacts not taken into account under the present regime. Another purpose is to show that small conceptual models can be used to good effect to derive rules concerning transport policy. Such models will usually only consist of an aggregate demand function, a cost function, and a budget constraint.



Transportøkonomisk Institutt, Gaustadalléen 21, 0349 Oslo Telefon 22 57 38 00 E-post: toi@toi.no www.toi.no

# Forord

Denne rapporten er en samling artikler av Harald Minken om metodespørsmål i samfunns-økonomisk analyse av transporttiltak. Til dels dreier det seg om utdrag fra tidligere TØI-rapporter og upubliserte arbeidsdokumenter, men her finns også manuskripter som aldri tidligere har blitt publisert i noen form.

Språket i hele rapporten er på norsk. Det foreligger imidlertid en tilsvarende rapport, TOI report 1934/2023, med ulike bidrag på engelsk om samme slags emner. Ingen av artiklene der er oversettelser av artiklene på norsk i denne rapporten. Ingen av artiklene i den engelske rapporten er heller oversatt til norsk i denne rapporten.

Det er forfatteren som har stått for utvalget av artikler i begge rapportene. Når han nå pensjonerer seg, vil han takke alle på TØI for vennskap, støtte og hjelp gjennom snart 30 år. Særlig vil han takke Farideh Ramjerdi, Lasse Fridstrøm og det gamle biblioteket på TØI for å ha lært ham transportøkonomi, og Olav Eidhammer og Kjell Werner Johansen, som etter-fulgte hverandre som avdelingsledere på instituttets avdeling for transportøkonomi, for å ha gitt ham rom til å praktisere det han hadde lært, hver på sin egen måte.

Andre arbeider av Minken og de andre forskerne på instituttet er å finne på https://www.toi.no/publikasjoner/. (Skift mellom norsk og engelsk øverst i høyre hjørne av vinduet.)

Oslo, februar 2023
Transportøkonomisk institutt

| | |
|---|---|
| Bjørne Grimsrud | Kjell W. Johansen |
| Administrerende direktør | Avdelingsleder |

# Innhold

Transportøkonomisk institutt
Stiftelsen Norsk senter for samferdselsforskning

NORSK **Sammendrag**

# Noen transportøkonomiske emner

TØI rapport 1936/2023 • Forfatter: Harald Minken • Oslo 2023 • 228 sider

Denne rapporten er en samling artikler om metodespørsmål i samfunnsøkonomisk analyse av transporttiltak. En av hensiktene med rapporten er å etablere regler for hvordan samfunnsøkonomiske analyser i samferdselssektoren skal beregnes og rapporteres. Grunntrekkene i disse reglene er allerede godt etablert i Norge, men vi fremmer også forslag om å beregne virkninger vi ikke tar hensyn til i dag. En annen hensikt er å vise at skissemessige små modeller kan brukes til å trekke slutninger om hva som er ønskelig samferdselspolitikk. Disse modellene består i regelen bare av en aggregert etterspørselsfunksjon, en kostnadsfunksjon og en budsjettbetingelse.

Om et så enkelt oppsett skal kunne gi regler for hele samfunnet, må jo folk være likere hverandre enn de normalt vil være – dvs. det må eksistere en representativ konsument, som det heter. Men gjør det ikke det, får vi klare oss med det vi kaller en gjennomsnittlig konsument. Gitt at vi også tar skritt for å utjamne de økonomiske og sosiale forskjellene, mener vi at vi stort sett kan gjøre det som er best for gjennomsnittskonsumenten (eller den gjennomsnittlige trafikanten, i vårt tilfelle). Det er på dette – ikke helt stødige – grunnlaget vi gir våre transportøkonomiske råd og anbefalinger.

ENGLISH Summary

# Some issues in transport economics

TØI Report 1936/2023 • Author: Harald Minken • Oslo 2023 • 228 pages

This report is a collection of on the methods of transport project appraisal. One of the purposes of the report is to establish rules for computing and reporting costs and benefits of all kinds of transport policies. The basics of these rules are already well established in Norway, but we also propose to calculate impacts not taken into account under the present regime. Another purpose is to show that small conceptual models can be used to good effect to derive rules concerning transport policy. Such models will usually only consist of an aggregate demand function, a cost function, and a budget constraint.

For such a simple design to produce valid recommendations and rules for the whole of society, people will have to be more equal than they usually are. That is, a so-called representative consumer will have to exist. But if not, we will have to make do with the so-called average consumer. Provided we take steps to reduce economic and social inequality, we think we may indeed recommend policies that improve the net benefits of the average consumer – or the average traveller, in our case. It is on this not totally stable foundation we base our transport economics advice and recommendations.

# 1 Innledning

Transportøkonomi er en normativ vitenskap. Gjennomgående tar den sikte på å finne ut hvilke priser, avgifter og investeringer som gir den største samfunnsøkonomiske lønnsomheten i en gitt situasjon. Dette er hovedmålet. Som regel er det også et mål at transportpolitikken i den konkrete situasjonen bidrar til jamnere fordeling av goder og byrder, at den bevarer naturverdier og kulturverdier, og at den reduserer ulykker, klimautslipp og andre skadelige utslipp.

I prinsipp måles nytte ved hva de som vil få oppleve nytten, er villig til å betale for å få den, og kostnaden måles ved hva de som påføres ulempene må ha i kompensasjon for å akseptere det. Men når det gjelder et konkret forslag om å bygge ny infrastruktur eller endre driftsopplegg, priser og avgifter, er det i praksis umulig å kartlegge fordelingen av nytte og kostnader på de berørte individene på forhånd. Det er heller ikke lett etter at et tiltak er satt ut i livet. Vi opererer derfor med gjennomsnittsverdier på ulike nyttevirkninger og kostnader. De er funnet i såkalte betalingsvillighetsstudier. Verdien av spart reisetid er et eksempel.

I et konkret tilfelle er det også vanskelig å finne ut på forhånd hvordan alle ulike mennesker vil tilpasse seg til et transporttiltak. Hvor mye mer vil de reise? Vil de endre hvor de reiser, og hva slags transportmiddel de vil bruke? Vi bruker derfor såkalte aggregerte etterspørselsfunksjoner, som i praksis sier hvordan gjennomsnittsmennesker vil reagere på tiltaket, ikke hvordan hver enkelt vil reagere. I artikkelen «The Pareto criterion and the Kaldor Hicks criterion», som er gjengitt i TØI-rapporten «Transport cost benefit analysis – basic assumptions and accounting rules», viser vi at det bare er i ganske spesielle tilfeller at dette gir et dekkende bilde av hva alle de berørte opplever. Noen vil mene at de vinner på tiltaket, andre at de taper. Så hvis vi vurderer saka på grunnlag av hvordan gjennomsnittsmennesket stiller seg til det, vil vi både kunne ta feil om hvor mange som vi bruke tiltaket, hva slags nytte de har av det, og hva slags kostnad som påføres de som ikke ser seg tjent med å bruke det.

Løsningen som artikkelen tar til orde for, er å prøve å innarbeide kompenserende tiltak til taperne når prosjektet utformes. Det vil i praksis innebære at forskjellen mellom vinnere og tapere blir mindre, eller med andre ord at flere vurderer prosjektet med gjennomsnittsmennesket briller. Om dette lykkes, vil en gjennomsnittsberegning kunne gi et tilstrekkelig dekkende bilde, både av hvordan brukerne tilpasser seg til tiltaket, og hvilken nytte de har av det. Det vil ikke alltid klaffe, men tilstrekkelig ofte til at vi kan basere nyttekostnadsanalysen på forhånd på det.

Artiklene i denne rapporten er i stor grad basert på en aggregert etterspørselsfunksjon, altså at de som påvirkes av et tiltak, reagerer som de skulle være ett eneste gjennomsnittsindivid, og at trafikantnytten til dette gjennomsnittsindividet er et godt uttrykk for hva alle de som påvirkes, får av nytte eller kostnader når tiltaket settes ut i livet. På dette grunnlaget gjør vi greie for hvordan den samfunnsøkonomiske analysen av tiltaket gjøres, eller snarere burde gjøres, og hvordan en skal velge mellom ulike alternative tiltak og sette dem sammen til en plan. Dette er hovedinnholdet i andre kapittel.

I tredje kapittel lager vi teoretiske modeller som i større grad har til hensikt å foreslå forbedringer av dagens praksis, og i fjerde kapittel behandler vi trafikale virkninger som ikke er med i dagens samfunnsøkonomiske analyser, men som nokså enkelt kunne innføres etter oppskrifta vi gir der.

**Andre kapittel: Samfunnsøkonomiske analyser og prosjektvalg**

Den første artikkelen her gir en oppskrift på hvordan en samfunnsøkonomisk analyse i transportsektoren skal beregnes og føres. Oppskrifta er i det vesentlige den samme som blei beskrevet i Minken og Samstads TØI-rapport 798/2005, og som her til lands nå kalles bruttometoden. Men Minken og Samstad har igjen sin metode fra The Common Appraisal Framework, som i en kort periode var den vedtatte oppskrifta på hvordan nyttekostnadsanalyser i samferdelssektoren i Storbritannia skulle rapporteres, før Webtag tok over. Hovedprinsippet er at nytte og kostnader føres separat for fire sektorer – trafikanter, trafikkselskaper, det offentlige og samfunnet for øvrig. Overføringer fra en sektor til en annen føres da som kostnad for den første og som inntekt for den andre. Dette gjør det lettere å skille mellom virkelige virkninger for samfunnet og reine overføringer, samtidig som det legger et godt grunnlag for å beregne hvem som vinner og hvem som taper. Metoden har stort sett vært fulgt i samferdselssektoren siden TØI-rapporten.

«Om noen tekniske forutsetninger …» er en kritisk diskusjon av analyseperiode, levetid, restverdi og kalkulasjonsrente, samt hva som må gjøres om levetida er kortere enn analyseperioden. Det er konklusjoner her som burde inn i veglederne.

«Prinsipper for vegprising i kombinasjon med miljøkrav» er et upublisert notat om hvordan optimal vegpris endrer seg ved en eksogen etterspørselsendring eller ved innføring av flere bibetingelser, som skyggepris på offentlige midler eller klimagassrestriksjoner.

Til tross for at eldre vegledere i samfunnsøkonomisk analyse er nøye med å anbefale usikkerhetsanalyser som verktøy til å identifisere forholdene som kan forrykke konklusjonene av analysen, har usikkerheten ofte liten oppmerksomhet i praksis. (De beste KS-analysene er et unntak.) Arbeidsdokumentet «Problemer i usikkerhetsanalyse av samfunnsøkonomiske beregninger» fra 2011 peker ut mangler ved daværende praksis, og foreslår tiltak for å forbedre situasjonen.

Det upubliserte notatet «Drivstoffavgift, bompenger og skattefinansiering» er en gjennomgang av hvilke former for avgift som egner seg til hva, og hvordan arbeidsdelingen mellom virkemidlene bør være. Når er for eksempel bompenger en god løsning? Hvordan skal drivstoffavgifta settes? Hvordan skal en kilometerbasert vegpris innafor et bestemt geografisk område utformes?

Den supplerende prosjektvalgsregelen i avsnitt 2.6 gjelder tilfellet der prosjektene i større eller mindre grad kan finansieres utenom samferdselsbudsjettet. Om det er fornuftig å gjøre, avhenger av lønnsomheten av slike prosjekter og lønnsomheten av det siste som får plass i samferdselsbudsjettet.

Siste avsnitt i andre kapittel, altså avsnitt 2.7, handler om prosjektvalget igjen.

Den normale framgangsmåten når et nytt prosjekt skal konsekvensutredes, er at det utredes flere forskjellige måter å realisere prosjektet på, og den presumptivt beste av dem går videre og blir eventuelt tatt inn i Nasjonal transportplan. For prosjekter over en viss kostnadsgrense er det faktisk et krav at det gjøres på denne måten – det kalles en konseptvalgutredning. Men gitt at det finnes en samlet investeringsramme som hele transportplanen må holde seg innafor, er det ikke utenkelig at et billigere, men litt mindre lønnsomt alternativ kan gjøre at det blir rom i planen til et nytt og mer lønnsomt alternativ i et annet prosjekt, slik at planen som helhet blir mer lønnsom. Det kan altså lønne seg å la den samlede lønnsomheten i hele planen overstyre valget av alternativ i det enkelte prosjektet. Det er faktisk ikke vanskelig å ta hensyn til denne muligheten. «Optimalt prosjektvalg» er en superkort redegjørelse for den teoretisk riktige framgangsmåten for å velge prosjekter til en plan med en gitt kostnads-ramme, gitt at hvert prosjekt kan foreligge i flere forskjellige alternative utforminger. En grundigere redegjørelse finnes på engelsk i artikkelen «Project selection with sets of mutually exclusive alternatives», i nr. 6/2016 av det vitenskapelige tidsskriftet «Economics of Transportation». Artikkelen er også gjenopptrykt i TØI-rapport 1934/2023.

Avsnitt 2.7 har også en kort redegjørelse for problemet når det finnes flere typer begrens-ninger, for eksempel et klimabudsjett i tillegg til kostnadsramma i NTP. Dette er et vesentlig mer komplisert problem.

**Teoretiske spørsmål**

Notatet om byvekst og transport inngår i en rapport til det såkalte Effektutvalget, som leverte sin innstilling i 2003. Det skisserer to ulike teoritradisjoner på området, nemlig «urban economics» og «ny geografisk økonomi». Det konkluderer med at byenes størrelse og vekst avgjøres i samspillet mellom agglomerasjonsfordeler på den ene sida, og trengsels-kostnader i form av kø i transporten, høye husleier og andre ubehagelige sider ved tette ansamlinger av folk på den andre sida. Transporttiltak i byene kan derfor gi grunnlag for større byer. Hvorvidt dette fører til økt økonomisk vekst kommer an på om det framleis er agglomerasjonsfordeler å hente ut.

Arbeidsdokumentet «Tidsverdienes inntektsavhengighet og velferdsfunksjonens form» fra 2011 drøftet sammenheng mellom inntektsutviklingen i samfunnet og verdien av spart reisetid, slik den studeres i tidsverdiundersøkelser og innarbeides i transportmodellene. Vi bruker en såkalt deSerpa-modell til å finne hvilke faktorer som påvirker tidsverdiene, og finner at det ikke er rimelig at de vil utvikle seg i takt med BNP pr. innbygger. Snarere burde den utvikle seg i takt med reallønna pr. time etter skatt. Samfunnsutviklingen de siste tiårene, med flere under utdanning, flere pensjonister og færre på timelønn, bør tilsi at tidsverdiene øker langsommere enn timelønna.

Arbeidsdokumentet har også en kort omtale av vilkåret for at det skal finnes en representa-tiv konsument, altså for bruken av en aggregert etterspørselsfunksjon som ikke er en funk-sjon av inntekt.

Tidsverdiene er skeivfordelt, dvs. at det er flere med lavere tidsverdier enn den gjennom-snittlige, og færre med høye verdier. Dette innebærer at i teorien vil flertallet av de berørte stemme nei til et prosjekt som gir tidsbesparelser som er marginalt større enn kostnaden. Det burde altså merkelig nok være flere lønnsomme prosjekter enn de som kan vedtas i en demokratisk avstemning blant brukerne. Det vises i artikkelen «Nyttekostnadsanalyse og flertallsavgjørelser».

Arbeidsdokumentet «Merknader om mernytte» går gjennom de ulike formene for mernytte og uttrykker behov for nye modeller som kan tallfeste disse effektene. Men det stiller også spørsmålet om hvorfor det ikke alternativt kan gjøres noe med markedsimperfeksjonene som skaper slik mernytte. Det ville jo være bedre.

«Samfunnsøkonomisk optimal køpris i det lange løp» er et kort utdrag fra et vedlegg til en av Dovre og TØIs KS1-rapporter fra 2017. Utdraget viser at en samfunnsøkonomisk riktig køpris vil måtte endre seg nokså raskt og nokså kraftig når antall bilreiser øker eller når enten vegsystemet eller kollektivsystemet blir bedre. Vi kan altså ikke love noen stabil pris eller gi prisgarantier for lang tid framover.

Dokumentet om kollektivselskapets kostnader, optimalt kollektivtilbud og verdien av forbedringer behandler de viktigste kostnadene ved å drive en kollektivlinje eller en samling av linjer, og viser hvordan selskapet kan tilpasse seg optimalt. Det viser seg blant annet at gitt samme billettpris er det ikke er noen forskjell mellom samfunnsøkonomisk og bedriftsøkonomisk beste tilpasning når det gjelder frekvens, flatedekning og kjøretøykapasitet. Forskjellen ligger i billettprisen, som blir for høy når best mulig bedriftsøkonomi er målet. Løsningen er et pristak. Dokumentet er for øvrig på hele 43 sider, og dekker flere problemer.

### Om ulike enkeltvirkninger

De to første dokumentene under dette temaet drøfter to slags køproblemer. Det første gjelder hastighetsforskjeller, altså at et saktegående kjøretøy samler opp en kø bak seg. Det andre gjelder hendelser som innskrenker kapasiteten i en kortere tid. Dokumentene er såpass presise i beskrivelsen av det som skjer at en beregning av nyttetapet for trafikantene er mulig.

Arbeidsdokumentet om statistiske stordriftsfordeler har mange anvendelsesområder. Det dreier seg om at behovet for reservekapasitet blir relativt mindre jo større skala virksomheten drives i. To busselskaper som deler markedet i en by, trenger flere busser og flere sjåfører enn om de slo seg sammen. En bydel som har hjemmesjukepleierne fordelt på flere roder, må ansette flere sjukepleiere enn en bydel som har samlet reservekapasiteten til en felles styrke. Dette prinsippet har overraskende mange anvendelser. (Det finnes imidlertid også motvirkende forhold, for eksempel at et sjukdomstilfelle blant reservepleierne kan smitte flere når de er samlet på samme sted, eller at reisevegen for sjåførene blir lengre.)

# 2 Samfunnsøkonomiske analyser og prosjektvalg

## 2.1 Rammeverk for nyttekostnadsanalyse og prosjektvalg[1]

[1] Dette er arbeidsdokument TØ/2156/2009, som er den endelige versjonen av det som i mange år har vært (eller var?) grunnlaget for nyttekostnadsberegninger med transportmodell i Norge.

# Rammeverk for nyttekostnadsanalyse og finansieringsanalyse

## Innhold

# 1. Innledning

Hensikten med dette arbeidsdokumentet er å gi et mest mulig riktig, men enkelt opplegg for nyttekostnadsanalyse og finansieringsanalyse på grunnlag av transportmodeller, enten det dreier seg om store modellsystemer eller enkle stiliserte modeller. Opplegget kan programmeres opp eller brukes som grunnlag for håndregning.

# 2. Hovedelementene i det samfunnsøkonomiske regnestykket[2]

All nytte og alle kostnader tilfaller en av fire sektorer. De fire sektorene er brukerne (trafikantene), produsentene (operatørene), det offentlige og samfunnet for øvrig. Vi kaller brukernytta B, produsentenes overskudd (operatørnytta) P, virkningen på offentlige budsjetter F og virkningene for samfunnet for øvrig E. Nytte og kostnader som tilfaller det offentlige multipliseres med 1 + S, skattekostnaden. Årlig netto nytte i år *n* er:

(1) $$V^n = B^n + P^n - (1+S)F^n + E^n$$

Operatørnytte består av inntekt J minus kostnad C og minus private infrastrukturinvesteringer og andre private infrastrukturkostnader K. Tilskudd T minus overføringer til staten Y kommer i tillegg:

(2) $$P^n = J^n - C^n - K^n + T^n - Y^n$$

Det offentliges finansieringsbehov (underskudd) består av offentlig kjøp av transporttjenester T, offentlige investeringer I og offentlige etaters driftskostnader D, minus overføringer fra bom- og parkeringsselskaper og andre inntektskilder, Y, og skatteinntekter, R:

(3) $$F^n = -Y^n - R^n + T^n + I^n + D^n$$

Nytte for samfunnet forøvrig består av sparte ulykkeskostnader U og sparte miljøkostnader M pluss annen nytte A. Den sistnevnte kategorien er satt inn for å ha et sted å føre inn tilfeldige effekter og justeringer i programmet. Dessuten fører vi Z, skrapverdien av investeringene ved utløpet av perioden, her.

(4) $$E^n = U^n + M^n + A^n + Z^n$$

Alle variable i likning (1) til (4) gjelder årlige verdier. For hver av disse variablene definerer vi nåverdien over alle N+1 år slik: La X være et hvilket som helst av variablene A, B, C, D, E, F, I, J, K, M, P, R, T, U, V, eller Y – med eller uten indeks n. For alle slike variable gjelder:

(5) $$X = \sum_{n=0}^{N} (1+r)^{-n} X^n$$

der r er kalkulasjonsrenta. År 1 er "startåret", det første året med trafikk. Ofte vil et tiltak medføre en investeringskostnad som vi regner faller i år 0. Derfor har vi tatt med år 0. For Z gjelder:

[2] For fullstendig liste over variablene, samt en tabell over hvilke virkninger som føres på hver av sektorene, se Minken og Samstad (2005).

(6) $$Z = (1+r)^{-N} Z^N$$

Skattefaktoren S er alltid den samme og skal ikke neddiskonteres. Naturligvis har vi nå:

(7) $$V = B + P - (1+S)F + E$$

(8) $$P = J - C - K + T - Y$$

(9) $$F = -Y - R + T + I + D$$

(10) $$E = U + M + A + Z$$

Endringen i virkemiddelbruk fra situasjon 0 til situasjon 1 kalles *tiltaket*. Netto nåverdi av tiltaket, eller kort sagt nåverdien, er V. Netto nytte pr. budsjettkrone, eller kort sagt nyttekostnadsbrøken, er $VF^{-1}$. Da har vi forutsatt at det er F, dvs. alle inn- og utbetalinger over offentlige kasser, som skal stå under brøkstreken (hvilket bl.a. innebærer at momsinntektene på investering og drift kommer inn, slik at investering og drift faktisk blir uten moms under brøkstreken). Hvis det er *etats*budsjettet som er den begrensende faktoren, stiller det seg annerledes.

Kaller vi nåverdien NNV og nyttekostnadsbrøken NNB (netto nytte pr. budsjettkrone), har vi altså under disse forutsetningene:

(11) $$NNV = V$$ $$NNB = VF^{-1}$$

## 2.1 Finansieringsforutsetninger

Anta vi skal nytteberegne en bypakke. To ulike finansieringsforutsetninger må slå til dersom pakka skal kunne realiseres etter intensjonen. For det første må kollektivselskapene kunne oppnå overskudd etter tilskudd, og for det andre må investeringsprogrammet kunne realiseres med de midlene som blir til overs etter at denne forutsetningen er oppfylt. Flere forhold kan bidra til å rokke ved finansieringsforutsetningene: Investeringene kan bli dyrere enn planlagt, inntektene fra brukerbetalingen kan bli mindre enn planlagt, kollektivselskapenes kostnader kan bli høyere enn planlagt, eller statlige og kommunale bevilgninger kan utvikle seg annerledes enn planlagt.

Både i denne sammenhengen og i andre sammenhenger der brukerfinansiering er forutsatt som en del av tiltaket, er det naturlig å regne statlige og kommunale bevilgninger som en gitt størrelse. Det innebærer at underskuddet på offentlige budsjetter maksimalt kan være et gitt beløp $F_0$, altså $F \le F_0$ eller $F - F_0 \le 0$. La $F_B = F - F_0$ være underskuddet i bypakka eller det brukerfinansierte tiltaket etter at statlige og kommunale midler er mottatt. Kravet om at brukerfinansieringen skal dekke de kostnadene ved tiltaket som ikke dekkes av offentlige midler, kan altså skrives $F_B \le 0$.

Operatørselskapene er av fire slag: kollektivselskaper, bomselskaper, parkeringsselskaper og OPS-selskaper. Bruker vi henholdsvis fotskrift K, B, P og O for å betegne overskudd, inntekter og kostnader for hver av dem, har vi åpenbart at $P = P_K + P_B + P_P + P_O$. Kravet om at kollektivselskapet ikke skal gå med underskudd etter tilskudd skriver vi $P_K \ge 0$.

De to finansieringsforutsetningene vi har omtalt er altså:

(12) $$F_B \le 0,\ P_K \ge 0$$

I et slikt kontosystem som vi bygger nyttekostnadsanalysen på her, vil det framgå av nytteberegningene om disse to forutsetningene er innfridd.

Nå er det klart at vi alltid kan gjøre $P_K \geq 0$ ved å øke overføringene fra det offentlige, $T_K$, tilstrekkelig. Vi står derfor overfor et valg: Skal vi behandle $T_K$ som gitt eller som en variabel som kan brukes til å sikre overskudd for kollektivselskapene? På den ene sida er tilskudd til kollektivtrafikken et eget virkemiddel i bypakkene, med egne økonomiske rammer. På den andre sida er det enklest å forutsette at det planlagte kollektivtilbudet blir finansiert uansett, slik at vi ikke må gå ekstra planleggingsrunder med nedskjæring av kollektivtilbudet før vi kan si at $P_K \geq 0$ er oppfylt.

Vårt valg vil være å anta at $T_K - Y_K$ tilpasser seg slik at $P_K = 0$. Det betyr at kollektivtilbudet i praksis er det offentliges ansvar. Trolig kan vi også anta at bompengeselskapet og parkeringsselskapene kan regnes som en del av det offentlige. Vi eliminerer $T - Y$ fra (8) og (9). Videre skjønner vi at inntektene fra drivstoffavgifter m.m., $R$, ikke tilfaller prosjektet. Det gjør derimot inntektene $R^+$ fra en eventuell lokal drivstoffavgift.

Når det offentlige har ansvaret for kollektivtransporten, har vi ikke bruk for å skille mellom offentlige og private investeringer. Om det ikke finns noe OPS-selskap, kan vi altså sette $K = 0$. Den gjenstående finansieringsbetingelsen for bypakka eller det brukerfinansierte prosjektet kan da skrives:

(13) $$F_B = (I + D - R^+) + (C - J) - F_0 \leq 0$$

Likning (13) er intuitivt riktig: Kostnaden til investering og drift av veg- og kollektivsystemet må være mindre enn lokale avgifter, billettinntekter, bominntekter, (offentlige) parkeringsinntekter og statlige og kommunale tilskudd.

Resten av det offentlige har i dette tilfellet underskuddet $F_0 - R$. Det kan argumenteres for at det bare er denne delen av det offentliges budsjett som skal belastes med skyggeprisen ($1 + S$). Gitt at (13) er oppfylt, vil det jo bare være denne delen som påvirker skattleggingsbehovet. Hvis det er et akseptabelt argument, har det store konsekvenser for beregning av optimal bomavgift, som er svært følsom for skyggeprisen på offentlige midler.

# 3. Elementene i trafikantnytten, $B^n$

Et reisemarked er kjennetegnet ved et startsted, et bestemmelsessted, en reisemåte (bil, kollektivt), den perioden på dagen det reises i (rush, utenom rush), og reisehensikten. Av grunner som vi kommer tilbake til, vil vi ikke skille mellom reisehensikter på annen måte enn at vi regner med en annen miks av reisehensikter i rush enn utenom rush, og dermed ulik gjennomsnittlig tidsverdi i de to periodene. Indekserer vi startsted med $i$, bestemmelsessted med $j$, reisemåte med $m$, og periode på dagen med $k$, vil mengden av alle reisemarkeder $W$ bestå av alle mulige vektorer $w = (i, j, m, k)$. Den generaliserte reisekostnaden i marked $w$, $g_w$, består av tre deler: kjørekostnaden $p_w$, tidskostnaden $\omega t_w$ og et direkte pengeutlegg $b_w$:

(14) $$g_w = p_w + \omega_k t_w + b_w$$

I (14) er $\omega_k$ tidsverdien og $t$ reisetida. For kollektive reisemarkeder vil $p_w$ være null, og for bilreisemarkeder som ikke krysser noe bomsnitt eller involverer parkeringskostnader vil $b_w$ være null.

Vi skal merke oss tre viktige fakta om $g_w$ i (14).

1. $g_w$ representerer reisekostnadene etter at trafikanten har valgt rute, dersom hun er bilist, eller kollektivlinje, dersom hun er kollektivreisende. Vi forutsetter at trafikantene velger den ruta som gir minst $g_w$, gitt de framkommelighetsforhold som eksisterer når de andre trafikantene har valgt sine ruter og linjer. Vi forutsetter m.a.o. at det eksisterer en Nash-likevekt i transportsystemet – en såkalt brukerlikevekt.

   Sjøl om $g_w$ vil være entydig bestemt når trafikksystemet er i likevekt, vil ikke nødvendigvis oppdelingen på de tre delene være entydig. Når det er køer, kan nemlig likevekten innebære at det finnes mer enn en rute som har kostnadsminimal $g_w$. En av dem kan for eksempel være kort men langsom, mens den andre er lang men rask. Den resulterende tvetydigheten i (14) kan bare løses dersom vi splitter opp reisemarkedet på ruter som er i bruk og definerer (14) rutevis. La oss se bort fra det praktiske problemet med å finne alle ruter som er i bruk, og bare forutsette at indeksen $w$ også skiller mellom ruter der hvor det finnes flere ruter i bruk i reisemarkedet $(i, j, m, k)$. Det er viktig for oss å beregne samfunnsøkonomisk lønnsomhet på en måte som gjør tvetydigheten i hvordan $g_w$ er sammensatt, minst mulig.

2. $g_w$ vil kunne inneholde kostnader som trafikanten ikke bryr seg om når hun treffer sitt valg av bestemmelsessted og reisemåte. Typisk vil det gjelde slike deler av $p_w$ som olje- og dekkostnader, vedlikeholdskostnader og kapitalkostnader for kjøretøyet, osv. Ved beregning av brukernytte (trafikantnytte) er det opplevde kostnader, altså kostnaden eksklusive de ”ikke opplevde” elementene, som er relevant. Besparelser i ”ikke opplevde” eller ”ikke atferdsrelevante” kostnader vil bli å behandle på samme måte som eksterne kostnader *i nyttekostnadsanalysen.*[3]

3. $g_w$ vil på den andre sida inneholde elementer som er kostnader for trafikantene, men inntekter for kollektivselskapene eller det offentlige. Dette gjelder bompenger, parkeringsavgifter ut over den marginale kostnaden ved drift av parkeringsplassen, samt skattene i drivstoffprisen. Ved beregning av brukernytte (trafikantnytte) er det opplevde kostnader, altså kostnaden inklusive disse overføringene, som er relevant.

La nå $x_w$ være antall reiser i reisemarked $w$, og la toppskrift 0 og 1 på variablene $g_w$ og $x_w$ betegne henholdsvis variabelen med og uten tiltaket som vi skal beregne brukernytten av. La oss samtidig midlertidig sløyfe toppskriften $n$ for året beregningen gjelder. I henhold til trapesformelen er da brukernytten $B$:

[3] I Minken og Samstad (2005) er denne måten å behandle ikke opplevde kostnader på kalt korreksjoner. Både ikke opplevde kjørekostnader, eksterne kostnader, skatter og billettkostnader blir der betraktet som en form for korreksjoner av brukernytten.

At ikke atferdsrelevante kostnader kan behandles som eksterne kostnader i nyttekostnadsanalysen, betyr ikke at de skal behandles slik når optimale avgifter skal bestemmes. De vil jo ikke være påført andre enn den reisende sjøl.

$$(15)\qquad \begin{aligned} B &= \frac{1}{2}\sum_{w\in W}\left(g_w^0 - g_w^1\right)\left(x_w^0 + x_w^1\right) \\ &= \sum_{w\in W}\left(g_w^0 - g_w^1\right)x_w^0 + \frac{1}{2}\sum_{w\in W}\left(g_w^0 - g_w^1\right)\left(x_w^1 - x_w^0\right) \\ &= \frac{1}{2}\sum_{w\in W}\left(g_w^0 + g_w^1\right)\left(x_w^1 - x_w^0\right) + \sum_{w\in W}\left(g_w^0 x_w^0 - g_w^1 x_w^1\right) \end{aligned}$$

De tre linjene i formel (15) er tre alternative måter å skrive trapesformelen på. Første linje i (15) representerer trapesformelen på normal form. Andre linje skiller mellom nytte for eksisterende trafikanter (første leddsum) og nytten av nyskapt og overført trafikk (andre leddsum). Med 'eksisterende trafikanter' mener vi de som ikke endrer atferd som følge av tiltaket, mens den nyskapte og overførte trafikken åpenbart er de som endrer atferd på grunn av tiltaket.

I den tredje linja i (15) er trapesformelen skrevet på en form som egner seg som utgangspunkt for korrigeringer for reelle kostnader som trafikantene likevel ikke har tatt hensyn til i sine beslutninger. Den siste leddsummen her er opplevde kostnader i førsituasjonen minus opplevde kostnader i situasjonen *med* tiltaket. Her er det altså bare å føye til kostnader i før- og ettersituasjonen som trafikantene ikke har tatt hensyn til, eller trekke fra opplevde kostnader som ikke er reelle ut fra et samfunnsøkonomisk synspunkt. Det er dette vi mener med korrigeringer i Minken og Samstad (2005).

Alle elementene i alle tre linjer, unntatt den siste leddsummen i tredje linje, involverer ledd der kostnader i den ene situasjonen skal multipliseres med volumer fra den andre situasjonen. Da er det umulig å beregne disse elementene bare ved å granske data om de to situasjonene hver for seg. Det er derimot mulig når det gjelder den andre summen i tredje linje, som består av totale opplevde kostnader uten tiltaket minus totale opplevde kostnader med tiltaket.

Summen av totale opplevde kostnader i en situasjon er – med to unntak! – den samme enten vi summerer over reisemarkeder, ruter eller lenker i nettverket. Å summere over lenkene er overlegent det enkleste, forutsatt at vi ikke skiller mellom ulike reisehensikter innen samme periode på dagen. De to unntakene gjelder kostnader som knytter seg til reisa uansett hvilke lenker eller ruter som brukes, dvs. parkeringskostnader og kollektivbilletten (i den grad den er uavhengig av hvilken linje en velger). Dermed blir det klart at brukernytten kan beregnes ved å legge sammen tre separate beregningsresultater:

1. Den første summen i tredje linje i (15), som kan beregnes ved å multiplisere en turmatrise-differanse med en kostnadsmatrisedifferanse og ta summen av diagonalelementene,

2. De lenkebaserte kostnadene i den andre summen, og

3. Billettkostnadene og parkeringskostnadene, som kan beregnes på tilsvarende måte som den første summen i tredje linje i (15), men med andre data.

De direkte pengeutleggene, $b_w$ i likning (14), består av bompenger, billettkostnad og parkeringskostnad. Bompengene fanger vi opp ved en lenkebasert beregning. Parkeringskostnadene i sone $j$ i periode $k$ på dagen kaller vi $b_{jmk}$, der $m$ er med fordi parkeringskostnader bare er aktuelt for bil. Billettprisen ved en reise fra $i$ til $j$ med reisemåte $m$ på tidspunkt $k$ kaller vi $f_w$, der $m$ er med fordi billetter bare er aktuelt ved kollektivreiser, og $k$ er med fordi det kan være aktuelt å skille mellom kollektivprisen i rush og utenom rush.

Kall mengden av lenker i nettverket for reisemåte $m$ for $A_m^0$ uten tiltaket og $A_m^1$ med tiltaket, og indekser lenkene i nettverket med $a$. Trafikkvolumet på lenke $a \in A_m$ i tidsrom $k$ er $v_{ak}$, og transporttida på lenke $a \in A_m$ i tidsrom $k$ er en funksjon av trafikkvolumet: $t_{ak} = t_a\left(v_{ak}\right)$. De opplevde kjørekostnadene på lenke $a \in A_m$ er $p_a$. Bompengesatsen på lenke $a \in A_m$ er $b_{ak}$. Vi kan da beskrive den andre summen i tredje linje i (15) mer nøyaktig slik:

$$
\begin{aligned}
&\sum_{w \in W} \left( g_w^0 x_w^0 - g_w^1 x_w^1 \right) \\
&= \sum_{m,k} \omega_k \left( \sum_{a \in A_m^0} v_{ak}^0 t_a^0 \left( v_{ak}^0 \right) - \sum_{a \in A_m^1} v_{ak}^1 t_a^1 \left( v_{ak}^1 \right) \right) \\
&+ \sum_{m,k} \left( \sum_{a \in A_m^0} p_a^0 v_{ak}^0 - \sum_{a \in A_m^1} p_a^1 v_{ak}^1 \right) \\
&+ \sum_{m,k} \left( \sum_{a \in A_m^0} b_{ak}^0 v_{ak}^0 - \sum_{a \in A_m^1} b_{ak}^1 v_{ak}^1 \right) \\
&+ \sum_{w \in W} \delta_{jmk}^w \left( b_{jmk}^0 x_w^0 - b_{jmk}^1 x_w^1 \right) \\
&+ \sum_{w \in W} \left( f_w^0 x_w^0 - f_w^1 x_w^1 \right)
\end{aligned} \tag{16}
$$

Andre linje i (16) er tidsbesparelsen ved tiltaket. Legg merke til at volume-delayfunksjonen $t_a$ kan ha endret seg som følge av tiltaket. Tredje linje er kjørekostnadsbesparelsen. Fjerde linje er endringen i bompengekostnadene (= bompengeinntekten for bomselskapet). Femte linje er parkeringskostnaden (hvorav en del er inntekt for parkeringsselskapet). Kroneckerdeltaket er 1 hvis *jmk* er en komponent i *w*, 0 ellers. Siste linje er billettkostnaden, som naturligvis også er billettinntektene for kollektivselskapene.

Styrken ved å bruke den tredje linja som grunnlag for beregningene er at vi ved denne oppdelingen får ut de elementene som bare er overføringer, og som derfor må føres med motsatt fortegn andre steder i regnestykket, dvs. bompengebetalingen, billettbetalingen, parkeringsavgiftene og skatt på drivstoff.

En stor styrke ved opplegget i *andre* linje i (15) er at de enkelte delene av trafikantenes nytte kan analyseres. Vi ser at en endring i bompengene, for eksempel, kan dekomponeres i tidsgevinsten for de som fortsatt velger å kjøre, minus bompengene, og minus nyttetapet for de som prises av vegen (dvs. den andre summen i andre linje i (15)).

Svakhetene med vår brukernytteberegning, uansett hvilken linje i (15) vi bruker, er at vi må operere med en gjennomsnittlig tidsverdi for hver av periodene på dagen, og kanskje at vi ikke tar høyde for at kjørekostnadene kan være en funksjon av trafikken på lenkene.

## 3.1 Provenyet av drivstoffavgiftene

Når det gjelder skatten, må vi da dele den samlede kjørekostnaden i en ressurskostnadsdel og en skattedel. Skattedelen tas til inntekt for det offentlige etter en korreksjon som skal ta hensyn til at når forbrukerne bruker mer drivstoff, slik at staten får større inntekt av avgifter og moms på drivstoff, må de også kjøpe mindre andre varer, slik at staten taper momsinntekter på andre varer.

La *R* være det offentliges inntekt av skatt og avgift på drivstoff i år *n*, *q* være drivstoffprisen i kroner pr. liter, *z* drivstoffeffektiviteten i kilometer pr. liter, og *d* det totale antall kjørte kilometer i år *n*. Drivstoffkostnadene i et alternativ er da $dqz^{-1}$. (Til sammenlikning: I tredje linje i (16) omfatter $p_a^0$ og $p_a^1$ implisitt literprisen multiplisert med lengden av lenke *a* i kilometer og delt på drivstoffeffektiviteten, og det er endringen fra nullalternativet som er beregnet.) Anta at literprisen *q* består av en ressurskostnad $q_0$ pluss en skattedel $q_s$, $q = q_0 + q_s$. Skattedelen av kjørekostnadene er da $dq_s z^{-1}$. La *m* være gjennomsnittsmoms på annet forbruk. Inntekten av skatt og avgift på drivstoff for det offentlige er da:

(17)
$$R = \frac{q_s - mq_0}{1+m} \cdot \frac{d}{z}$$

Denne formelen er lik formel (4.1) i Minken og Samstad (2005), men med litt annen notasjon og mer eksplisitt angivelse av antall liter forbrukt.

### 3.2 Eksterne kostnader

Endelig vil de samlede kjørekostnadene gi et enkelt grunnlag for beregning av eksterne kostnader, dersom vi forenklet regner disse som en form for kilometeravhengige kostnader. Det er bare å multiplisere kjørekostnadene med en faktor lik forholdet mellom vedkommende eksterne kostnad pr. kilometer og drivstoffkostnaden pr. kilometer. På denne måten behandler vi ulykkeskostnadene, støykostnadene, utslippskostnadene og kostnaden ved forbruket av ikke opplevde ressurser pr. kilometer, som olje og dekk, reparasjoner og service og kilometerdelen av bilenes kapitalkostnader.

## 4. Operatørnytte *P* og det offentliges nytte *F*

Vi har allerede dekket inntektssida for kollektivselskapene, parkeringsselskapene og bomselskapet (hhv. linje 6, 5 og 4 i (16), med motsatt fortegn). Til sammen utgjør de *J*. Vi har også behandlet *R*, inntektsprovenyet fra drivstoffavgiftene. Gjenstår *K* + *C*, dvs. kollektivselskapenes kostnader, bomselskapenes innkrevingskostnader og parkeringsselskapenes kostnader, i den grad det sistnevnte er nødvendig. Gjenstår også det offentliges investerings- og driftskostnader, *I* og *D*.

### 4.1 Kollektivselskapenes kostnader

Som når det gjelder overskuddet *P*, kan også operatørenes kostnader *C* deles i kostnadene til kollektivselskapene, $C_K$, og kostnadene til henholdsvis bomselskapet, $C_B$, parkeringsselskapet, $C_P$, og eventuelt OPS-selskapene, $C_O$. Vi behandler her $C_K$.

Kollektivtilbudet er oppdelt i et antall kollektivlinjer. Vi antar at hver linje bruker rullende materiell og mannskap som er dedikert til denne linja. Driften på linja består av rundturer. Kollektivselskapets kostnader er summen av kostnadene på linjene (vi ser bort fra at felleskostnadene kan variere med tilbudet).

I Minken (2009), kapittel 2, er kostnadsfunksjonen for en linje utledet under enkle, men ikke urealistiske forutsetninger. Toperiodetilfellet er behandlet i kapittel 7 i samme dokument. Dette arbeidsdokumentet er lagt ved rapporten om samfunnsøkonomiske analyser i konseptvalgsutredninger (Minken m.fl. 2009) som vedlegg 5. Kollektivselskapets kostnader kan beregnes/programmeres med utgangspunkt i det dokumentet.

I henhold til Minken (2009) kan kostnaden pr. driftstime for en samling av *N* like kollektivlinjer skrives

(18) $$C = C_1 + C_2 = \left[\left(r_0/h + w\ell_0\right) + g_0 s\right] N \frac{af}{s} + \left[r_1/h + g_1 s\right]\varphi^{-1}\frac{m}{s}x$$

Her er *a* antall kilometer pr. rundtur, *s* gjennomsnitthastigheten pr. rundtur inklusive snutid og stopp, *f* frekvensen (avganger pr. time pr. linje), *m* gjennomsnittslengda pr. reise, *x* den samlede etterspørselen pr. time for alle *N* linjer, $\varphi$ forholdstallet mellom gjennomsnittsbelegg og belegget over dimensjonerende snitt, *h* antall driftstimer pr. år, *w* timelønnskostnaden, og $\ell_0$ mannskapsbehovet pr. buss/togsett. $r = r_0 + r_1 c$ er kapitalkostnaden pr. år for en buss/et togsett, oppdelt i en kostnad $r_0$ for minste realistiske kapasitet pr. buss/togsett og et tillegg $r_1$ for hver kapasitetsenhet (passasjerplass) ut over det. Den tilsvarende oppdelingen av de kilometer-avhengige kostnadene er $g = g_0 + g_1 c$.

Kapasiteten pr. avgang, *c*, framkommer ikke direkte av formelen. Det gjør heller ikke materiellbehovet. Imidlertid er materiellbehovet (antall busser eller togsett i drift) lik $afs^{-1}$, de kilometeravhengige kostnadene pr. år *hafg* og de tidsavhengige kostnadene pr. år $(r + hw\ell_0)afs^{-1}$. Utkjørt distanse pr. år er *haf*. Dette er indikatorer som det kan være hensikts-messig å rapportere i et dataprogram. Utkjørt distanse og energidelen av de kilometeravhengige kostnadene gir også grunnlag for å beregne eksterne virkninger.

Når tilbudet varierer over døgnet, kan vi under visse vilkår beregne de årlige kostnadene separat for hver type av driftstime. Anta for eksempel at det er *h* driftstimer pr. år, fordelt med $h_H$ høybelastningstimer og $h_L$ lavbelastningstimer. Frekvensen, målt i antall avganger pr. time, er $f_H$ i høybelastningsperioden og $f_L$ i lavbelastningsperioden. Vi kan skille mellom tre tilfeller, alle behandlet i Minken (2009). Det første er når det finnes et grunntilbud som går hele drifts-døgnet, pluss ekstraavganger i rush. Det andre er når det er samme trekkenhet som brukes hele tida, men kapasiteten kan tilpasses ved at vogner koples av og på. Det tredje er når kapasiteten pr. avgang er lik både i høy- og lavbelastningsperioden. Tilfellene der man ikke kan beregne kostnadene i de to periodene separat, knytter seg til grunntilbud pluss ekstraavganger dersom en bestemt bibetingelse er bindende i optimum, og til situasjonen med samme kapasitet pr. avgang hele tida, dersom en annen bibetingelse er bindende.

Det er viktig at kapitalkostnader for rullende materiell bare, eller i det alt vesentlige, påløper i høybelastningsperioden, ettersom høybelastningsperioden er den som bestemmer kapasitets-behovet. Det er også viktig at klargjøringskostnaden, som påløper en gang pr. dag, utelukkende legges til kapitalkostnaden i den mest belastede perioden, ellers oppstår dobbeltføringer. Når det gjelder vedlikeholdet, fordeles kostnadene til det på kilometeravhengige og tidsavhengige kostnader. De kilometeravhengige vedlikeholdskostnadene er et tillegg til energikostnadene, som i Minken (2009) er betegnet med variabelen *g*. De tidsavhengige vedlikeholdskostnadene regnes som et tillegg til kapitalkostnadene, og påløper derfor i høybelastningsperioden.

Samlede kostnader for kollektivtransporten, $C_K$, framkommer ved å summere over alle linjer og driftsperioder. Hvis **c** er en vektor med kapasitetene pr avgang i hver periode på hver linje og $\mathbf{f}_H$ og $\mathbf{f}_L$ er tilsvarende vektorer for frekvensene (men her oppdelt i høy- og lavtrafikk-perioder), kan vi forenklet skrive $C_K = C_K(\mathbf{c}, \mathbf{f}_H, \mathbf{f}_L)$.

For hver enkelt kollektivlinje eller samling av likeartede linjer kan det være av interesse å kontrollere at tilbudet i den mest belastede timen ikke er for lite til å dekke etterspørselen. Dersom det skal unngås, må følgende ulikhet gjelde:

(19) $$\varphi cf \geq \frac{m}{a}\frac{x}{N}$$

At (19) er oppfylt med likhet er et vilkår for effektiv tjenesteproduksjon. Vi skal ikke kjøre rundt med overflødig kapasitet. For å bruke (19) til kontroll må man ha et uavhengig anslag på kapasiteten pr. avgang, *c*. Formel (18) kan ikke brukes til kontroll, for den bygger på at (19) er oppfylt med likhet.

Likning (18) er likevel brukbar uansett om kollektivselskapet har tilpasset frekvens og flatedekning optimalt eller ikke. De etterfølgende kapitlene i Minken (2009) gir formler som kan brukes til å beregne kostnadene når selskapet tilpasser seg optimalt, eventuelt under bibetingelser og restriksjoner.

## 4.2 Det offentliges kostnader

Investeringskostnaden *I* er gitt i og med beskrivelsen av tiltaket. Verre da med vedlikeholds-kostnadene *D*, men i mangel av en egen modell til å beregne dem, vil vi akseptere anslag som gjøres i EFFEKT e.l.

# 5. Det fullstendige samfunnsøkonomiske regnestykket

## 5.1 Årlig nytte i år n

Vi husker at årlig nytte i år n er $V^n$:

(20) $$V^n = B^n + P^n - (1+S)F^n + E^n$$

**Brukernytte**

La antall timer i året av type *k* være $h_k$. Brukernytten, basert på tredje linje i likning (15), er da:

$$B^n = \sum_i \sum_k h_k B^n_{ik}$$

der

$$B^n_{1k} = \frac{1}{2}\sum_{ijm}\left(g^{n0}_{ijmk} + g^{n1}_{ijmk}\right)\left(x^{n1}_{ijmk} - x^{n0}_{ijmk}\right)$$

$$B^n_{2k} = \sum_{ijm}\left(b^{n0}_{jmk}x^{n0}_{ijmk} - b^{n1}_{jmk}x^{n1}_{ijmk}\right)$$

$$B^n_{3k} = \sum_{ijm}\left(f^{n0}_{ijmk}x^{n0}_{ijmk} - f^{n1}_{ijmk}x^{n1}_{ijmk}\right)$$

$$B^n_{4k} = \sum_m \omega^n_k\left(\sum_{a\in A^{n0}_m} v^{n0}_{ak}t^{n0}_a\left(v^{n0}_{ak}\right) - \sum_{a\in A^{n1}_m} v^{n1}_{ak}t^{n1}_a\left(v^{n1}_{ak}\right)\right)$$

$$B^n_{5k} = \sum_m\left(\sum_{a\in A^{n0}_m} p^{n0}_a v^{n0}_{ak} - \sum_{a\in A^{n1}_m} p^{n1}_a v^{n1}_{ak}\right)$$

$$B^n_{6k} = \sum_m\left(\sum_{a\in A^{n0}_m} b^{n0}_{ak} v^{n0}_{ak} - \sum_{a\in A^{n1}_m} b^{n1}_{ak} v^{n1}_{ak}\right)$$

(21)

I (21) er $B_1$ det eneste elementet som trenger en matrise med realiserte generaliserte kostnader etter rutevalget som input. $B_2$ er parkeringskostnader og $B_3$ er billettkostnader (differansen mellom totale utlegg før og etter tiltaket). Disse to elementene er også basert på matriser. $B_4$ er tidskostnader, $B_5$ er kjørekostnader og $B_6$ er bompengebetaling (også dette differansen mellom før og etter). Disse elementene er basert på lenkedata.

Det er hensiktsmessig for å kunne beregne inntektene for operatørselskapene og skatteinntektene for det offentlige at en kan rapportere de seks elementene hver for seg i tillegg til summen av dem, som jo er trafikantnytten. På den måten trengs det ikke noen egen kollektivnyttemodul, bortsett fra til kollektivselskapets kostnader.

Samtidig er det også av og til nyttig å beregne brukernytten på grunnlag av *andre* linje i likning (15). I det tilfellet er det den andre leddsummen (trekanten i trapeset) som krever en matriseberegning, mens den første leddsummen (nytten for eksisterende trafikken) i prinsippet kan beregnes lenkebasert. Ved å bygge på andre linje kan en dekomponere den første leddsummen i spart tid, sparte kjørekostnader og sparte bomavgifter. Andre leddsum representerer nytten eller nyttetapet til dem som velger å endre atferd. Ved analyse av køprising kan en slik inndeling i fire komponenter – tid, kjørekostnad, bomavgift/billett og tap for de som endrer atferd – være spesielt opplysende, som vist i Eliasson og Mattson (2006).

Hvis en gjør beregningen også av den første leddsummen i andre linje på grunnlag av tur- og kostnadsmatriser, vil en kunne gjøre enkle analyser av den geografiske fordelingsvirkningen av et tiltak. Vi anbefaler altså at både den andre og den tredje linja i (15) programmeres, og at programmering av den andre linja baseres helt ut på matriser.

### Nytte for det offentlige

I tråd med avsnitt 3.1 og likning (13) inkluderer vi operatørnytten under det offentliges nytte. Operatørnytten er summen av nytten for kollektivselskapet, bomselskapet og parkeringsselskapet (vi regner med ett selskap av hver type):

(22) $$P^n = J^n - C^n = P_K^n + P_B^n + P_P^n$$

Anta det er Q kollektivlinjer indeksert med s. *Kollektivselskapets* nytte er

(23) $$P_K^n = -\sum_k h_k B_{3k}^n - C_K^n\left(\mathbf{c}, \mathbf{f_H}, \mathbf{f_L}\right)$$

der $B_{3k}^n$ er fra (21) og $C_K{}^n(\mathbf{c}, \mathbf{f}_H, \mathbf{f}_L)$ er beskrevet i avsnitt 5.1. I tillegg til denne kostnaden ved å drive kollektivsystemet, kan det beregnes et påslag pr. år for reserver av rullende materiell. Størrelsen på reservene av en spesiell type materiell avhenger av hvor mange kjøretøy (busser, togsett) av denne typen som totalt er i bruk i systemet, se Minken og Samstad (2005).

*Bomselskapets nytte* er

(24) $$P_B^n = -\sum_k h_k B_{6k}^n - C_B^n$$

der $B_{6k}^n$ er fra (21) og $C_B$ avhenger av antall bompasseringer og antall filer med bompengeinnkreving, pluss sjølsagt av innkrevingsteknologien.

*Parkeringsselskapets nytte* er

(25) $$P_P^n = -\sum_k h_k B_{2k}^n - C_P^n$$

der $B^n_{2k}$ er fra (21) og $C_P$ avhenger av antall parkeringsplasser i hver sone. Parkeringsplassene i sonene kan ha ulik driftskostnad.

Likning (22) til (25) er beregninger som gjøres for den enkelte kjøring av transportmodellen, og er altså ikke *differansen* mellom tiltaksalternativet og nullalternativet, slik som likning (21). Ved den endelige nytteberegningen av $F^n$ må kostnaden i nullaternativet (toppskrift 0) trekkes fra kostnaden i tiltaksalternativet (toppskrift 1). I tråd med (13), der $K$ settes lik 0, har vi da

(26)
$$F^n = \left[(I^{n1} + D^{n1} - R^{n1}) - P^{n1}\right] - \left[(I^{n0} + D^{n0} - R^{n0}) - P^{n0}\right]$$

der $I^n$ og $D^n$ er input til beregningene, og $R^n$ følger av likning (17), der parametrene $p_s$, $p_0$, $m$ og $s$ kan være spesifikke for år n, og kanskje til og med for tiltaket, og det totale antall utkjørte kilometer, $d$, er en indikator som må tas ut av transportmodellkjøringen.

**Nytte for samfunnet for øvrig**

I likhet med skatten $R$ vil de eksterne kostnadene være proporsjonale med utkjørte kilometer $d$ i alternativet, skal vi anta. Proporsjonalitetsfaktoren $e$ er summen av ikke opplevde kjørekostnader pr. kilometer, ulykkeskostnaden pr. kilometer, støykostnaden pr. kilometer og utslippskostnaden pr. kilometer, og er input til beregningene. $e$ vil utvikle seg med teknologien og er derfor spesifikk for det enkelte år. Vi ignorerer annen nytte $A$, men føyer til restverdien Z, som bare er aktuell i det siste året. Det markerer vi med et delta som er 1 hvis n er det siste året, null ellers. Vi har:

(27)
$$E^n = E^{n1} - E^{n0} = e^{n1}d^{n1} - e^{n0}d^{n0} + \delta Z^n$$

## 5.2 Netto nytte

Hvis vi nå hadde beregninger for hvert år i analyseperioden, ville likning (5) og (6) vært tilstrekkelig til å beregne netto nytte over hele perioden. I praksis har vi bare beregninger for ett, to eller i høyden tre år, og må interpolere. Ved interpolasjonen er det aktuelt å ta hensyn til utviklingen i befolkningsmengde, inntekt, teknologi og utenfra gitt nasjonal skatte- og avgiftspolitikk på en forenklet måte. Det lar seg ikke gi noen fast regel for dette – beregningsprogrammet må gi åpning for å definere hvordan de beregnede årene skal regnes sammen.

# 6. Finansieringsanalyse

Beregningsprogrammet må kunne presentere beregningen av $F$, dvs. summen av årlige neddiskonterte ”likning (26)”-resultater, som en separat indikator. Vi legger til grunn at det offentlige kan låne og plassere midler i et perfekt kapitalmarked til rente 4.5 %. Tiltaket er da finansiert hvis og bare hvis $F_B \leq 0$.

# 7. Tidsverdien og kjørekostnadene over tid

Vi veit at tidsverdien vokser med inntekten. Alle anslag tilsier at tidsverdien er en funksjon $\omega_k = \omega_k(y)$ av gjennomsnittsinntekten i husholdene, og at elastisiteten av tidsverdien med hensyn på $y$ er mellom ½ og 1. Vi antar forsøksvis at elastisiteten er 2/3.

Det betyr at i om noen år vil tidsverdien ha vokst i henhold til denne formelen. Å se bort fra det vil bety å bryte med Finansdepartementets veileder i samfunnsøkonomiske analyser, som sier at det er forventningsverdier som skal anvendes.

På samme måte vil kjørekostnadene utvikle seg fra periode til periode. Det er ikke husholdningsinntektene som driver dette, men kjøretøyteknologien og avgiftene. Hvis antall kilometer pr. liter drivstoff øker, reduseres $p_w$ tilsvarende. Hvis derimot bensin- og diesel-avgifter øker, eller produksjonskostnaden øker, vil $p_w$ øke. Samstad m.fl. (2005) gir grunnlag for å framskrive drivstoffeffektiviteten, men drivstoffkostnaden pr. liter må anslås på annet vis.

# 8. Rammeverk for nyttekostnadsanalyse av utforskende beregninger

Utforskende beregninger benytter seg bare av modellkjøringer for et enkelt år. Det vi da trenger, er bare avsnitt 6.1 pluss en regel for hvordan investeringen $I^n$ og restverdien $Z^n$ skal føres. Setter vi $Z^n = 0$ og regner investeringens levetid til 40 år, kan vi bruke annuiteten over 40 år av investeringskostnaden $I$ i alternativet som vårt anslag på $I^n$. Vi har:

(28) $$I^n = I \frac{r}{1-(1+r)^{-40}}$$

Med $r = 0.045$ er $I^n$ lik $0.054*I$.

## Litteraturliste


Eliasson, J. and L.-G. Mattsson (2006) Equity effects of congestion pricing. Quantitative methodology and a case study for Stockholm. *Transportation Research A* **40**, 602-620.

Minken, H. (2009) Kollektivselskapets kostnader, optimalt kollektivtilbud og verdien av forbedringer. Arbeidsdokument ØL/2157/2009, TØI. (Dette arbeidsdokumentet er også tatt inn i andre kapittel i den foreliggende rapporten.)

Minken, H., Larsen, O.I., Braute, J.H., Berntsen, S. og Sunde, T. (2009). Konseptvalgsutredninger og samfunnsøkonomiske analyser. TØI-rapport 1011/2009.

Minken, H. og H. Samstad (2005) Nyttekostnadsanalyser i samferdselssektoren: Rammeverk for beregningene. TØI-rapport 798/2005.

Samstad, H., M. Killi og R. Hagman (2005) Nyttekostnadsanalyse i samferdselssektoren: Parametre, enhetspriser og indekser. TØI-rapport 797/2005.

## 2.2 Om noen tekniske forutsetninger[4]

[4] Dette er arbeidsdokument 51078 fra 2017.

# Om noen tekniske forutsetninger for de samfunnsøkonomiske beregningene i KS1 Oslonavet

## Innhold



# Innledning og sammenfatning

Utgangspunktet for dette arbeidsdokumentet er den pågående kvalitetssikringen av KVU for Oslonavet som Dovre og TØI utfører på oppdrag av Finansdepartementet.[5] Arbeidsdokumentet berører ikke direkte kvalitetssikringen av KVU-en, men snarere de grunnleggende reglene for samfunnsøkonomisk analyse og hvordan de tolkes ulikt i de ulike etatene. Det gjennomgående eksemplet vi drøfter, er likevel KVU og KS1 av Oslonavet.

I første kapittel drøfter vi begrepsbruk og forutsetninger i Jernbanedirektoratets metodehåndbok og Statens vegvesens veileder i konsekvensanalyse.[6] Vi finner at det eksisterer forskjeller i praktiseringen og tolkningen av en del begreper i de to håndbøkene, og at dette vanskeliggjør sammenlikninger på tvers av etatene. Vår vurdering er at disse forskjellene ikke alltid lar seg begrunne med at forholdene på de to områdene er ulike, men oftest skyldes ulik lesning av NOU 2012:16 og Finansdepartementets rundskriv R-109/2014. Men ikke bare etatene tolker ting forskjellig, på ett punkt finnes det faktisk også en viktig tolkningsforskjell mellom NOU-en og rundskrivet. Slik vi ser det, er det behov for at alle parter sammen setter seg ned og kommer fram til en enhetlig tolkning og praksis.

I den grad det ikke fører fram, er det behov for metoder for å korrigere for forskjellene og vurdere hvordan det påvirker resultatet. I kapittel 2 redegjør vi for hvordan man kan sammenlikne prosjekter med ulik levetid eller når prosjektene er beregnet med ulike analyseperioder. Hovedprinsippet er alltid å bruke samme analyseperiode på alle prosjektene som skal sammenliknes, med om man velger å sette analyseperioden lik den lengste levetida

[5] Dette oppdraget pågikk naturligvis i 2017, ikke nå.

[6] Det som kalles kapitler i dette arbeidsdokumentet, er naturligvis underkapitler i denne TØI-rapporten.

eller kortere, er mer et spørsmål om hvor langt fram i tid en kan si noe presist om nytte og kostnader. Formelen som vi utvikler, er uansett brukbar.

I kapittel 3 legger vi til grunn at det foreligger sammenliknbare nytteberegninger for et analyseår, men alle prosjektene som skal sammenliknes, bruker ikke samme analyseår, og prosjektene er behandlet forskjellig når nytten fra det enkelte året er ekstrapolert til hele analyseperioden. Vi utvikler vi en praktisk metode for å bedømme hvor stor del av forskjellen i brutto nytte mellom to eller flere prosjekter som skyldes forutsetningene for denne ekstrapoleringen. Beregningene kan gjennomføres med lommeregner eller i et lite regneark. Eksemplet som vi viser, gjelder KVU og KS1 av Oslonavet.

Både i kapittel 2 og kapittel 3 beholder vi altså de særegne forutsetningene for hvert prosjekt. Forskjellen er at i kapittel 2 setter dem inn i en felles ramme som gjør det mulig å sammenlikne dem likevel, mens vi i kapittel 3 nøyer oss med å undersøke hvordan ulikhetene i forutsetningene påvirker resultatet, uten egentlig å oppheve ulikhetene og etablere en gyldig ramme for sammenlikningen.

Det er ellers grunn til å merke seg hvor mye mer komplisert det blir å sammenlikne prosjekter når rundskriv R-109/2014 skal følges til punkt og prikke. Mye skyldes at analyseåret er utgangspunkt for rentenedtrappingen.

# 1. Analyseperiode, levetid, restverdi og kalkulasjonsrente

## 1.1 KS1 av Oslonavet: Våre valg

Alle priser er i 2017-kroner. Alle kostnader og nyttevirkninger i alle konsepter og varianter er henført til henføringsåret (diskonteringsåret) 2017.

Vi har brukt en analyseperiode på 40 år fra og med åpningsåret, dvs. første driftsår av tiltaket. Første driftsår (åpningsåret) er satt til 2030 for alle konsepter og varianter som er nytteberegnet. Analyseperioden er altså 2030-2070.

Vi har satt levetida lik analyseperioden, altså 2030-2070. Vi har dermed ingen restverdier. Nedtrappingen av diskonteringsrenta fra 4 til 3 prosent etter 40 år er foretatt med utgangspunkt i analyseåret 2017, slik at 3 prosent rente er anvendt på nytte og kostnader fra 2058 til 2070.

Vi har vurdert om konseptene og variantene vi har analysert, er så sammensatt at det ikke er naturlig å operere med en enkelt analyseperiode. Vi konkluderer med at sjøl om konseptene består av elementer som ikke vil bli realisert simultant, er det ikke behov for å plassere dem entydig ut i tid på det nåværende tidspunktet. De tunge investeringene i de fleste konseptene faller i årene mellom midten av 2020-årene og midten av 2030-årene. Dermed har vi valgt 2030 som åpningsår for alle konsepter og varianter. Et unntak gjelder noen av kjøringene som bare inneholder prisvirkemidler – de kan naturligvis realiseres nærmest når som helst.

## 1.2 Våre valg vurdert mot Finansdepartementets rundskriv, etatenes praksis og KVU

KVU bruker 2014 som prisår og 2022 som henføringsår. Det er en relativt liten justering som vi kommer tilbake til i siste kapittel av dette dokumentet. Her skal vi først behandle forskjeller i tolkninger og praksis når det gjelder viktigere ting som analyseperiode, levetid, restverdi og kalkulasjonsrente.

*Analyseperiode, levetid, åpningsår, oppstartsår og restverdi*

Lærebøker i bedriftsøkonomi framhever at når flere forskjellige tiltak skal sammenliknes med hverandre, er det nødvendig å bruke samme analyseperiode. Hvis man nemlig skal sammenlikne et kortvarig tiltak med et mer langvarig, vil avkastningen i det kortvarige tiltaket kunne investeres på nytt og gi en ekstra avkastning i tida fram til det langvarige tiltaket har virket ferdig. Denne gevinsten ser man feilaktig bort fra om man velger ulike analyseperioder for de to tiltakene.

Man må derfor enten ha en analyseperiode som er like lang som levetida til det langvarige prosjektet, eller man kan velge en felles, kortere analyseperiode, men da må man beregne realistiske restverdier av det (eller de) tiltakene som kan gi avkastning etter analyseperiodens utløp. Man kan også velge en lengre analyseperiode enn tiltakene som skal nytteberegnes. Når levetida er kortere enn analyseperioden, vil vi få en eller flere reinvesteringer innen analyseperiodens utløp, pluss eventuelt en restverdi dersom levetida etter siste reinvestering først utløper etter analyseperioden.[7] Uansett må man alltid bruke samme tidshorisont for at sammenlikningen mellom tiltakene skal bli økonomisk korrekt.

Ved samfunnsøkonomisk analyse vil en stor del av avkastningen ikke foreligge i form av penger som kan reinvesteres. Og sjøl den delen av avkastningen som har pengeform, vil i regelen være spredt på mange eiere, og dermed ikke være like lett å reinvestere. Men i svært mange tilfeller vil det være mulig å gjenta det kortvarige tiltaket og få samme avkastning en eller flere ganger til, inntil levetida for det langvarige prosjektet er utløpt. Også når det gjelder samfunnsøkonomiske analyser vil det derfor være viktig å bruke samme analyseperiode. 40 år er i den sammenhengen en rimelig avveining mellom hvor langt fram i tid det er mulig å vurdere bruken av tiltaket og hvor langt framover vi kan regne med at større investeringer i samferdselssektoren vil være brukbare, reint teknisk.

Analyseperioden for investeringsprosjekter i samferdselssektoren er fastsatt i Finansdepartementets rundskriv R-109/2014 til 40 år. Imidlertid praktiseres det ikke på samme måte i vegvesenet og i jernbanesektoren. Vegvesenet bruker en analyseperiode på 40 år fra og med *åpningsåret*, dvs. første driftsår av tiltaket (SVV 2015). Jernbaneverket (og nå Bane NOR) bruker en analyseperiode på 40 år fra og med byggingen begynner og spaden stikkes i jorda. Dette tidspunktet kaller de prosjektets *oppstartsår* (Jernbaneverket 2015). KVU følger Jernbaneverkets praksis.

Dette har to konsekvenser: For det første medfører det at samfunnsøkonomiske analyser av jernbaneprosjekter og vegprosjekter ikke blir helt sammenliknbare. Blant annet vil altså KVUs utredning av Oslonavet ikke være helt sammenliknbar med vegprosjektene som er tatt inn i NTP (2018-2026), mens vår utredning i KVU vil være sammenliknbar med vegprosjektene, men ikke med jernbaneprosjektene i NTP. Riktignok er analyseperioden 40 år i begge tilfeller, men den faller tidligere i tid i jernbaneanalysene og KVU, og den har et annet innhold, siden en lang byggeperiode er inkludert, mens driftsperioden er kortere. Dette kompenseres ved at jernbaneanalysene gis en restverdi. Dersom denne restverdiberegningen hadde gått over like mange år som byggetida, ville vi likevel hatt sammenliknbarhet. Men den går over hele 35 år. Dette er en vesentlig kilde til manglende sammenliknbarhet.

Men for det andre vil bestemmelsen i rundskriv R-109/2014, i kombinasjon med bruken av en analyseperiode som begynner allerede når spaden stikkes i jorda, innebære at kalkulasjonsrenta

---

[7] Den riktige måten å regne investeringskostnader og restverdi på i et slikt tilfelle, er gitt i Hauge (2010), avsnitt 6.2. Vi gjengir dette avsnittet i noe lenger ned i dette vedlegget.

faller til 3 prosent allerede fra år 2057. Dette gir en viss økning av neddiskontert netto nytte i perioden 2031-2070 i forhold til å bruke 4 prosent til 2070. I tillegg kommer ytterligere et påslag i nåverdien på grunn av den delen av restverdien som faller i perioden 2070-2091. Virkningen blir ytterligere kraftig forsterket av at en regner med realprisjustering av tidsverdiene også i restverdiperioden. Samlet sett er virkningen så kraftig at vi nesten kan se bort fra å sammenlikne resultater fra KVU og andre jernbaneanalyser med analysene av vegprosjekter. Til slutt i dette dokumentet gjør vi likevel et anslag på den tallmessige forskjellen.

Hele differansen mellom oss og vegvesenet på den ene sida og BaneNOR og KVU på den andre kan føres tilbake til ulike tolkninger av to begreper i rundskriv R-109/2014. Det ene er begrepet *oppstartsår*. Jernbanesektoren tolker det som året da spaden blir stukket i jorda, men veg-sektoren tolker det som åpningsåret. Det heter i rundskrivet: «Analyseperiode og kalkulasjons-rente bør som hovedregel ta utgangspunkt i tiltakets oppstartsår.» Vi er ikke i tvil om at den riktige tolkningen er vegvesenets:

For det første har en slik tolkning av begrepet oppstartsår ikke noe grunnlag i tidligere begreps-bruk eller praksis, verken i NOU 2012:16 eller i tidligere praksis i sektoren. NOU-en definerer jo analyseperioden som perioden der en analyserer prosjektets virkninger i detalj. Det er høyst uvanlig å regne byggingen som en del av prosjektet *virkninger*. Tvert imot tenker man vel på byggingen som prosjektet, og det som skjer seinere som virkningene.

For det andre er rundskrivet opptatt av å få best mulig samsvar mellom analyseperioden og levetida, og levetida som begrep er helt klart knyttet til perioden der prosjektet gir nytte, altså etter åpning. Det er usannsynlig om man har villet innført byggetida som en fast kile mellom levetida og analyseperioden. Figur 6.1 i NOU viser da også at det ikke er tilfelle: man tenker at levetida og analyseperioden starter samtidig.

Det andre begrepet som tolkes forskjellig, er *restverdi*. NOU-en antar at restverdien er en avtrapping over ganske få år av netto nytte slik den var i siste år i analyseperioden. Vegvesenet antar at restverdier hovedsakelig vil oppstå i prosjekter bestående av flere delprosjekter som starter til ulik tid. Den avtrappes ikke i de åra den varer, men varer heller ikke særlig lang tid. Og den øker ikke med antatt trafikkvekst eller på andre måter. Jernbaneverket (2015), på den andre sida, oppfatter restverdiperioden som en tid hvor man fremdeles kjenner mange forhold rundt nyttevirkningen i detalj. Det gjelder spesielt trafikkveksten og den relative økningen i verdien av tid og miljøgoder i forhold til andre goder. Og restverdien varer lenge, helt til utløpet av antatt gjennomsnittlig levetid i jernbaneinfrastrukturen (75 år).

Vårt valg er hovedsakelig bygget på vegvesenets praksis, som ser ut til å samsvare best med tankegangen i NOU og rundskrivet. Vi bemerker ellers at det er uholdbart om disse tolknings-forskjellene får fortsette.

*Analysetidspunktet som utgangspunkt for rentenedtrappingen*

Vi har lagt til grunn NOU-ens argumentasjon for avtrappende rente, nemlig at usikkerheten om den makroøkonomiske utviklingen, og dermed avkastningen i alternative anvendelser, er økende på særlig langt sikt. I tråd med det foreslår NOU-en at renta skal trappes ned etter 40 og 75 år, og at utgangspunktet for avtrappingen er analysetidspunktet, altså ikke åpningsåret eller året da byggingen starter. Dette synspunktet står imidlertid faktisk i motsetning til rundskrivet, der det heter at analyseperiode og kalkulasjonsrente som hovedregel bør ta utgangspunkt i tiltakets oppstartsår. Med oppstartsår mener man da sannsynligvis åpningsår, ikke året da byggingen starter, slik JBV (2015) gjør.

Vi regner med at denne motsetningen vil bli avklart av Finansdepartementet.

*Supplerende litteratur om begrepet levetid*

Begrepet levetid i forbindelse med samferdselsprosjekter er utredet bl.a. i Minken m.fl. (2008), avsnitt 2.5, Minken m.fl. (2011), og Minken (2015, særlig vedlegg 2, der Vegard Østli og Marius Fossen er medforfattere).

# 2 Gjentatte reinvesteringer: Nåverdien av investeringer og restverdi når analyseperioden er lengre enn levetida

Det er et poeng å bruke en analyseperiode som mest mulig samsvarer med levetida til prosjektene som skal vurderes. Men det er et enda viktigere poeng å bruke samme analyseperiode på alle prosjekter som skal sammenliknes. I de fleste tilfeller oppnås det tilnærmet ved å anta at prosjekter med kortere levetid enn analyseperioden, kan videreføres ved å gjenta samme investering til analyseperiodens slutt. Deretter vil det som hovedregel oppstå en restverdi, fordi siste reinvestering gir nytte som varer ut over analyseperioden. Dette kan håndteres i et regneark, eller vi kan bruke formelverket nedenfor, som lett modifisert er hentet fra avsnitt 6.2 i Hauge (2010). (Hauges arbeidsdokument er tilgjengelig ved henvendelse til TØI.)

*Nåverdi av gjentatte investeringer med restverdi*

Vi skiller mellom analyseperiode og tiltakets økonomiske levetid. Tiltakets økonomiske levetid er det minste av to ting; den tekniske levetiden til et bestemt tiltak, dvs. tiden til det er utslitt og ubrukelig, og tiden fram til behovet for objektet bortfaller (for eksempel fordi det er funnet opp noe nytt og bedre). For eksempel kan man anslå at den tekniske levetiden til et leskur i gjennomsnitt er 12 år. Etter 12 år vil man da regne med å skifte ut leskuret. Analyseperioden viser til tidshorisonten benyttet i nytte- og kostnadsanalysen. Dersom levetiden til et tiltak er lavere enn analyseperioden, må man legge til grunn i analysen at det reinvesteres i tiltaket. I eksemplet med ovenfor om leskur, må man i en nytte- og kostnadsanalyse med en analyseperiode på 40 år legge til grunn at det investeres i et nytt leskur i årene 12, 24, og 36. Ved analyseperiodens slutt vil leskuret ha vært brukt i 4 år, og fortsatt kunne brukes i ytterligere 8 år. Denne restverdien av leskuret skal trekkes fra kostnadene i nytte- og kostnadsanalysen. Det kan beregnes en faktor som kostnaden ved én investering multipliseres med for å gi nåverdi av summen av investeringskostnader og eventuelle reinvesteringer over analyseperioden, fratrukket restverdien ved analyseperiodens slutt. Denne faktoren vil vi kalle **investeringsfaktoren** ($i$).

La levetida være $n$ år og analyseperioden $N$ år. Dersom vi kaller tiltakets investeringskostnad for $C_i$, og $m$ er største heltall som tilfredsstiller $n(m-1) \leq N$, vil *nåverdien av reinvesteringene* være:

$$C_{i+} \frac{C_i}{(1+r)^n} + \frac{C_i}{(1+r)^{2n}} + \cdots + \frac{C_i}{(1+r)^{(m-1)n}}$$

$$= C_i[1 + \frac{1}{(1+r)^n} + \frac{1}{(1+r)^{2n}} + \cdots + \frac{1}{(1+r)^{(m-1)n}}]$$

$$= C_i \left[ \frac{1 - \left(\frac{1}{(1+r)^n}\right)^m}{1 - \left(\frac{1}{(1+r)^n}\right)} \right] = C_i \left[ \frac{1-(1+r)^{-nm}}{1-(1+r)^{-n}} \right]$$

Restverdien settes til investeringskostnaden multiplisert med andelen av gjenværende levetid. *Nåverdien av restverdien* blir da den neddiskonterte verdien av denne restverdien:

$$C_i\left[\frac{nm-N}{n}*\frac{1}{(1+r)^N}\right]$$

Setter vi sammen nåverdien av reinvesteringene og nåverdien av restverdien, får vi da nåverdien av investeringene inkludert restverdien:

$$\begin{aligned}NV(C_i) &= -C_i\left[\frac{1-(1+r)^{-nm}}{1-(1+r)^{-n}}\right]+C_i\left[\frac{nm-N}{n}*\frac{1}{(1+r)^N}\right]\\ &= -C_i\left[\frac{1-(1+r)^{-nm}}{1-(1+r)^{-n}}-\frac{nm-N}{n}*\frac{1}{(1+r)^N}\right]\\ &= -C_i\left[\frac{1-(1+r)^{-nm}}{1-(1+r)^{-n}}-\frac{N-nm}{n}*\frac{1}{(1+r)^N}\right]=-C_i*i\end{aligned}$$

hvor $i=\left[\frac{1-(1+r)^{-nm}}{1-(1+r)^{-n}}-\frac{N-nm}{n}*\frac{1}{(1+r)^N}\right]$

**Investeringsfaktoren** (*i*)**,** avhenger av diskonteringsrenten, analyseperioden, tiltakets levetid, og heltallet *m*.

## 2.1 Sammenlikning av brutto nytte i KVU og KS1

*Henføringsår og prisår*

Hva vi bruker som henføringsår har ingen reell virkning, men det har en *illusorisk* virkning: Jo seinere henføringsår, jo mer imponerende vil alle nytte- og kostnadsposter se ut. Ved 4 prosent rente får henføringsåret 2022, som er brukt i KVU, alle tall til å se 22 prosent større ut enn når henføringsåret er 2017, som det er i KS1.

Hvilket prisår vi bruker, har heller ingen reell virkning. KVU bruker 2014 og KS1 2016. Prisstigningen på disse 2 årene var 5,8 prosent. Det innebærer at KVUs tall kan blåses opp med 5,8 prosent for å bli sammenliknbare med KS1 sine tall, eller omvendt, at KS1-tallene kan nedjusteres tilsvarende. Samlet for henføringsår og prisår får vi sammenliknbarhet om vi beholder KS1-tallene og justerer ned KVU-tallene med faktoren 1,058/1,22 = 0,87, eller omvendt justerer opp KS1-tallene med 15 prosent.

*Åpningsår, beregningsår og analyseår*

Analyseåret er tidspunktet når den samfunnsøkonomiske analysen er utarbeidet, altså når utrederen sammenstiller sine resultater. Beregningsåret er året (eller den typiske dagen i dette året) som simuleres i transportmodellen. Det kan finnes flere beregningsår, men det ser vi bort fra her.

Tidligere hadde ikke analyseåret noen annen betydning enn at jo nærmere opp til gjennomføringen av prosjektet som analysen var gjort, jo bedre anslag på trafikk og kostnader kunne man regne med å ha. Bortsett fra dette kunne to prosjekter med ulikt analyseår lett gjøres sammenliknbare ved neddiskontering til et felles henføringsår. Med bestemmelsene om diskonteringsrente i rundskriv R-109/2014 har det blitt mer komplisert. Differansen mellom analyseåret og åpningsåret har nå en reell betydning, siden jo lengre det er mellom dem, jo tidligere i levetida vil kalkulasjonsrenta synke til 3 prosent (og kanskje også videre til 2 prosent). (Et nokså merkelig resultat av dette er at et prosjekt vil bli mindre lønnsomt om man regner en gang til på det etter at det har gått noen år, helt uten at man bruker andre tall for virkningene og kostnadene.)

Nåverdien over hele levetida av en årlig netto nytte på 1 krone, kaller vi diskonteringsfaktoren. I neste kapittel (kapittel 2) viser vi hvordan andre faktorer enn renta kan innarbeides i diskonteringsfaktoren til en *justert* diskonteringsfaktor. Det forenkler beregningene for analyseperioden som helhet i forhold til et langt og uoversiktlig regneark.

KVU har 2022 som åpningsår og bruker en levetid på 75 år, mens KS1 har 2030 som åpningsår og bruker en levetid på 40 år.

Om vi nå regner alle priser i 2017-kroner og tar utgangspunkt i KS1s henføringsår 2017, analyseår 2017 og åpningsår 2030, vil virkningene etter år 2057 bli neddiskontert til 2057 med 3 prosent rente, mens eventuelle virkninger etter 2092 vil bli neddiskontert til 2092 med 2 prosent rente. Hvis derimot åpningsåret hadde vært 2022, som i KVU (men henføringsåret og analyseåret fremdeles 2017), ville tre prosent rente gjelde etter år 2062, og to prosent etter år 2097.

I kapittel 3 ser vi nærmere på et prosjekt med levetid 40 år og et annet prosjekt med levetid 75 år under ulike forutsetninger om åpningsår. Vi finner én justert diskonteringsfaktor på for prosjektet med levetid 40 år, og en helt annen for prosjektet med levetid 75 år. Ved å endre på forutsetningene om åpningsår og levetid kan vi imidlertid beregne hver av prosjektene på en måte som er mer sammenliknbar, sjøl om vi regner med avtrappende renter.

# 3 Diskonteringsfaktorer

Vi vil finne et enkelt uttrykk for den neddiskonterte verdien av en årlig nytte på 1 krone over en periode på $N$ år når diskonteringsrenta er $\rho$.

Vi veit fra en matematisk formelsamling at formelen for summen av de $N$ første elementene i en geometrisk rekke er:

$$a + ak + ak^2 + ... + ak^{N-1} = a \cdot \frac{1-k^N}{1-k}$$

Her er det forutsatt at $k \neq 1$. I tilfellet der $a = k = (1-\rho)^{-1}$ og $\rho \neq 0$ gir dette:

$$\frac{1}{1+\rho} + \frac{1}{1+\rho} \cdot \frac{1}{1+\rho} + \frac{1}{1+\rho} \cdot \left(\frac{1}{1+\rho}\right)^2 + ... + \frac{1}{1+\rho} \cdot \left(\frac{1}{1+\rho}\right)^{N-1}$$

$$= \sum_{i=1}^{N} \left(\frac{1}{1+\rho}\right)^i = \frac{1}{1+\rho} \cdot \frac{1-\left(\frac{1}{1+\rho}\right)^N}{1-\frac{1}{1+\rho}} = \frac{1-\left(\frac{1}{1+\rho}\right)^N}{1+\rho-1} = \frac{1}{\rho} \cdot \left\{1-\left(\frac{1}{1+\rho}\right)^N\right\}$$

Hvis vi ikke har neddiskontering, altså $\rho = 0$, vil summen av $N$ ett-tall, som det blir i det tilfellet, være lik $N$. Samlet har vi da følgende enkle formel for nåverdien pr. krone $\Delta$ av et hvilket som helst konstant nytteelement over $N$ driftsår:

(1) $$\Delta(\rho, N) = \frac{1}{\rho}\left(1 - \left(\frac{1}{1+\rho}\right)^N\right) \quad \text{for } \rho \neq 0$$

$$\Delta = N \text{ for } \rho = 0$$

Hvis årlig netto nytte er konstant lik *U*, er den neddiskonterte nytten over *N* år naturligvis $U\Delta$ .

## 3.1 Faste årlige vekstrater

Hva så hvis årlig netto nytte ikke er en konstant, men vokser eller avtar med en konstant prosentsats? Dette kan være på grunn av trafikkvekst, overført trafikk til eller fra det området vi analyserer, eller på grunn av voksende enhetspriser. Vi kan da danne en justert diskonteringsrente som gjør det mulig å bruke samme formel som før, men med en justert rente.

Kall vekstraten for trafikken for *g*, den relative prisveksten for *p*, og kalkulasjonsrenta for *r*. Da vil den justerte diskonteringsfaktoren være:

(2) $$\frac{1}{1+\rho} = \frac{(1+g)(1+p)}{1+r}$$

hvilket gir følgende formel for den justerte diskonteringrenta:

$$\rho = \frac{1+r}{(1+g)(1+p)} - 1$$

La oss beregne $\rho$ i noen aktuelle tilfeller. Ta Oslo og Akershus som eksempel. Vi veit at befolkningsøkningen i Oslo og Akershus er anslått i SSBs fylkesfordelte befolkningsprognose til å være omtrent 1 prosent, og i mangel av noe bedre kan vi også anta at dette er trafikkøkningen i samme område, altså *g* = 0,01. Perspektivmeldingen 2014 antar at gjennomsnittlig BNP-vekst per innbygger i samme område i tidsrommet 2014-2060 vil ligge på 1,3 prosent pr. år. Hagenutvalget anbefaler å bruke denne satsen til realprisjustering av enhetspriser som stammer fra betalingsvillighetsundersøkelser, altså *p* = 0.013.

**(*r*, *g*, *p*) = (0,04; 0,01; 0,013):**

$$\frac{1}{1+\rho} = 0,984$$

$$\rho = 0,0165$$

$$\rho^{-1} = 60,65$$

**(*r*, *g*, *p*) = (0,03; 0,01; 0,013):**

$$\frac{1}{1+\rho} = 0,993$$

$$\rho = 0,0067$$

$$\rho^{-1} = 148,93$$

**(*r*, *g*, *p*) = (0,02; 0,01; 0,013):**

$$\frac{1}{1+\rho}=1,003$$
$$\rho=-0,0031$$
$$\rho^{-1}=-326,88$$

**(*r*, *g*, *p*) = (0,04; 0,01; 0):**

$$\frac{1}{1+\rho}=0,971$$
$$\rho=0,0297$$
$$\rho^{-1}=33,67$$

**(*r*, *g*, *p*) = (0,03; 0,01; 0):**

$$\frac{1}{1+\rho}=0,9806$$
$$\rho=0,0198$$
$$\rho^{-1}=50,5$$

**(*r*, *g*, *p*) = (0,02; 0,01; 0):**

$$\frac{1}{1+\rho}=0,9902$$
$$\rho=0,0099$$
$$\rho^{-1}=101$$

**(*r*, *g*, *p*) = (0,04; 0; 0):**

$$\frac{1}{1+\rho}=0,9615$$
$$\rho=0,04$$
$$\rho^{-1}=25$$

**(*r*, *g*, *p*) = (0,03; 0; 0):**

$$\frac{1}{1+\rho}=0,9709$$
$$\rho=0,003$$
$$\rho^{-1}=33,33$$

**(*r*, *g*, *p*) = (0,02; 0; 0):**

$$\frac{1}{1+\rho}=0,9804$$
$\rho=0,02$ [2018-2022] 5 år, ingen nytte
$$\rho^{-1}=50$$

## 3.2 Anvendelse på KVU og KS1 av Oslonavet

Vi regner 2017 som henføringsår og analyseår. Det betyr at renta går fra 4 prosent til 3 prosent etter 2057, og videre til 2 prosent etter 2091. *g* og *p* vil forutsetningsvis være den samme i hele analyseperioden, og lik i KVU og KS1. Videre veit vi at vi har ulikt åpningsår i KVU og KS1, nemlig 2022 og 2030, og at vi har ulike forutsetninger om levetid – 75 år i KVU og 40 år i KS1. Det gir følgende atskilte perioder for hver av de to analysene:

**KVU:**

[2018-2022] 5 år, ingen nytte, neddiskontering til 2017 med r = 4

[2023-2057] 35 år, nytte neddiskontert til 2022 med r = 4

[2058-2092] 35 år, nytte neddiskontert til 2057 med r = 3

[2093-2097] 5 år, nytte neddiskontert til 2092 med r = 2

Den neddiskonterte nytten i hver periode må diskonteres videre ned trinn for trinn til år 2017. Til dette formålet innfører vi funksjonen $R(\rho,i)$:

(3) $$R(\rho,i)=\left(\frac{1}{1+\rho}\right)^i$$

Her er naturligvis $\rho$ den justerte kalkulasjonsrenta og *i* er antall år i perioden. Om det trengs, kan vi bruke regneregelen $R(\rho,j)\cdot R(\rho,k)=R(\rho,jk)$.

La $\rho_4$ være den justerte kalkulasjonsrenta med r = 4, $\rho_3$ den justerte kalkulasjonsrenta med r = 3, og $\rho_2$ den justerte kalkulasjonsrenta med r = 2.

Samlet nytte over hele analyseperioden blir nå V(KVU):

(4) $$V(KVU)=R(\rho_4,5)\cdot\Delta(\rho_4,35)+R(\rho_4,40)\cdot\left\{\Delta(\rho_3,35)+R(\rho_3,35)\cdot\Delta(\rho_2,5)\right\}$$

I første periode (fra 2017 til 2022) er det ingen nytte. I andre periode er det nytte over 35 år, som er diskontert videre til 2017. I tredje periode, også den på 35 år, blir 3 prosent kalkulasjonsrente brukt under til neddiskonteringen til 2057, og hele beløpet er henført til 2017 ved hjelp kalkulasjonsrenta på 4 prosent. I siste periode er 2 prosent brukt til neddiskonteringen til 2092, deretter brukes produktet av $R(\rho_3,35)$ og $R(\rho_4,40)$ for å føre det hele 75 år bakover i tid til 2017.

Fordelen med metoden er at vi, sjøl i dette kompliserte tilfellet, bare har seks tall å beregne, og de kan beregnes med lommekalkulator om man vil.

**KS1:**

For KS1 er levetida bare 40 år og periodeinndelingen enklere:

[2018-2030] 13 år, ingen nytte, neddiskontering til 2017 med r = 4

[2031-2057] 27 år, nytte neddiskontert til 2030 med r = 4

[2058-2070] 13 år, nytte neddiskontert til 2057 med r = 3

Nåverdien blir:

(5) $$V(KS1)=R(\rho_4,13)\cdot\Delta(\rho_4,27)+R(\rho_4,40)\cdot\Delta(\rho_3,13)$$

Vi kan tilnærmet regne med at alle elementer i nyttekostnadsregnestykket er proporsjonale med antall trafikanter. Siden vi også har antatt at trafikkveksten er like stor som befolkningsveksten, beholder vi $g = 0{,}01$ for all nytte. Derimot skiller vi mellom elementer som er utsatt for realprisvekst og elementer som ikke er det. De førstnevnte er spesielt nytte av spart reisetid, der vi har $p = 0$. Altså får vi to beregninger (en for tidsnytte og en for annen nytte) for hver av de to analysene, KVU og KS1.

Resultatene er gjengitt i tabell 1. Vi ser at tidselementene rundt regnet er 50 prosent større enn de andre elementene. Det gjelder både KVU og KS1. Dessuten er den neddiskonterte nytten rundt 70-80 prosent høyere i KVU enn i KS1. En del av denne virkningen, ca. 20 prosentpoeng, skyldes at KS1 antar seinere åpningsår.

*Tabell 1 Justerte diskonteringsfaktorer i KVU og KS*1

| | Tid | Annet |
|---|---|---|
| KVU | 42,6 | 27,6 |
| KS1 | 23,6 | 16,1 |

Men sjøl med samme åpningsår er differansen mellom KVU og KS1 – eller mellom Jernbanedirektoratets og SVVs diskonteringsfaktorer – betydelig. Og den er uten grunnlag. Som vi har vist, er ulik levetid nemlig ikke noe argument for ulik analyseperiode.

# Litteratur


NOU 2012:16 Samfunnsøkonomiske analyser. Finansdepartementet.

Finansdepartementet (2014) Prinsipper og krav ved utarbeidelse av samfunnsøkonomiske analyser mv. Rundskriv R-109/2014.

Hauge, K.E. (2010) Dokumentasjon av revidering av virkningsberegninger av enklere kollektivtiltak. Arbeidsdokument ØL/2272/2010, TØI.

Minken, H., Dahl, G. og Steinsland, C. (2008) Samfunnsøkonomisk analyse av vedlikeholdsstrategier, oppgradering og standardheving i vegnettet. TØI-rapport 957/2008.

Minken, H., Frislid Meyer, S., Veisten, K. og Bai, Y. (2011) Samfunnsøkonomisk analyse av vedlikehold – hva trengs i etatene? TØI-rapport 1185/2011.

Minken, H., (2015) Samfunnsøkonomisk vurdering av innsats innen drift og vedlikehold. TØI-rapport 1460/2015.

Jernbaneverket (2015) Metodehåndbok. Samfunnsøkonomiske analyser for jernbanen.

Statens vegvesen (2015) Konsekvensanalyser. Veileder. Håndbok V712.

## 2.3 Vegprising i kombinasjon med miljøkrav[8]

[8] Dette er et upublisert notat fra 2017, med en kort tilføyelse fra 2022 helt til slutt. Det dreier seg om henvisningen til TØI-rapport 1314/2014 når det gjelder egenskapene ved funksjonen S(x).

# Vegprising i kombinasjon med miljøkrav

## Modellen

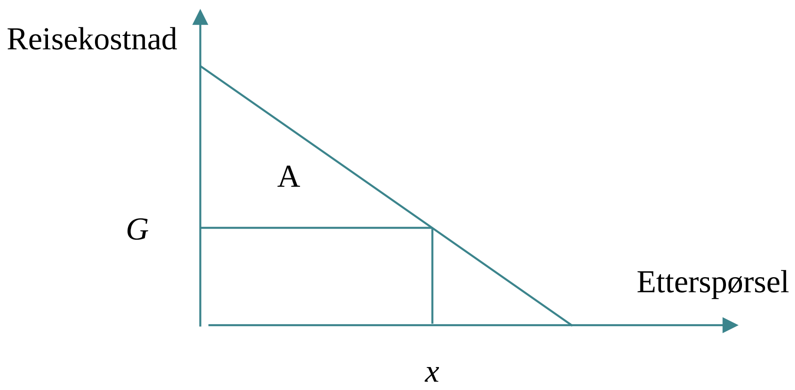
Vi skal se på hvordan en optimal vegavgift endrer seg ved en eksogen etterspørselsendring eller ved innføring av flere bibetingelser, som skyggepris på offentlige midler eller klimagassrestriksjoner. For dette formålet formulerer vi en modell som er så enkel som overhode mulig, bestående av en etterspørselsfunksjon, en tilbudsfunksjon og en mulig restriksjon på etterspørselen.

La $x$ være etterspørselen etter bilreiser i området vi studerer. Etterspørselen er en avtakende funksjon $D(G)$ av generaliserte kostnader $G$. Generaliserte kostnader er summen av trafikantbetalingen $B$ pr. reise i området og reisekostnadsfunksjonen $S(x)$, som omfatter både tidskostnader og kjørekostnader. Det er kø og trengsel som gjør at reisekostnaden er en funksjon av $x$. Både teori og empiri underbygger at funksjonen $S(.)$ skal være tiltakende og konveks.

Vi har altså:

$$x = D(G),\ D' \leq 0$$
$$G = B + S(x),\ S' \geq 0, S'' \geq 0$$

Samfunnsnytten $W$ av reiseaktiviteten i dette området er summen av brukernytten og det offentliges nytte. Brukernytten (konsumentoverskuddet) er arealet A mellom etterspørselskurven og prislinja $G$, se Figur 1.




*Figur 1 Brukernytten*

$G$ vil alltid være strengt positiv, mens derimot $B$ kan antas å være ikke-negativ. Vi vil maksimere samfunnsnytten gitt to bibetingelser, nemlig for det første at reisekostnaden kan beskrives som en generalisert kostnad $G = B + S(x)$, som nevnt, og for det andre at det foreligger en politisk målsetning om at etterspørselen ikke skal overstige nivået $\bar{x}$. Dette kan f.eks. være en klimamålsetning brutt ned på lokalt nivå eller et lokalt miljømål.

Dette optimeringsproblemet er et Kuhn-Tuckerproblem med ikke-negative variable. Skyggeprisen på offentlige midler kaller vi $\lambda$, og skyggeprisen på de to bibetingelsene kaller vi $\mu_1$ og $\mu_2$, slik at hele problemet kan skrives slik:

$$\underset{G,B}{Max} W = \int_G^\infty D(y)\,dy + (1+\lambda) BD(G)$$

gitt

$$B + S\big(D(G)\big) = G \qquad (\mu_1)$$

$$D(G) \le \bar{x} \qquad (\mu_2)$$

Vi danner Lagangefunksjonen

$$L = \int_G^\infty D(y)\,dy + (1+\lambda) BD(G) - \mu_1\Big(B + S\big(D(G)\big) - G\Big) - \mu_2\big(D(G) - \bar{x}\big).$$

Kuhn-Tucker-betingelsene for optimum er:

$$(\text{KT1}) \quad \frac{\partial L}{\partial G} = -D + (1+\lambda) BD' - \mu_1 (S' \cdot D' - 1) - \mu_2 D' = 0$$

$$(\text{KT2}) \quad \frac{\partial L}{\partial B} = (1+\lambda) D - \mu_1 \le 0 \, (= 0 \text{ for } B > 0)$$

Grunnen til at KT1 har likhetstegn er at det er utenkelig at *G* skal kunne bli 0. Vi skal også anta at *B* er større enn 0, dvs. at det er optimalt å innkreve en avgift, og dermed at det også er likhet i KT2.[9] Ved å bruke KT2 til å innsette for $\mu_2$ i KT1 kommer vi da fram til følgende vilkår for optimum:

$$\lambda D + (1+\lambda) D' \cdot \big(B - S' \cdot D\big) - \mu_2 D' = 0 \qquad (1)$$

Vi har nå fire mulige tilfeller, avhengig av om $\lambda$ og $\mu_2$ er 0 eller ikke.

**Tilfelle 1: $\lambda = \mu_2 = 0$.**

I dette tilfellet er altså størrelsen på bomavgifta *B* uten innvirkning på andre offentlige budsjetter eller det offentliges finansieringsbehov i det hele tatt, og det finnes ingen bindende begrensninger på transportaktiviteten. Den optimale avgifta er da gitt ved $B = D \cdot S'$, eller uttrykt på en annen måte:

$$B^* = x \cdot S'(x) \qquad (2)$$

[9] *B* kan bare være null om det ikke finnes køer og trengsel, og om i tillegg klima- eller miljømålet vil la seg oppfylle, sjøl uten avgift.

*B* vil *alltid* være større enn null om det finnes køer og trengsel. Om det finns kø og trengsel, er det et empirisk spørsmål om miljømålet blir oppnådd med køprising, eller om det trengs noe mer. Om det *ikke* finnes kø og trengsel, blir det et empirisk spørsmål om miljømålet kan oppnås uten avgift, eller om det trengs en avgift spesielt for det formålet.
Om etterspørselen er mindre enn $\bar{x}$ når det ikke er noen avgift, vil både *B* og $\mu_2$ være lik 0, og KT1 blir en likning til å finne $\mu_1$. Om denne $\mu_1$ ikke tilfredsstiller KT2, er denne løsningen umulig. Slik vil det sannsynligvis være i et konkret tilfelle – det vil trenges en avgift for å oppnå klimamålet eller det lokale miljømålet. Hvis ikke, er målet som ligger til grunn for KT2, allerede oppnådd uten avgift, slik mange teknologioptimister ser for seg at vil skje. Vi antar altså det motsatte.

Her er $S'(x)$ den marginale ekstrakostnaden som påføres hver av de reisende når det totale antall reisende økes litt fra nivået *x*. $x \cdot S'(x)$ er altså den samlede ekstrakostnaden som en ny reisende påfører de reisende som helhet. Dette er den velkjente regelen om at de reisende skal betale en avgift som får dem til å ta inn over seg de marginale eksterne kostnadene de påfører andre. (Stjerna på $B^*$ viser at det dreier seg om en optimal verdi, altså etter at optimeringsproblemet er løst.)

Det vi kan legge merke til, er at etterspørselsnivået *x* ikke er nivået uten avgiftsbetaling eller nivået ved en for høy eller for lav avgift, men etterspørselen akkurat slik den blir *etter at* den optimale avgifta er innført. I praksis betyr det at for å fastslå om avgifta er optimal eller ikke, må man først estimere funksjonen $S(x)$ samt anslå etterspørselsvirkningen av ulike aktuelle nivåer på avgiften. Man kan ikke regne ut avgifta utelukkende på grunnlag av empiriske data om situasjonen *før* den er innført.

**Tilfelle 2: $\lambda = 0$, $\mu_2 > 0$**

I dette tilfellet er andre bibetingelse oppfylt med likhet, hvilket betyr at vi kan finne optimal G ved å løse den framkommende likningen med hensyn på G. Vi skriver $G^* = D^{-1}(\bar{x})$. Deretter kan vi finne optimal avgift $B^*$ av første bibetingelse: $B^* = G^* - S(\bar{x})$. De optimale Lagrangeparametrene beregnes så av Kuhn-Tucker-betingelsene.

Alternativt kan vi illustrere løsningen ved å beholde $\mu_2$ som en parameter og bruke likning (1). Vi finner da at den optimale avgifta skal være:

$$B^* = \bar{x} \cdot S'(\bar{x}) + \mu_2 \tag{3}$$

Lagrangeparametrene vil alltid være ikke-negative i denne typen optimeringsproblemer. Her skal man altså øke avgifta ut over nivået i tilfelle 1, eller mer presist: ut over det nivået den samfunnsøkonomisk riktige avgifta ville hatt om etterspørselen var $\bar{x}$ og miljømålsetningen hadde latt seg oppnå ved dette nivået.

Det gjør en forskjell om tolkningen av andre bibetingelse er et klimamål eller et lokalt miljømål. I tilfellet med et klimamål kunne vi tilnærmet anta at alle kjøretøyer slipper ut like mye $CO_2$ uavhengig av tidspunktet på dagen og trengselen på det aktuelle tidspunktet. I så fall blir $\mu_2$ et fast påslag på vegprisen, uavhengig av tid på dagen og kjøreforhold. Dette vil gjøre den relative forskjellen mellom avgifta i rush og avgifta utenom rush mindre, slik det (kanskje litt overdrevet) er bestemt i tilleggsavtalen til den nye bompengeavtalen i Oslopakke 3.[10] Men om en ser litt mer nøyaktig på det, er det faktisk en tydelig forskjell på $CO_2$-utslippene under forhold med mye køer og utslippene med fri tilpasning av farta. Så sjøl om $\mu_2$ sikkert vil modifisere den relative forskjellen mellom rushtidsavgifta og avgifta utenom rush, vil den ikke fungere helt om et fast påslag.

Igjen må vi minne om at det første leddet i avgifta, $\bar{x} \cdot S'(\bar{x})$, skal vurderes ut fra de faktiske køforholdene etter at den optimale avgifta er innført. I dette tilfellet vil det bety at det blir mindre enn i tilfelle 1, slik at den relative størrelsen mellom de to leddene forskyves ytterligere i retning av siste ledd ($\mu_2$). Å finne det rette punktet dersom $\mu_2$ kan antas å være konstant over

[10] Se Oslos bompengereform – et lovende kompromiss - Samferdsel (toi.no).

døgnet, betyr å finne gir den generaliserte kostnaden som holder etterspørselen nede på $\overline{x}$, og deretter bestemme kostnaden som gir denne etterspørselen. Fordelingen av avgifta på de to leddene den består av, følger da ved å beregne $\overline{x} \cdot S'\left(\overline{x}\right)$, som likning (3) viser.

Om andre bibetingelse er et lokalt miljømål, slik at det f.eks. kreves en større prosentvis trafikkreduksjon i rush enn utenfor rush for å oppnå målet, blir det noe vanskeligere å beregne optimal avgift. Trolig må man bruke en iterativ prosedyre der man vekselvis tilpasser avgifta i og utenom rush slik at målet nås til et minst mulig samlet nyttetap.

**Tilfelle 3: $\lambda > 0$, $\mu_2 = 0$**

I Norge er det jo fastsatt at vi skal bruke en skyggepris på offentlige midler på 0,2 i samfunns-økonomiske analyser, dvs. multiplisere alle endringer i offentlige budsjetter med faktoren 1,2. I dette tilfellet er det litt vanskeligere å finne et oversiktlig uttrykk for optimal $B^*$ fra likning (1). De to triksene vi bruker, er å danne elastisiteten av $x$ med hensyn på $G$, $El_G x$, og splitte den $G$ som da blir til overs, i de to delene $B$ og $S(x)$ før vi flytter $B$-leddene over på venstre side av likhetstegnet. Vi får:

(4) $$B^* = \frac{xS'\left(x\right) - \frac{\lambda}{1+\lambda} \cdot \frac{1}{El_G x} \cdot S\left(x\right)}{1 + \frac{\lambda}{1+\lambda} \cdot \frac{1}{El_G x}}$$

Siden $El_G x$ er negativ, innebærer eksistensen av en skyggepris på offentlige midler at vi skal plusse på den optimale avgifta med et ledd som tilnærmet utgjør en fast andel (trolig omtrent en sjettedel) av tids- og kjørekostnadene $S(x)$. Men ikke nok med det: Vi skal også jekke opp det hele ved å dele på en faktor som er mindre enn 1 (trolig rundt 5/6). Det vil si at vi skal gange opp avgifta (etter påslaget av en andel av tids- og kjørekostnadene) med en faktor på rundt 1,2.

Likning (4) kan utrykkes på en mer praktisk måte ved å dele på $S(x)$. Dermed blir den optimale avgifta en bestemt andel av tids- og kjørekostnadene. Alt annet likt er denne andelen den samme enten vi vil beregne en kilometerbasert avgift eller en avgift basert på en gjennomsnittlig reiselengde i et område. Vi får nemlig:

(5) $$\frac{B^*}{S\left(x\right)} = \frac{El_x S\left(x\right) - \frac{\lambda}{1+\lambda} \cdot \frac{1}{El_G x}}{1 + \frac{\lambda}{1+\lambda} \cdot \frac{1}{El_G x}}$$

Hvis vi kjenner køfunksjonen $S(x)$, trafikkvolumet $x$, skyggeprisen på offentlige midler og elastisiteten av bilreiser med hensyn på generaliserte kostnader, kan vi åpenbart anslå optimal avgift.

Denne framgangsmåten kan i det minste brukes som en første tilnærming. Så vil mer nøyaktige modellkjøringer og empiriske undersøkelser kunne forbedre opplegget etter hvert.

**Tilfelle 4: $\lambda > 0$, $\mu_2 > 0$**

I dette tilfellet får vi samme løsning som i tilfelle 2, dvs. $B^* = D^{-1}(\bar{x}) - S(\bar{x})$. Løsningen framkommer av de to bibetingelsene med likhet. Siden $\lambda$ ikke inngår i bibetingelsene, har det heller ingen virkning på løsningen om den er null eller positiv.

Verken løsningen i tilfelle 2 eller tilfelle 4 har egentlig noe med vegprising å gjøre – bibetingelse 2 overstyrer fullstendig prinsippet om internalisering av marginale eksterne kostnader. Ingen spor av strukturen fra den samfunnsøkonomisk optimale løsningen kan gjenfinnes i formelen $B^* = D^{-1}(\bar{x}) - S(\bar{x})$. Men hvis vegprising med differensiering av avgiften i og utenom rushtid hadde gitt et resultat som var i nærheten av miljøkravet, ville vi likevel kunne oppnå både samfunnsøkonomi og miljømål ved å legge prisingsprinsippene i vegprising til grunn for avgiften. Det motsatte er ikke tilfelle: En flat avgift som oppnår miljømålet kan ikke gi noen god samfunnsøkonomisk løsning.

Med sikte på å knytte en viss kontakt mellom den samfunnsøkonomiske løsningen og miljøløsningen, kan vi finne løsningen i tilfelle 4 når vi betrakter $\mu_2$ som en parameter. Som i tilfelle 2 og 3 tar vi da utgangspunkt i likning 1. Vi får:

$$\frac{B^*}{S(\bar{x})} = \frac{El_x S(\bar{x}) - \frac{\lambda}{1+\lambda} \cdot \frac{1}{El_G \bar{x}}}{1 + \frac{\lambda}{1+\lambda} \cdot \frac{1}{El_G \bar{x}}} + \frac{1}{1 + \lambda\left(1 + \frac{1}{El_G \bar{x}}\right)} \cdot \frac{\mu_2}{S(\bar{x})} \tag{6}$$

Denne likningen viser at den optimale $B^*$ i dette tilfellet har strukturelle trekk både fra løsningen i tilfelle 2 (likning 3) og løsningen i tilfelle 3 (likning 5). Første ledd er nemlig den optimale vegprisen dersom den optimale etterspørselen er så lav som $\bar{x}$, jfr. likning (5). Andre ledd er et påslag, tilsvarende påslaget i likning (3), som faktisk medfører at etterspørselen blir så lav som $\bar{x}$.

Hvis miljøkravet er meget stramt, blir etterspørselen så lav at første ledd blir ganske lite – vegprising på dette nivået har liten betydning og innebærer små avgifter. Det aller meste består da av påslaget for å komme ned på den lave etterspørselen. Omvendt blir vegprisingen viktigere om miljømålet lar seg oppfylle ved et etterspørselsnivå som ikke er særlig mye lavere enn det vegprisingen aleine gir.

## Kalibrering av modellen

### Skattefinansieringskostnaden 1 + $\lambda$

En av faktorene som påvirker optimal vegpris i vår modell, er skattefinansieringskostnaden, dvs. merkostnaden for samfunnet ved å bruke vridende skatter til å finansiere et offentlig tiltak av marginal størrelse. I Norge er denne merkostnaden satt til 0,2 i samfunnsøkonomiske analyser siden 1997.[11] Den samlede kostnaden pr. krone av å bruke en slik finansieringsmåte, altså 1 + $\lambda$, kalles skyggeprisen på offentlige midler, den marginale skattefinansieringskostnaden, kort og godt skattefinansieringskostnaden, eller bare skattefaktoren. Den er altså 1,2.

[11] Se NOU 1997:27, Finansdepartementet (2014), DFØ (2014).

Det er svært usikkert hva denne faktoren egentlig skal være. Den kan ikke observeres direkte, så det en må gjøre, er å formulere en modell av hele økonomien og forsøke å estimere parametrene i den. Modellen vil imidlertid alltid være en forenkling av virkeligheten. Blant de forenklingene som ofte er gjort, er å anta at det bare finns ett menneske i landet, eller eventuelt at alle mennesker i økonomien er like og kan representeres av en såkalt representativ konsument, at det bare finns ett skattbart gode og bare en slags skatt (som regel inntektsskatt), og at denne skatten innkreves med en flat sats.

Litteraturen om skattefinansieringskostnaden deler seg i to teoritradisjoner. Den første kalles Stiglitz-Dasgupta-Atkinson-Stern (SDAS) og den andre Pigou-Harberger-Browning (PHB) (Holtsmark og Bjertnæs 2015).

I SDAS-teorien er utgangspunktet en liten reduksjon i inntekten til konsumenten eller konsumentene i privat sektor, f.eks. innført for å finansiere et prosjekt som frambringer et perfekt offentlig gode. Reduksjonen medfører lavere etterspørsel etter et sammensatt privat gode – la oss kalle det privat forbruk. Nyttetapet som det medfører måles som tapt konsument-overskudd i markedet for det private godet. Til dette bruker vi den vanlige, ukompenserte etterspørselskurva. Siden det ikke gis noen kompensasjon for nyttetapet, har konsumenten et nyttetap.

I PHB-teorien, derimot, bruker vi den kompenserte etterspørselskurva. Det vil si at vi forutsetter at konsumenten bringes opp på samme nyttenivå som før skatteøkningen også etter den marginale skatteøkningen. Det kan f.eks. skje ved et tilskott som ikke avhenger av hvor mye konsumenten arbeider eller forbruker – et såkalt lump-sum tilskott. Det kan også skje ved å produsere mer av et perfekt offentlig gode – antakelig det samme godet som skatteøkningen skulle finansiere.

På bakgrunn av denne skissemessige framstillingen av forskjellen mellom SDAS og PHB kan vi allerede trekke en viktig konklusjon, nemlig at i transportanalyse er det SDAS som er relevant, mens PHB er irrelevant. Årsaka er at tiltak i transportsektoren ikke frambringer noe perfekt offentlig gode i det hele tatt.

Et perfekt offentlig gode skal være slik at den enes forbruk av godet ikke går på bekostning av andres forbruk, og at det ikke går an å utestenge noen fra forbruket av godet. Det er derfor heller ikke mulig å ta noen pris for å bruke det. Når det gjelder transport, betaler imidlertid forbrukeren alltid en pris, som regel i form av penger til billett eller kjørekostnader, men alltid i form av bruk av egen tid. Det er også slik at den enes forbruk helt eller delvis utelukker eller fortrenger andres forbruk, f.eks. i form av kø og trengsel. Transport er altså på mange måter ganske så likt vanlige private goder. Siden vi ikke kan kompensere konsumenten ved å gi ham et gratis kollektivt gode, måtte vi altså supplere skatteøkningen med en lump-sum overføring. Men det er jo aldri det som skjer når vi innfører et transporttiltak.

Vi mangler altså helt forutsetningene i PHB-teorien – og i hele teorien for nytte-kostnads-analyse, i den grad den legger til grunn at det er et perfekt offentlig gode som blir frambrakt gjennom tiltaket. I stedet søker vi å beregne trafikantenes nytte av tiltaket i en analyse av endringene i transportmarkedene som tiltaket skaper. Til det bruker vi (som regel) ukompenserte etterspørselsfunksjoner.

Men transport er heller ikke et privatprodusert gode. Som regel vil et transporttiltak kreve bruk av skattepenger – samtidig som bruksavgifter også vil spare skattebetalerne for økte skatter. Vi må derfor se nærmere på SDAS-teorien om skattefinansieringskostnaden.[12]

### Hva kan vi anta om *S*(*x*)?

Vi viser i TØI-rapport 1314/2014 at volume-delayfunksjoner generelt kan sees som kombinasjonen av Littles formel og en antakelse om hvordan bilistene avpasser farten for å oppnå ønsket avstand til bilen foran. Det innebærer blant annet at volume-delayfunksjoner er tiltakende og konvekse.

Avsnitt 4.1 og 4.2 i den rapporten utleder for øvrig regler for køprising med og uten en flaskehals. Forskjellen mellom flaskehalsmodellen og den miljøkravmodellen vi har utledet her, er at de som ikke uten videre kommer gjennom flaskehalsen, samler seg opp i en kø.

## Litteraturliste

Direktoratet for økonomistyring (2014) Veileder i samfunnsøkonomiske analyser. https://dfo.no/fagomrader/publikasjoner/veiledere/.

Finansdepartementet (2014) Rundskriv R-109/14. Prinsipper og krav ved utarbeiding av samfunnsøkonomiske analyser mv. https://www.regjeringen.no/globalassets/upload/fin/vedlegg/okstyring/rundskriv/faste/r_109_2014.pdf

Holtsmark, B., and Bjertnæs, G. H. M., 2015, The size of the marginal cost of public funds: a discussion with special relevance to Norway, Reports, 2015/ 13, Statistics Norway.

Minken, H. (2014) Statisk trafikkteori og køprising. TØI-rapport 1314/2014.

NOU 1997:27. Nytte-kostnadsanalyser. Prinsipper for lønnsomhetsvurderinger i offentlig sektor.

[12] Her ender manuskriptet fra 2017.

## 2.4 Problemer i usikkerhetsanalyse av samfunnsøkonomiske beregninger[13]

[13] Dette er arbeidsdokument ØL/2319/2011.

# Problemer i usikkerhetsanalyse av samfunnsøkonomiske beregninger av samferdselsprosjekter

## Innhold



# 1 Innledning

Dette arbeidsdokumentet vil sette søkelys på noen metodiske problemer som dukker opp når en vil drive usikkerhetsanalyse av samfunnsøkonomiske beregninger i transport. Problemene henger sammen med at modellapparatet som vi bruker ved samfunnsøkonomiske analyser i transport, er mer komplisert enn det de har hatt i tankene, de som laget Finansdepartementets og SSØs vegledere i samfunnsøkonomisk analyse og usikkerhetsanalyse. Det består ikke bare av valgte forutsetninger og en enkel kalkyle på det grunnlaget, men involverer som regel en transportmodell, dvs. en partiell likevektsmodell der alle transportstrømmer og de tilhørende reisekostnadene innen et geografisk område blir fastlagt. Å anta andre verdier på de usikre faktorene enn de som i utgangspunkt er lagt inn i modellen, kan kreve omkoding, og derfor være et større og mer tidkrevende arbeid enn i enklere modeller. Men om man på den andre sida ikke går dypt inn og endrer den usikre faktoren i selve modellsystemet, men nøyer seg med å gjøre endringen i det avsluttende nyttekostnadsregnestykket, er det fare for inkonsistens mellom modellsystemet og kalkylene, hvilket naturligvis i seg sjøl gir opphav til usikkerhet om resultatet er gyldig.

Før vi kan definere problemene mer presist må vi gi en kort beskrivelse av hva de generelle vegledere sier om usikkerhetsanalyse. Det gjøres i kapittel 1. Deretter beskriver vi hvordan samfunnsøkonomiske analyser i transportsektoren er lagt opp i kapittel 2. Kapittel 3 gir en første definisjon og klassifisering av hvilke usikkerhetsfaktorer som finns i slike analyser, og forteller hvordan vi vanligvis gjør usikkerhetsanalyser i samferdsel. Kapittel 4 diskuterer metode-usikkerhet. Vi er da klar til å ta opp hovedspørsmålet, nemlig hvordan vi best kan legge opp en usikkerhetsanalyse i transport, herunder hvilke former for usikkerhet som kan behandles uavhengig av hverandre, hvordan leddene i nyttekostnadsanalysen påvirkes ulikt av usikkerheten, og hvordan vi skal redusere problemene med avhengighet mellom ulike usikre faktorer. Dette dekkes i kapittel 5. En kort konklusjon er tatt inn i kapittel 6.

# 2 Offisielle anbefalinger om usikkerhetsanalyse

I Finansdepartementet (2005) og SSØ (2010, 2006) er det gitt generell vegledning om hvordan usikkerheten i nyttekostnadsanalysene skal analyseres. I Finansdepartementet (2011) er det gitt egne regler og krav som gjelder for store statlige prosjekter som faller inn under KS-ordningen.

SSØ skiller ikke mellom usikkerhet og risiko. Vi vil følge SSØs begrepsbruk på dette punktet.

Usikkerheten skal framfor alt *synliggjøres* og om mulig *tallfestes*. Hensikten med usikkerhetsanalysen er ifølge SSØ å klargjøre hva som vil skje dersom sentrale forutsetninger endres, og identifisere hvilke faktorer som påvirker resultatet sterkest. Det hører med til analysen å vurdere behovet for risikoreduserende tiltak og hvilke tiltak som bør tas på dette stadiet for å redusere usikkerheten. En annen vurdering som bør gjøres, er om risikoen kan kontrolleres eller unngås ved å utsette tiltaket, for eksempel i påvente av mer pålitelig informasjon om behov, virkninger osv.

I alle vegledere sies det at nyttekostnadsanalysen skal bygge på *forventningsverdier* for de usikre variablene. Tankegangen er her at *først* skal vi finne virkningene målt i fysiske størrelser, *deretter* skal vi finne den samfunnsøkonomiske verdien (prisen, kostnaden) pr. enhet av virkningen. Så skal vi multiplisere virkning med verdi for å finne den samfunnsøkonomiske effekten målt i kroner. Dersom verdi og fysisk virkning ikke er stokastisk avhengige av hverandre eller står i funksjonssammenheng med hverandre, er det fullt ut mulig å gjøre tingene i denne rekkefølgen, og forventningsverdien vil da være forventet verdi ganget med forventet fysisk virkning. Men nettopp dette er uholdbare forutsetninger når det gjelder mange av virkningene i samferdsel.[14] Verdi og virkning må finnes simultant, og forventningsverdien er ikke produktet av forventingsverdiene.

Etter at nyttekostnadsanalysen er gjennomført med forventningsverdier for de usikre variablene, anbefales det i veglederne å beskrive de usikre faktorene i regnestykket og velge ut de viktigste av dem for nærmere analyse. Man kan for eksempel angi en høyeste og laveste verdi som de usikre faktorene realistisk sett kan anta, og gjøre nyttekostnadsregnestykket om igjen med høyeste og laveste verdi på de usikre faktorene, en etter en. Dette er en form for *følsomhetsanalyse*. En litt mer avansert analyse tar for seg flere usikre faktorer samtidig, og ser på ulike kombinasjoner av ekstremverdier for faktorene. Dersom kombinasjonene man analyserer er framkommet ved en systematisk undersøkelse av hvordan de usikre faktorene henger sammen med hverandre, kan det kalles *scenarioanalyse*.

Man innrømmer at det å definere ulike mulige utfall med hver sin sannsynlighet, og dermed også en forventningsverdi, ofte vil måtte gjøres subjektivt og skjønnsmessig, men mener det

[14] Etterspørselen er for eksempel en funksjon av reisekostnad og tidsbruk, hvilket gjør at forventet verdi av spart tid involverer kovariansen mellom tidsbesparelsen og etterspørselen. Se et seinere kapittel for en nærmere drøfting av dette.

likevel vil gi innsikt. Et slikt ”bayesiansk” syn, som betrakter sannsynligheter som uttrykk for hva vi på det nåværende tidspunkt mener å vite om et saksforhold, står i motsetning til en ”frekventilistisk” filosofi, der en bare vil snakke om sannsynlighet og sannsynlighetsfordelinger når det kan begrunnes i observasjoner eller teori. En alternativ form for følsomhetsanalyse, som ikke krever noen eksplisitt bayesiansk ex ante sannsynlighetsfordeling, kunne gått ut på å etablere hvor feil vi kan ta om enkeltfaktorene, enkeltvis eller i kombinasjon, uten at konklusjonen om samfunnsøkonomisk lønnsomhet eller ulønnsomhet endres.

Når usikkerheten totalt er kartlagt og kanskje kvantifisert, vil det være grunnlag for å dele den inn i to slag, *systematisk* og *usystematisk* usikkerhet. Systematisk risiko innebærer at årlig netto nytte i prosjektet avhenger av konjunkturene. Usystematisk risiko er alt mulig annet, både usikkerhet som man kan jobbe med å redusere og ting som det er vanskelig å gjøre noe med. De store tall lov tilsier imidlertid at jo flere prosjekter man har nytte av, jo mindre blir den usystematiske risikoen pr. nytte- eller kostnadsenhet, gitt et bestemt nivå på arbeidet med å redusere og kontrollere risiko.

Risiko er jo noe man helst vil unngå. Men den systematiske risikoen er uunngåelig. Det betyr at et prosjekt med høy systematisk risiko er mindre verdt for samfunnet enn et annet som gir like stor forventet netto nytte, men har lavere risiko. Vi kan justere for systematiske risikoforskjeller ved å omregne årlig netto nytte til såkalte sikkerhetsekvivalenter, eller alternativt ved å inkludere en risikopremie i kalkulasjonsrenta. Begge deler gjør årlig netto nytte i et framtidsår litt mindre verdt. Sikkerhetsekvivalenter er i prinsipp å foretrekke dersom risikoen endrer seg fra år til år. Men den enkle metoden som de offisielle veilederne angir for å beregne sikkerhetsekvivalenter, gjør i praksis sikkerhetsekvivalentmetoden helt identisk med metoden med risikopremie i kalkulasjonsrenta.

Risikopremien i kalkulasjonsrenta er estimert for jernbane, veg, fly og sjøfart, samt for godstransport, i Minken (2005). På grunnlag av disse estimatene, som ikke var så veldig forskjellige fra hverandre, fastsatte Samferdselsdepartementet i 2006 risikopremien til 2,5 prosent, og dermed kalkulasjonsrenta i samferdselsprosjekter til 4,5 prosent. Det framstilles ofte som om satsen 4,5 prosent er fastsatt av Finansdepartementet, men det er altså ikke riktig. Noen steder hevdes det også at estimatet på 4,5 prosent bygger på CAPM-modellen. Det er heller ikke riktig – det bygger på en modell som opprinnelig stammer fra Lund (1987, 1993). Estimering av risikopremier er imidlertid ikke noe som kan gjøres en gang for alle. Mye har endret seg siden begynnelsen av 2000-tallet, og det kan være behov for å reestimere denne modellen nå.[15]

For store prosjekter er det bestemt at det skal gjøres egne analyser av den systematiske usikkerheten i hvert prosjekt, og at metoden med sikkerhetsekvivalenter skal anvendes.[16] I mange tilfeller bruker man likevel en kalkulasjonsrente på 4,5 prosent også der. Grunnen er at en ikke har datagrunnlag for å forbedre anslaget på 4,5 prosent fra 2006. Man har oftest heller ikke data som tilsier at man vil oppnå andre resultater med å bruke sikkerhetsekvivalentmetoden enn de man får med risikopremiemetoden.

[15] Vi noterer dette som den første praktiske konklusjonen i dette arbeidsdokumentet: Den økte markedsprisen på risiko etter finanskrisa, samt nytt empirisk materiale om sammenhengen mellom oljepris, børs og lønnsutvikling, tilsier at risikopremien i samferdsel snart blir estimert på nytt.

[16] Finansdepartementet (2011), som ikke bare krever egne usikkerhetsanalyser for hvert prosjekt, men også har den mest detaljerte beskrivelsen av hvordan usikkerhetsanalyse skal gjennomføres.

Vi er naturligvis enig i at det bør gjennomføres en usikkerhetsanalyse av den samfunnsøkonomiske beregningen, blant annet for å identifisere usystematisk risiko som det kan gjøres noe for å redusere allerede før tiltaket iverksettes. Vi mener også at ulik grad av usystematisk usikkerhet bør telle med når det skal prioriteres mellom tiltaksalternativer. Men offisielle vegledere har en litt for naiv tilnærming til tallfesting av risikoen i mindre prosjekter. Når det gjelder systematisk risiko i store prosjekter, på den andre sida, legger vegledern opp til framgangsmåter som ser sofistikerte ut, men ofte ikke gir noen tilleggsgevinst i forhold til de enklere retningslinjene for mindre prosjekter.

# 3 Samfunnsøkonomiske analyser i samferdselssektoren

Det metodiske apparatet i en samfunnsøkonomisk analyse i samferdselssektoren består av følgende deler:

- Et scenario
- Et transportmodellsystem
- Et nytteberegningsverktøy

## 3.1 Scenario

*Et scenario* er en beskrivelse av hvordan de ytre forholdene som påvirker transportsystemet, utvikler seg fra år til år i analyseperioden (den perioden vi skal gjøre nytteberegning for). Vi kan også kalle det *forutsetningene* for den etterfølgende analysen med transportmodellen og nytteberegningsverktøyet. Delvis dreier det seg om generelle forutsetninger, som gjelder landet og verden som helhet, og delvis om spesielle forutsetninger, som gjelder innenfor det valgte studieområdet. Studieområdet er det geografiske området der tiltaket vi skal nyttekostnadsberegne, kan ventes å ha virkninger.

De generelle forutsetningene gjelder blant annet inntektsutviklingen i samfunnet, den teknologiske utviklingen, for eksempel når det gjelder drivstoffeffektivitet og kjøretøyteknologi, og den politikken som vil bli ført på nasjonalt og internasjonalt plan, for eksempel drivstoffavgifter og klimapolitikken. De felles forutsetningene kan også dreie seg om hvilke enhetspriser på markedsgoder og ikke-omsatte goder som skal brukes i nytteberegningene, og hvordan de relative prisforholdene utvikler seg over tid. Et eksempel som har blitt satt på dagsorden i den siste tida, er hvordan tidsverdiene utvikler seg over tid. Det er grunn til å merke seg at en rekke kalkulasjonspriser på tid, helse og miljøgoder kan tenkes å utvikle seg mer eller mindre i takt med inntektsutviklingen, men det er også grunn til å merke seg de andre samfunnsendringene som også vil påvirke disse enhetsprisene (Minken 2011).

Dersom man skal sammenlikne prosjekter på tvers av geografiske områder og transportmåter (f.eks. i forbindelse med den nasjonale transportplanen), bør de generelle forutsetningene være de samme. Det er grunnen til at man, i land som er litt mer systematisk anlagt enn vårt, har bygget opp et eget utredningsapparat og et beslutningsorgan som for hver planleggingsomgang gjennomgår alle enhetspriser og parametre og fastlegger felles retningslinjer for alle prosjektene som skal konkurrere om offentlige midler.

De spesielle eller lokale forutsetningene omfatter i første rekke utviklingen av folketallet og befolkningens sosio-økonomiske sammensetning i studieområdet, næringsutviklingen, som bestemmer hvor arbeidsreisene går hen, utviklingen av kjøpesentre og andre attraksjoner, hvilke andre infrastrukturprosjekter som vil være på plass i studieområdet i analyseperioden, eventuelle bompenger og bompengesatser, samt hvordan kollektivtrafikken er utbygd og drives. Endelig må det også gjøres forutsetninger om lokale myndigheters politikk, kanskje spesielt arealbrukspolitikken.

Det som er viktig å merke seg om de lokale forutsetningene, er at vi ikke bare trenger dem på studieområdenivå, men på sonenivå. Studieområdet må nemlig inndeles i soner, ofte på ganske detaljert geografisk nivå (grunnkretser). En reise eller en transport har opphav i en sone og målpunkt i en annen sone. Dette er grunnlaget for transportmodellen.[17] Derfor trenger vi å vite hvor mange og hva slags folk som bor i hver av sonene, hvor mange arbeidsplasser og andre attraksjoner det er i hver av sonene, i hvilke soner lokalpolitikerne har tenkt å bygge nye boligområder, osv.

Det vi kjenner godt, er demografien og næringslivet i sonene *slik det er i dag*. Vi kan vel også regne med at demografien ikke endrer seg veldig raskt. Men arbeidsplassenes lokalisering kan endre seg ganske raskt i mindre samfunn med én eller noen få hjørnesteinsbedrifter. Og straks vi vil si noe om befolkning og næringslivet på sonenivå ti, tjue eller tredje år fram i tid, blir usikkerheten betydelig. Til grunn for prognosene på nasjonalt nivå og fylkesnivå ligger det faktisk innbakt ganske mye skjønn og tru på at framtida vil lik fortida. Og når tallene skal brytes videre ned på sonenivå, øker usikkerheten knyttet til flytting, nedleggelser, nye næringer osv. raskt jo finere soneinndelingen blir.

Scenarioet inngår i transportmodellsystemet på en slik måte at det kan kreve et betydelig kodearbeid å etablere en ny framtidssituasjon i modellen. Dette er grunnen til at man bare vil kjøre modellen for de bestemte årene som man på forhånd har arbeidet mye med å kode inn i modellen. Det er også en vesentlig hindring for konsistente følsomhetsanalyser, sjøl om hindringen er mindre jo enklere og grovere modellen er.

Modeller for førerkortinnehav og bilhold er noe midt imellom ytre forutsetninger og en del av transportmodellsystemet. Grunnen er at de involverer mer grunnleggende valg enn de daglige valgene av *om* man skal reise, og i tilfelle hvor, når og hvordan. Prognoser for førerkort og bilhold brukes derfor som regel til å danne de faste ytre forutsetningene for transportmodellkjøringene. Men samtidig er dette viktige deler av de valgene som gir opphav til den daglige trafikken. Når en skal teste for eksempel virkemidler for å oppnå klimapolitiske mål på lengre sikt, burde en legge større vekt på å modellere at husholdningene kan endre bilholdet (og lokaliseringen) enn det man ofte har gjort til nå, for eksempel i prosjektet Klimakur, som skal gi grunnlaget for regjeringens klimapolitikk.

## 3.2 Transportmodellsystemet

Transportmodellsystemet består av en etterspørseldel og en tilbudsdel. Nå til dags er det vanlig å bygge *etterspørselsmodellen* på stokastisk nytteteori (random utility modelling). Det vil si at

[17] Under bestemte forutsetninger kan vi klare oss uten opplysninger om reisenes opphav og målpunkter, og nøye oss med opplysninger om trafikkvolumene på *lenkene* i transportnettverket. Disse forutsetningene er fast etterspørselsmatrise (ingen endring i antall reiser på noen reiserelasjoner som følge av tiltaket) og et transportnettverk med få eller ingen alternative ruter for noen av reisene. Ofte sier vi da at vi ikke bruker noen transportmodell, men det er vel riktigere å si at modellen er så enkel at vi klarer oss med reint lokale data fra den vegen som blir forbedret og noen få veger i nærheten. Verktøyet for å analysere slike situasjoner heter EFFEKT.

modellen er estimert på grunnlag av data om enkeltindividers valg – enten hypotetiske valg eller faktiske valg slik de er kartlagt i reisevaneundersøkelser. For hver reisehensikt velger individene reisehyppighet, destinasjon, transportmåte og kanskje tid på dagen. Valgmulighetene avhenger av bl.a. biltilgang. Sosioøkonomiske kjennetegn ved individet (kjønn, alder, inntekt) påvirker valget. Det samme gjør naturligvis reisekostnadene, som i regelen er modellert som *generaliserte kostnader*, dvs. summen av pengeutlegg og tidsforbruket verdsatt i kroner.

Den aggregerte transportetterspørselen fra hele befolkningen i en sone modellerer vi ved å la hvert individ i vårt opprinnelige datamateriale representere en undergruppe av sonebefolkningen, og tilpasse størrelsen på disse undergruppene slik at de sosioøkonomiske kjennetegnene ved sonebefolkningen som helhet gjenskapes best mulig. Dette kalles 'prototypisk utvalg'. På mellomlangt og langt sikt vil de transportrelevante egenskapene ved sonebefolkningene kunne endre seg relativt dramatisk, som når den befolkningen som opprinnelig flyttet inn i en drabantby, dør ut og erstattes med nye innflyttere.

Vi ser derfor at om det skal gjøres en konsistent følsomhetsanalyse av slike scenariovariable som gjennomsnittsinntekt, inntektsfordeling, kjønnssammensetning og alderssammensetning, må vi gjøre en jobb inni transportmodellen med å sette sammen sonebefolkningene på nytt. Samtidig vil endringer i gjennomsnittsinntekt også eventuelt påvirke tidsverdier og dermed generaliserte reisekostnader, hvilket krever enda mer omkoding om vi skal ta hensyn til det.

Tilbudssida i modellsystemet består av en nettverksmodell for biltrafikken og eventuelt egne nettverksmodeller for kollektivtrafikk (tog, buss, fly, båt, eller skinnegående kollektivtrafikk og buss i et byområde). Slike modeller består av lenker og noder. Nodene kan være vegkryss, kollektivterminaler e.l., og lenkene kan være vegstrekninger eller banestrekninger fra node til node. Her vil vi bruke vegnettverket som eksempel.

All trafikk fra en sone antas å oppstå i en særskilt node, en såkalt sentroide, og ha målpunkt i en annen sentroide. Fra startpunktet til målpunktet kan den følge ruter bestående av et utvalg av sammenhengende lenker. På lenkene vil det gå trafikk fra mange forskjellige reiserelasjoner (sonepar). I modellen kan det dessuten finnes svært mange ulike ruter mellom to sentroider. Spørsmålet er da hvilken rute som de reisende på hver av reiserelasjonene vil bruke. Vi antar at de reisende ikke har annen interesse av rutene enn å minimere reisekostnadene. Dette i motsetning til valgene i etterspørselsdelen av modellen, som dreier seg om det folk virkelig er interessert i med reisene: muligheten til å drive aktiviteter som ikke kan drives hjemme, reisekomfort osv.

Hvis det ikke finns kø og trengsel, kan folk velge rute uten at de andres valg påvirker rutekostnadene. De antas da å velge billigste rute (målt i generalisert kostnad). Ved kø, derimot, antas folk å velge rute slik at de ikke vil angre på sitt valg med mindre andre endrer sine ruter. Dette prinsippet for rutevalget kalles brukerlikevekt og er en form for Nash-likevekt. Først når en slik likevekt er oppnådd, vil de generaliserte reisekostnadene være entydig bestemt.[18]

---

[18] I kollektivtransport kreves det litt andre framgangsmåter for å bestemme hvilke linjer og lenker trafikantene vil bruke. Et vanlig brukt prinsipp er at trafikanten utser seg noen linjer som alle vil føre til målet, og velger den linja blant dem som først dukker opp på holdeplassen. Hvilke linjer som finns å velge mellom, og hvor ofte de kommer, antas da for gitt. Men videre kan vi ha en situasjon hvor kollektivtilbudet blir bestemt av kollektivselskapets økonomiske tilpasning, for eksempel ved profittmaksimering. Noen ganger burde vi modellere imperfekt konkurranse mellom konkurrerende selskaper, enten innen samme transportmåte eller mellom transportmåter. For eksempel vil innføring av høyhastighetstog utvilsomt ha konsekvenser for prisene i flymarkedet og avgangsfrekvensene fra de enkelte flyplassene. Det er flyselskapenes økonomiske tilpasning

La **X** være en vektor av antall reiser på alle reiserelasjoner, og la **G** være de tilhørende generaliserte kostnadene etter at brukerlikevekt er oppnådd. Vår etterspørselsmodell er en funksjon **X = D(G)** som gir volumet i alle reisemarkedene som en vektor av funksjoner som alle i prinsippet er funksjoner de generaliserte reisekostnadene i alle reisemarkeder. Grunnen er at individene som velger destinasjon, transportmiddel osv. vil vurdere kostnadene i mange markeder før de bestemmer seg for ett av dem. Men **G** er på den andre sida i prinsippet en tiltakende og konveks funksjon av reisevolumene i alle markedene. Den kan vi skrive **G = S(X)**. Dette er tilbudsdelen av modellen. I likevektspunktet der **X = D(G)** og **G = S(X)** begge gjelder, vil ingen angre på at de la ut på reise dersom ingen andre endrer sine reisebeslutninger. Vi har altså en dobbelt likevekt – både i rutevalget og i reisemarkedene totalt.

Dette er et idealbilde av en transportmodell. I praksis er det ikke sikkert modellene blir kjørt til likevekt, og dermed kan de generaliserte kostnadene som kommer ut av rutevalget avvike fra kostnadene som bestemmer etterspørselen. Et annet inkonsistensproblem er at de mange ulike brukergruppene som finns i etterspørselsdelen, hver med sine tidsverdier og generaliserte kostnader, plutselig har blitt til én eneste gruppe med en felles tidsverdi i rutevalget. Slike inkonsistenser kan løses i teorien, men det er ikke vanlig å gjøre det i praksis.

Når det gjelder usikkerhetsanalysen, vil slike scenariovariable som kjøretøyteknologi og nasjonal avgiftspolitikk åpenbart påvirke generaliserte kostnader. En følsomhetsanalyse av slike faktorer krever derfor at kostnadene på lenkene kodes om og at modellen kjøres til en ny likevekt. Dette er nokså tidkrevende, hvilket kan være grunnen til at slike følsomhetsanalyser ikke gjøres i praksis. Mye det samme vil gjelde for følsomhetsanalyse av tidsverdiene og pengekostnaden ved å reise. Det stiller seg annerledes med enhetsprisene på ulykker og miljø – som ikke antas å påvirke etterspørselen.

Jo enklere og grovere transportmodellen er, jo enklere vil det være å tilpasse den til endringer i scenariovariablene. Det gjør det også lettere å gjennomføre usikkerhetsanalyse. I det enkle tilfellet hvor trafikantene ikke endrer atferd som følge av tiltaket, vil usikkerhetsanalysen kunne gjennomføres stort sett på den måten som de offisielle veglederne har tenkt seg.

## 3.3 Nyttekostnadsberegningen

Persontransport er hjemmeproduksjon. Folk bruker innkjøpte innsatsfaktorer (husholdningens egne transportmidler, drivstoff m.m., eller tilgang til et kollektivt transportmiddel) i kombinasjon med egen tid og innsats til å produsere *reiser*. Reiser er på sin side stort sett bare innsatsfaktorer i produksjonen av det man *egentlig* har nytte av, nemlig aktiviteter som ikke kan drives hjemme.

Dette at folk må betale og på annen måte yte noe sjøl for å kunne reise, gjør at reisemuligheter ikke er et offentlig gode i egentlig forstand. For det første: Gjennom kø og trengsel påvirkes kostnaden og kvaliteten av min reise av hvor mange *andre* som reiser – mitt forbruk og de andres forbruk er *rivaliserende*.[19] For det andre: Gjennomføring av bilreiser forutsetter bil, dvs. et større utlegg, og gjennomføring av kollektivreiser forutsetter gyldig billett. Med andre ord er

som vil avgjøre hvilke endringer som vil skje. Men uansett vil trafikanten i den typen modeller vi bruker i Norge ikke ha andre interesser av rutevalget enn å minimere generaliserte reisekostnader.

[19] Dette er tilfelle hvis det er kø og trengsel. Hvis vi har en tilbudsmodell der kollektivselskapet bruker økte inntekter til å forbedre tilbudet, kan det også hende at kostnadene, for eksempel i form av ventetid mellom avgangene, blir *mindre* jo flere andre som reiser.

det mulig å *ekskludere* de som ikke betaler fra å forbruke godet reiser. Et offentlig prosjekt som forbedrer reisemulighetene er altså ikke det samme som å produsere mer av et offentlig gode i økonomisk forstand.

Dette har noen implikasjoner for hvordan vi tenker om samferdselsprosjekter. Av kostnader finns det minst tre slag: kostnader som det offentlige pådrar seg ved å produsere og drive infrastruktur og kollektive tilbud, kostnadene til private selskaper som tilbyr transporttjenester, og for det tredje de kostnadene som individene og husholdningene pådrar seg for å ta de offentlige og private tilbudene i bruk. Ressurser går med både på offentlig og privat side. Hvor mye ressurser som medgår, og hva de koster, vil påvirke hvor mye som tilbys av de private bedriftene og hvor mye som konsumeres. Noen av ressursene er markedsgoder, andre ikke.

Sammenliknet med standardsituasjonen i lærebøker i nyttekostnadsanalyse, er det mange flere ting å holde styr på. Standardsituasjonen er at det offentlige bekoster økt produksjon av et reint offentlig gode, som befolkningen kan konsumere uten å måtte betale eller ofre egne ressurser på annen måte.

Det finnes i regelen svært mange reisemarker å holde styr på – hver kombinasjon av et startsted, et bestemmelsessted, en reisemåte, en reisehensikt og kanskje også en tid på dagen er et eget reisemarked. Reiser i ulike markeder er delvis nære substitutter, slik at priser og andre kostnader i det ene reisemarkedet ofte vil inngå i etterspørselsfunksjonen etter reiser i et annet marked. Og reiser i ulike markeder påvirker hverandre gjennom at de bruker felles knappe ressurser, som gategrunn, T-baneavganger osv. Bruken av tid på bilreiser i en købelastet by er derfor en spillteoretisk likevekt der ingen vil angre på sitt rutevalg gjennom byen gitt at de andre heller ikke angrer på sine valg.

Transportmodellen skal hjelpe oss til å holde oversikt over dette mylderet, og spesielt over hvordan det endrer seg og nye likevekter oppstår når det gjennomføres et infrastrukturprosjekt eller en annen endring av politikken. Nyttekostnadsanalysen bruker resultater fra transportmodellen. Det har vist seg hensiktsmessig å dele inn i fire segmenter som påvirkes av tiltaket, nemlig trafikantene, operatørselskapene, det offentlige og samfunnet for øvrig. Dette kalles *bruttometoden*.

*Trafikantene* får en endring i sitt konsumentoverskudd. For å beregne det på en riktig måte, må vi naturligvis bruke de prisene som de faktisk opplever. Bensinen må inngå med skatt og avgift, billetten må inngå med faktisk beløp, osv. *Operatørselskapene* vil ha inntekter i form av billetter, parkeringsavgifter osv., og kostnader ved å drive tilbudet. Når vi ser trafikantene og operatørene under ett, vil trafikantenes billettkostnader og operatørenes billettinntekter utlikne hverandre. Men det betyr slett ikke at vi kan stryke billettinntektene på begge sider, for de er med å bestemme, både hva trafikantene velger å gjøre og hva operatørene velger å gjøre. I en modell der aktørene tilpasser seg endringer i kostnader og inntekter, finns det ikke noe som heter ”reine overføringer”, og vi er nødt til å føre alle overføringer eksplisitt som utgift på den ene gruppas konto og inntekt på den andre gruppas konto. Det er bruttometoden.

Vi kan også betrakte trafikantenes betaling av billetter på en annen måte, nemlig som en *korrigering* fra opplevde kostnader til virkelige samfunnsøkonomiske kostnader, idet kostnadene som trafikantene betaler, blir *byttet ut* med kostnadene ved å drive transporttilbudet når vi summerer over de to gruppene. Men en slik summering over de to gruppene kan altså ikke gjennomføres før gruppene har foretatt sine tilpasninger!

Det offentlige har inntekter fra skatter og avgifter og kostnader til infrastruktur. Eventuelt har offentlig sektor også inntekter fra parkeringsselskaper og bomselskaper, og kostnader til subsidier av kollektivsystemet. Her kommer en ny grunn til at vi ikke kan eliminere overføringer, inn i bildet. Endringer i netto inn- og utbetalinger over offentlige kasser skal nemlig multipliseres med en *skyggepris på offentlige midler* (skattefaktor). Uten å føre overføringene eksplisitt som inntekt for den ene sektoren og utgift for den andre, vil vi ikke vite hva vi skal multiplisere med skyggeprisen.

Endelig vil vi føre endringer i ulykkes- og miljøkostnader på kontoen for "samfunnet for øvrig". Noen vil hevde at det er bare den delen av de eksterne kostnadene som ikke er internalisert gjennom avgifter, som skal føres inn i regnestykket. Men vi har allerede ført avgiftene som kostnad for trafikantene og inntekt for det offentlige. Det var nødvendig for å få et riktig grunnlag for å beregne trafikantenes tilpasninger og for å beregning av skyggekostnaden på offentlige midler. Så det er ingen tvil om at vi må føre hele ulykkes- og miljøkostnaden, ikke bare den delen som ikke er internalisert, på kontoen til "samfunnet for øvrig".

Det skulle ikke være vanskelig å skjønne disse reglene. Likevel opplever vi gjentatte ganger at økonomer som ikke til daglig jobber med transport, implisitt eller eksplisitt legger til grunn at trafikantene ikke tilpasser seg tiltaket, sjøl når det er helt urimelig å gjøre det, og at de eliminerer det de kaller "reine overføringer". Det gjelder for eksempel ECON Pöyry i nyttekostnadsanalysen av Regionpakke Bergen. De drøfter til og med TØI-rapport 797/2005, som behandler bruttometoden grundig, men de forstår visst ikke det de leser.

# 4 Typer av usikkerhet i nyttekostnadsanalyse i samferdselssektoren

Anta at vi skal nyttekostnadsberegne et tiltak i samferdselssektoren. Det kan være et infrastrukturtiltak eller et virkemiddel av annet slag. I utgangspunkt gjør vi våre beregninger for ett eller flere framtidsår med de enhetspriser, trafikkvolumer og virkninger av tiltaket som vi mener mest sannsynlig representerer virkeligheten i dette framtidsåret. Tilsvarende gjør vi beregninger med de mest sannsynlige verdiene i tilfelle tiltaket *ikke* gjennomføres. Nyttekostnadsanalysen er en sammenlikning mellom situasjonen med og uten tiltaket i et visst antall år framover, og en påfølgende sammenregning av årene ved hjelp av en kalkulasjonsrente.

En slik beregning er usikker – og det av mange forskjellige grunner. Vi kan skille mellom følgende typer av usikkerhet i det samfunnsøkonomiske regnestykket:

1. Teknologiusikkerhet, som er usikkerhet om hvordan den generelle teknologiske utviklingen i samfunnet virker inn på reisebehov og reisevaner, samt hvordan kjøretøyteknologi og drivstoffeffektivitet vil utvikle seg.

2. Demografisk usikkerhet, som er usikkerhet om befolkningsutviklingen i det geografiske området vi studerer og om befolkningens fordeling på soner, og videre om fordelingen på demografiske og sosioøkonomiske grupper, i den grad det har betydning for reiseaktiviteten og reisemønsteret.

3. Næringsusikkerhet, som er usikkerhet om den fortsatte eksistensen av konkrete bedrifter og (mer generelt) om antall arbeidsplasser i det geografiske området vi studerer, og om hvordan arbeidsplassene og andre målpunkter for reisevirksomheten fordeler seg på geografiske delområder (soner).

4. Politisk usikkerhet om hvordan nasjonale myndigheter, og eventuelt også andre beslutningstakere utenfor tiltaksområdet vårt, vil bruke slike virkemidler som skatter og avgifter, infrastrukturutbygging, restriksjoner og påbud for å styre trafikkutviklingen i vårt analyseområde eller andre faktorer av betydning for regnestykket. Hvis det ikke

hører inn under transportmodellen som er brukt, faller også usikkerhet om kollektiv-tilbudet (ruteopplegg, frekvens, priser) inn her.

5. Politisk usikkerhet om lokale myndigheters politikk, kanskje spesielt arealbruks-politikken.
6. Økonomisk usikkerhet, som er usikkerhet om inntektsutviklingen og utviklingen av relative priser, og om konsekvensene av dette for størrelser som inngår i beregningene.
7. Modellusikkerhet, som er usikkerhet om hvorvidt modellen som har vært brukt til å beregne endringer i aktørenes økonomiske tilpasning som følge av tiltaket, faktisk gir riktige resultater.
8. Usikkerhet om førerkortinnhavet og om bilholdet. Dette er bakgrunnsvariable til transportmodellen, normalt beregnet i en egen førmodell.
9. Parameterusikkerhet, som er usikkerhet om parametre til nytteberegningene og om enhetspriser på goder som ikke er markedsgoder, og om hvordan disse parametrene og enhetsprisene utvikler seg over tid. Eksempler er tidsverdier, enhetspriser på miljø-faktorer, kalkulasjonsrente.
10. Kostnadsusikkerhet, som er usikkerhet om hvor mye infrastrukturtiltaket vil koste å bygge, vedlikeholde og drive.
11. Usikkerhet om konjunkturene i det enkelte år, altså svingninger i forhold til en normal økonomisk veksttakt.

Den ellevte typen er åpenbart den systematiske risikoen, som dreier seg om *konjunktur-usikkerhet*. Den har vi egne metoder for å behandle, som vi allerede har sett. Vi skal komme tilbake til det i avsnitt 3.1.

De øvrige ti er usystematisk usikkerhet. De seks første (1-6) vil vi samlet kalle *scenario-usikkerhet*. De tre neste (7-9) vil vi kalle *metodeusikkerhet*. Den tiende inneholder både scenariousikkerhet og metodeusikkerhet, men denne usikkerheten handler om noe helt annet enn årlig netto nytte. Det kan analyseres helt uavhengig av de ni første typene av usikkerhet, siden kostnadsusikkerheten er helt oppløst før den økonomiske usikkerheten om nytte-virkningene av tiltaket begynner å melde seg. Hvis man bruker risikopremie i kalkulasjonsrenta, er det ikke på noen måte gitt at den skal være like stor for infrastrukturkostnadene som for årlig netto nytte. Det er helt ulike faktorer som bestemmer hvor usikkert hvert av disse to elementene er. Analysen krever en annen form for kompetanse om et helt annerledes marked. Derfor har vi skilt ut kostnadsusikkerheten som en egen form.

Analysen av de seks første usystematiske formene for usikkerhet vil i første rekke ha som mål å innhente kunnskap som kan avklare usikkerheten, å vurdere om prosjektet bør utsettes til usikkerheten er avklart, identifisere gjenværende risikofaktorer som en må være særlig opp-merksom på i gjennomføringen av prosjektet, og å søke å finne måter å utforme prosjektet på som reduserer disse risikofaktorene eller gjør at de får mindre betydning.

Bortsett fra modellusikkerheten og usikkerheten om bilholdet, som påvirkes av nær sagt alle de andre faktorene, kan vi grovt sett regne alle andre former for usystematisk usikkerhet som *uavhengige* kilder til usikkerhet om årlig netto nytte, sjøl om det nok er visse forbindelser mellom dem.

Problemet vi har i forbindelse med usikkerhetsanalysen, er at scenariousikkerheten reflekteres i modellen på en måte som gjør det både tungvint og kostbart å undersøke virkningen av at scenarioparametrene antar andre verdier enn de vi har forutsatt. Et ytterligere problem er at scenarioparametrene av og til påvirker flere faktorer i nyttekostnadsanalysen samtidig, hvilket medfører at forventningsverdier ikke blir et så enkelt begrep som de offisielle veglederne forutsetter.

I tillegg til at scenariousikkerheten påvirker modellresultatet, vil naturligvis også metode-usikkerheten gjøre det. Disse to formene for usikkerhet kan betraktes som uavhengige kilder til den samlede usikkerheten, og de må analyseres og kontrolleres med ulike metoder.

*Nyttekostnadsanalysen* har sine egne usikkerhetsfaktorer, og blant de viktigste er kalkulasjons-renta, risikopremien i kalkulasjonsrenta, skattefaktoren, investeringskostnaden, infrastruktur-ens levetid.

I neste kapittel skal vi behandle metodeusikkerheten spesielt. Vi anfører bare her at tidsutvik-lingen til de såkalte enhetsprisene (blant annet på tid, ulykker, miljø og kjørekostnadens bestanddeler) til en viss grad vil være påvirket av inntektsutviklingen og andre scenariofaktorer. Usikkerhet om enhetsprisene vil igjen forplante seg til usikkerhet om de generaliserte kostnad-ene i modellen og de eksterne kostnadene i nyttekostnadsregnestykket. Dette er trolig den sammenhengen som det er vanskeligst å håndtere på en konsistent og forsvarlig måte i en usik-kerhetsanalyse, både fordi vi veit lite om tidsutviklingen av enhetsprisene og i hvilken grad den avhenger av inntekstutviklingen, og fordi det krever omfattende inngrep i modellen for å endre de generaliserte kostnadene og deres bestanddeler. Vi veit heller ingenting om i hvilken grad denne typen av ”tukling” med modellen er teoretisk forsvarlig.

## 4.1 Systematisk usikkerhet

Samferdselsdepartementets fastsettelse av risikopremien i kalkulasjonsrenta til 2,5 prosent bygger som nevnt på Minken (2005). For å vurdere hva som er med i systematisk risiko og hva som ikke er med, er det verdt å se nærmere på modellen og estimeringsmetoden i Minken (2005). Modellen sier at en kan investere i et sikkert objekt, to verdipapirer (som til sammen representerer børsen), egen kompetanse (humankapital) og samferdselsinfrastruktur. Samferd-selsinfrastruktur har en avkastning som avhenger av de to usikre faktorene *tidsverdien* og *trafikkvolumet*. Avkastningen er faktisk antatt å være proporsjonal med trafikkvolumet for gitt tidsverdi og proporsjonal med tidsverdien for gitt trafikkvolum, altså av forma $kva$, der $k$ er et mål på effekten av tiltaket, $v$ er tidsverdien og $a$ er trafikkvolumet. Tidsverdien antas på sin side å utvikle seg proporsjonalt med timelønna $w$. På grunn av forutsetningene vil konstanten $k$ ikke spille noen rolle og er satt til 1.

Årlig statistikk er brukt til estimeringen. De tjue årene med data som er brukt, kan deles inn i fire klasser: år der verdensøkonomien er god men oljeprisen lav, år der oljeprisen er høy men verdensøkonomien dårlig, år med høy oljepris og god verdensøkonomi og år med lav oljepris og dårlig verdensøkonomi. Det er så gjort antakelse om sannsynlighetsfordelingen over slike år i framtida. For at datamaterialet skulle representere fire mulige, men ulike sannsynlige fram-tidsår, måtte man ta vekk all trendmessig utvikling fra materialet.

Resultatet er en modell som fanger opp usikkerhet om avkastningen på transportprosjekter i den grad denne usikkerheten skyldes at konjunkturutviklingen på kort sikt virker inn på timelønn og trafikkvolumer.

Dette er en bedre modell enn CAPM når det gjelder infrastrukturprosjekter, for i motsetning til CAPM forutsetter den slett ikke at alle verdipapirer og formuesgjenstander er omsettelige. Men den har den svakheten at det er en toperiodemodell – de eneste tidspunkter som finns er nå og seinere. Den skiller derfor ikke mellom kort og langt sikt, sjøl om estimeringen er gjort på en måte som gjør at vi kan si at den handler om kortsiktige svingninger. Det er derfor litt uklart om vi kan behandle den langsiktige økonomiske trenden som en usystematisk risiko som vi må ta hensyn til, sjøl etter at vi har lagt inn en risikopremie i kalkulasjonsrenta.

Det er også en svakhet at man har forutsatt at nytteelementene utvikler seg over tid i takt med timelønnsutviklinga. Det teoretiske og empiriske belegget for det er faktisk svakere enn antatt i Minken (2005), se Minken (2011). Jo svakere sammenhengen er mellom tidsverdien (og kjørekostnader, ulykkes- og miljøverdier) på den ene sida og generelt privat konsum på den andre, jo mindre skal risikopremien i kalkulasjonsrenta være. Kanskje burde altså risikopremien vært redusert.[20]

## 4.2 Vanlig praksis

For å si det litt stygt, er vanlig praksis å understreke verbalt at beregningene er beheftet med usikkerhet, men at tid og ressurser ikke har tillatt en omfattende usikkerhetsanalyse. I beste fall gjøres det følsomhetsanalyse av følgende faktorer:

1. Anleggskostnader. Hvor høye kan anleggskostnadene være uten at prosjektet blir ulønnsomt?
2. Kalkulasjonsrenta. Hva er den renta som gir prosjektet nåverdi lik null?
3. Trafikkutviklingen. Man kan eksperimentere med ulike årlige vekstrater for trafikkutviklingen i studieområdet i nullalternativet, og/eller med ulik styrke på de umiddelbare effektene av tiltaket. Dette kan gjøres ved å endre data i nyttekostnadsberegningen, uten å endre data som går inn i eller kommer ut av transportmodellen. Det er klart at dette ikke er den ideelle måten å gjøre det på, men det er den enkleste.

Blant de tingene som gjør det vanskelig å gjøre noe mer enn dette, er at det slett ikke er lett å angi sannsynlighetsfordelinger for noe av det som kommer ut av modellen. Ut fra måten modellen og parametrene er estimert på, kan man vel si at de verdiene vi bruker og får ut, er *de mest sannsynlige*. Det er også det vi sa i innledningen til dette kapitlet. Men vi kan ikke si om de også er forventningsverdier, eller hvor mye forskjellige fra den mest sannsynlige verdien de kunne vært.[21] I tillegg til problemet med å si noe fornuftig om sannsynlighetsfordelinger, har vi de praktiske problemene med å endre modellen slik at den bygger på andre forutsetninger.

---

[20] Dette framgår av formel (32) i Minken (2005). Det har i det seinere vært brukt som argument for å sette ned risikopremien i forhold til anbefalingen i 2006 at det er feil å bruke gjennomsnittsavkastningen på børsen som grunnlag for å estimere beta i transport. En mye lavere beta burde vært brukt, siden trafikken varierer mye mindre med konjunkturene enn børskursene gjør. Minken (2005) viser at det er feil: Gjennomsnittsavkastningen på børsen har sin rettmessige plass i formelen sjøl om den dreier seg om avkastningen på uomsettelige objekter. I den grad avkastningskrevet på børsen er skjerpet etter finanskrisa, vil det motvirke en nedjustering av risikopremien i transport på grunn av en svakere enn antatt sammenheng mellom tidsverdi og timelønn.

[21] Muligens kunne vi gå tilbake til estimeringen av noen av parametrene og plukke opp konfidensintervaller derfra, men konfidensintervaller har ikke vært publisert fra verdsettingsstudiene til nå. Og når det gjelder for eksempel tidsverdiene, så er de estimert med ikke-parametriske metoder, slik at konfidensintervaller ikke kan beregnes. Å si noe mer om sannsynlighetsfordelingen til parametrene fordrer faktisk et nytt

Det beste vi kan gjøre hvis vi vil gjøre andre følsomhetsanalyser enn de tre vi har angitt, er nok derfor å gjøre endringene i nyttekostnadsverktøyet aleine, og se bort fra inkonsistensen med modellsystemet som det skaper.

# 5 Modellusikkerhet

Alle nyttekostnadsanalyser av samferdselsprosjekter bygger på en eller annen transportmodell. Til grunn for modellen ligger forutsetninger. For eksempel forutsettes det når man bruker *EFFEKT* at trafikantene ikke endrer lokalisering, bilhold, reisehyppighet, destinasjon og reisemåte som følge av prosjektet, men de kan eventuelt endre reiserute. Når man bruker *RTM* forutsettes normalt at trafikantene ikke endrer lokalisering eller bilhold som følge av tiltaket, men de kan endre reisehyppighet, destinasjon og reisemåte. Dersom tiltaket er i en by hvor det kan forekomme alvorlige køproblemer – nå eller i de framtidsårene man vil si noe om – bør valget av reiserute bygge på prinsippet om brukerlikevekt, men om det ikke er køer, vil rutevalget i modellen bygge på prinsippet om raskeste veg eller laveste generaliserte kostnad.

Skal vi bedømme usikkerheten av en gjennomført beregning, må vi altså først kontrollere om forutsetningene som er lagt til grunn i modellen, er realistiske i det konkrete tilfellet. Jo lengre fram i tid vi ser, jo mer urealistisk blir det å anta at det ikke er køproblemer, slik alle offisielle norske modeller har gjort til nå, med unntak av noen modeller for Osloområdet. Slik modellsituasjonen er nå, burde seriøse utredere avholde seg fra å spå noenting som helst om bytrafikken mer enn tjue år fram i tid. Først må det bygges opp en realistisk modell med kø!

Men på den andre sida er det heller ikke realistisk å spå om enorme køproblemer om tjue år, slik en lett kan komme til å gjøre i nullalternativet om en bruker en modell som har kø, men som ikke tar hensyn til at når køene blir store nok, vil folk finne på noe annet, for eksempel flytte, arbeide hjemmefra eller i alle fall reise mindre med bil.

Om forutsetningene er sjekket og funnet å være realistiske, blir det neste spørsmålet om modellen er på høyde med gjeldende teori på området, og om den er *validert*. Med validering mener vi to ting som må gjøres. Først en undersøkelse av om de estimerte tidsverdiene i modellen er realistiske og om de implisitte elastisitetene med hensyn på tid og kostnad samsvarer med gjengs kunnskap. Dernest en undersøkelse om modellen faktisk spår riktig når den brukes på et annet datamateriale enn den som den er estimert på.

Validering tas det stort sett for lett på her til lands. Nye modeller brukes nesten før de er ferdige, og rettes opp gradvis der hvor man finner at de virker feil. Gjennom denne gradvise tilpasningen til virkeligheten drukner gjerne også det teoretiske grunnlaget for modellen i pragmatiske og usystematiske ad hoc-løsninger.

En kvalitetssikring som ikke bedømmer om modellen holder mål og om forutsetningene for å bruke den er tilstede, kan ikke skille mellom usikkerhet som skyldes verktøyet og usikkerhet som skyldes virkeligheten, og vil derfor medføre at ingen konklusjoner fra usikkerhetsanalysen blir til å stole på.

Analysen av usikkerheten ved modellen og bilholdet kan som regel ikke gjøres i det enkelte prosjektet, men må være en mer grunnleggende aktivitet med sikte på å forbedre, oppgradere, reestimere og validere modellsystemet. Det vil som regel ikke være mulig å endre de viktigste

oppfølgingsprosjekt til den verdsettingsstudien fra 2010. Å si noe om sannsynlighetsfordelingen til modellresultatene er enda vanskeligere, for det krever at vi systematisk forandrer hele modellen.

parametrene i modellsystemet uten i forbindelse med et slikt større forbedringsarbeid. En del andre variable kan endres fra kjøring til kjøring, men det krever som regel et visst kodearbeid. Dette er ting som de generelle offisielle veglederne ikke tar høyde for, trolig fordi de ser bort fra at vi i transport anvender et større modellsystem for å gi inndata til analysen. Men om vi ikke kan gjøre enkle følsomhetsanalyser av variable som inngår fast i modellsystemet, vil det likevel være viktig å tenke gjennom hva ved modellen som gir opphav til usikkerhet og tvil om resultatet, og gjøre en innsats for å redusere disse formene for usikkerhet.

# 6 Usikkerheten i nytteberegningens ulike ledd

I avsnitt 2.3 så vi at nyttekostnadsregnestykket er summen av trafikantnytte (konsumentoverskudd), operatørnytte (produsentoverskudd), nytte for det offentlige (budsjettvirkning) og nytte for samfunnet for øvrig. I kapittel 3 så vi hva slags usikkerhet som fantes i regnestykket, og i kapittel 4 gikk vi nærmere inn på ett av aspektene ved metodeusikkerheten, nemlig modellusikkerheten. Vi skal nå vise hvordan de ulike usikre faktorene inngår i hvert av leddene som nyttekostnadsregnestykket består av. Hensikten er å danne et bilde av usikkerheten dersom ulike typer usikkerhet virker sammen.

Usikkerhet i scenariovariablene vil forplante seg til usikkerhet i modellresultatet og nytteberegningen. En annen og helt annerledes usikkerhet, som også medvirker til den totale usikkerheten i nyttekostnadsregnestykket, er modellusikkerheten, dvs. usikkerheten om modellen gjengir virkningene av tiltaket på en god måte (kapittel 4). Disse to typene av usikkerhet kan nok påvirke hverandre, i den forstand at modellen fungerer bedre eller dårligere for ett sett av scenariovariable enn for et annet. Men sjøl om det absolutt går an å tenke rundt slike ting som modellens gyldighetsområde, vil det være mye vanskeligere å kvantifisere det, dvs. å angi hvor mye mer usikkert modellresultatet blir når scenarioet endrer seg.

Vi skal derfor legge følgende trinnvise framgangsmåte til grunn for usikkerhetsanalysen:

1. Det beste vi kan gjøre med modellusikkerheten er å minimere den gjennom et arbeid på forhånd for å gjøre modellen best mulig.

2. Konjunkturusikkerhet kan slå ut i både markedspriser og volumer. Den er imidlertid håndtert ved et standardisert risikotillegg til kalkulasjonsrenta eller ved en standardisert form for beregning av sikkerhetsekvivalenter.

3. Deretter kan det gjennomføres en kvantitativ undersøkelse av den usystematiske usikkerheten som stammer fra scenariovariablene, med forutsetning om at modellen i seg sjøl er sikker.

4. Til slutt kan man om ønskelig problematisere modellresultatene ved å endre virkningene av tiltaket direkte i nyttekostnadsregnestykket. Man kan for eksempel anta at tidsbesparelsen blir større eller mindre enn beregnet av modellen, eller at tiltaket påvirker reisemiddelfordelingen sterkere eller svakere enn beregnet i modellen. Antakelig er den beste måten å gjøre det på å endre virkningene så mye at tiltaket får lønnsomhet lik null, og så spørre om det er mulig at dette kan skje.

Trinn 1 er drøftet i kapittel 4. Trinn 2 er drøftet i avsnitt 3.1. Trinn 4 er det ikke mye mer å si om. Resten av kapitlet vil derfor handle om trinn 3. For å kunne komme videre i analysen av scenariousikkerheten, trenger vi et formelverk, og det blir emnet i avsnitt 5.1.

## 6.1 Faktorene i nytteberegningens ulike ledd

Kall konsumentoverskuddet *B*, produsentoverskuddet (operatørnytten) *P*, den samfunnsøkonomiske verdien av det offentliges netto overskudd *F* og nytten for samfunnet for øvrig *E*. Skyggeprisen på offentlige midler kaller vi *S*. Vi bruker toppskrift 0 på variable uten tiltaket, og toppskrift 1 på variable med tiltaket, og ser for enkelhets skyld bort fra diskontering. Nytten av et tiltak, *V*, kan da skrives[22]

$$V = \left(B^1 - B^0\right) + \left(P^1 - P^0\right) - \left(1 + S\right)\left(F^1 - F^0\right) + \left(E^1 - E^0\right)$$

Vi skal videre se bort fra at det finnes mange ulike reisemarkeder. Etterspørselen *X* vil være en funksjon $X = D(G)$ av generaliserte reisekostnader *G*, som er summen av et pengeutlegg p (billett eller kjørekostnader) og en tidskostnad. Tidskostnaden er reisetida *t* multiplisert med en tidsverdi *v*.[23] Med andre ord:

$$G = p + vt$$

Det er vanlig å tilnærme konsumentoverskuddet med den såkalte trapesformelen[24], som ser slik ut i tilfellet med bare ett reisemarked:

$$\begin{aligned}\left(B^1 - B^0\right) &= \frac{1}{2}\left(G^0 - G^1\right)\left(X^0 + X^1\right) \\ &= \frac{1}{2}\left(p^0 - p^1\right)\left(X^0 + X^1\right) + \frac{1}{2}v\left(t^0 - t^1\right)\left(X^0 + X^1\right)\end{aligned}$$

Hvis vi hadde tatt hensyn til at det finnes flere reisemarkeder, ville konsumentoverskuddet bestått av en sum av mange ledd med akkurat denne forma. Hvis modellen har mange reisemarkeder, må trapesformelen ha like mange ledd. Å beregne konsumentoverskuddet med gjennomsnittskostnad over alle markeder og med totale trafikktall, gir nemlig ikke riktig resultat.

Hvis det finnes køer, er oppdelingen i siste linje av den generaliserte reisekostnaden *G* på tids- og pengekostnad faktisk ikke noe som kan gjøres entydig.[25] I brukerlikevekt kan det nemlig

---

22 Minken (2009).

23 Mer generelt vil vi ha en etterspørselsmatrise **X** = **D(G)**, der etterspørselen i hvert marked er en funksjon av generaliserte kostnader i eget og (i prinsippet) alle andre markeder.

24 En førsteordens taylorutvikling med restledd av den indirekte nyttefunksjonen til en representativ konsument.

25 I det generelle tilfellet med mange reisemarkeder vil kostnadsmatrisen **G** i prinsippet være en funksjon **G = S(X)** av etterspørselen i alle andre markeder. Med kø og trengsel vil denne funksjonen være konveks og tiltakende. I hvert punkt langs tilbudskurven $S_w(G_w, \mathbf{G_{-w}})$ i reisemarked w vil det herske brukerlikevekt, dvs. at ingen vil angre på *sitt rutevalg* gitt at ingen andre gjør det. Dermed er kostnadene **G** entydig fastlagt. I det spesielle punktet der etterspørselskurva og tilbudskurva møtes, dvs. **X = D(G)** og **G = S(X)**, hersker det i tillegg

være mer enn en rute i bruk for hver reiserelasjon. Alle ruter som er i bruk, vil ha samme generaliserte kostnad, men den ene ruta kan være lang og rask, mens den andre er kort og tar lang tid. Men dette er et problem vi kan se bort fra her.

Vi merker oss at konsumentoverskuddet er en sum av produkter av to eller tre faktorer. Den ene faktoren er en *differanse* mellom kostnaden i nullsituasjonen og den tilsvarende kostnaden i tiltakssituasjonen. Den andre er *et gjennomsnitt* av trafikkvolumet i nullsituasjonen og i tiltakssituasjonen. Den tredje, som bare opptrer i det siste leddet i formelen, er tidsverdien *v*.

Som en forenkling vil vi heretter tolke alle differanser mellom en variabelverdi i tiltaksalternativet og samme variabels verdi i nullalternativet enten som et resultat av tiltakets utforming og umiddelbare konsekvenser eller som et resultat av modellens virkemåte aleine. I usikkerhetsanalysens trinn 3 skal vi følgelig regne slike differanser som sikre. Her gjelder det pengeutlegget *p* og tidsbruken *t*, men som vi skal se, kan det også gjelde reisevolumet *X*.

Dette betyr at når det gjelder det første leddet i formelen, har scenariousikkerheten bare betydning i den grad den påvirker gjennomsnittsvolumet med og uten tiltaket. Når det gjelder det siste leddet, samvirker usikkerheten som påvirker tidsverdien og usikkerheten som påvirker gjennomsnittsvolumet. Tidsverdien kan dessuten påvirke volumet, slik at de to usikre faktorene har en kovarians. Som nevnt i avsnitt 3.1 kan det være uklart om samvariasjonen mellom den trendmessige utviklingen av tidsverdien og den trendmessige utviklingen av reisevolumet allerede er fanget opp i risikotillegget til kalkulasjonsrenta, men siden tidsverdien inngår som en parameter i funksjonsuttrykket for generalisert kostnad, og siden etterspørselen i transportmodellen er en funksjon av generalisert kostnad, virker det rimelig å ta hensyn til dette også ved vurdering av den usystematiske usikkerheten om inntektsvekst, tidsverdi og reiseaktivitet på langt sikt.

Nå over til produsentoverskuddet. For bilmarkedene eksisterer det ikke – bilistene produserer sin egen reise og betaler det den koster, slik at bortsett fra skatter og avgifter er det ingen forskjell mellom det de betaler og det reisa virkelig koster samfunnet. For kollektivreiser er produsentoverskuddet differansen mellom billettinntektene med og uten tiltaket, minus differansen mellom operatørkostnadene med og uten tiltaket. Hvis kostnadene er C, har vi:

$$\left(P^1 - P^0\right) = \left(p^1 X^1 - p^0 X^0\right) - \left(C^1 - C^0\right)$$

De differansene som opptrer her, er ikke differanser mellom to verdier av samme variabel, men involverer produktet av flere variable. Her har vi differansen mellom produktet av en prisvariabel og en volumvariabel i situasjonen med og uten tiltaket, minus en differanse mellom kostnadene med og uten tiltaket. Den sistnevnte differansen ser enklere ut, men også $C^0$ og $C^1$ vil i virkeligheten begge bestå av en enhetspris (på drivstoff, for eksempel) multiplisert med et volum. De ulike leddene i produsentoverskuddet består altså av en pris multiplisert med et volum. Volumet vil generelt være en funksjon av prisen.[26]

nok en likevekt, dvs. at ingen vil angre på at de besluttet seg for å reise, gitt at ingen andre gjør det. Ingen vil oppleve etter reisa at kostnadene blei annerledes enn det de trudde da de bestemte seg for å reise.

[26] Volumet X er en funksjon D(G) av G, og G er igjen en funksjon av p. Om kostnadene på innsatsfaktorene påvirker operatørenes bruk av innsatsfaktorer, er mer usikkert.

Også nytten for det offentlige og for samfunnet for øvrig vil bestå av ledd som er produkter av en pris og et volum, men når det gjelder usikkerhetsanalysen må vi skille mellom ledd der volumet er en funksjon av prisen (eller prisen er en funksjon av volumet) og ledd der det ikke er tilfelle.

Investeringskostnadene og kostnadene til drift og vedlikehold av infrastrukturen består av pris ganger volum, men volumet er bestemt av tiltakets utforming og ikke primært av prisen. Om disse to tingene er usikre stokastiske variable, så kan prisvariabelen og volumvariabelen som regel regnes som stokastisk uavhengige.

Annerledes med inntektene fra skatter og avgifter. Her vil volumet enten være turer i bestemte markeder, som i tilfellet med bompengeinntekter, eller knyttet til drivstofforbruket eller antall kjøretøykilometer. I alle disse tilfellene vil enhetsprisen (taksten, avgiften, skatten) være en av faktorene som bestemmer volumet.

Når det gjelder overføringer og subsidier til operatørselskapene, vil de ikke påvirke usikkerheten, men bare hvem som må bære den. Subsidiene overfører usikkerheten helt eller delvis fra operatøren til det offentlige.

Ser vi tilslutt på de eksterne kostnadene, så vil enhetskostnadene ikke påvirke volumet. Det ligger i ordet ekstern. Avgiften som skal internalisere den eksterne kostnaden, vil imidlertid påvirke volumet (det ligger i begrepet internalisering). Usikkerheten her er imidlertid i høy grad drevet av usikkerhet om hvor skadelig vedkommende eksterne virkning er, og av myndighetenes framtidige politikk med hensyn til grønne skatter.

Vi kan oppsummere vår drøfting av analysen av den usystematiske scenariousikkerheten (trinn 3) slik:

Vi forutsetter at modellusikkerheten og den systematiske usikkerheten er håndtert på forhånd, og at modellens virkninger nå er å betrakte som sikre. Vi har da noen ledd i nytteberegningen som bare er usikre på grunn av volumet, noen som er usikre på grunn av volumet og en prisvariabel som virker inn på volumet, og noen som er usikre på grunn av volumet og en prisvariabel som ikke innvirker på volumet. De ulike delene av regnestykket fordeler seg slik på de tre tilfellene:

- *Usikre på grunn av volumet*: prisdelen av konsumentoverskuddet
- *Usikre på grunn av volum og en pris som påvirker volumet*: tidsdelen av konsumentoverskuddet, produsentoverskuddet, det offentliges avgiftsinntekter
- *Usikre på grunn av volumet og en pris som er uavhengig av volumet*: Investeringskostnader, drift- og vedlikeholdskostnader ved infrastrukturen, de eksterne kostnadene og sannsynligvis driftskostnaden til operatørene.

## 6.2 Et problem og en omveg rundt det

Tatt i betraktning at de ulike variablene som inngår i scenariousikkerheten og parameterusikkerheten kan betraktes for uavhengig fordelte stokastiske variable, er usikkerhetsanalysen forholdsvis ukomplisert, med unntak av usikkerheten på grunn av volum og en pris som påvirker volumet (andre kulepunkt).

Leddene av denne typen har typisk forma $pD(p)$, der $p$ er en pris eller en parameter som inngår i prisen, og $D(p)$ er etterspørselsfunksjonen. Det er da problematisk både å beregne forventningsverdien og hele sannsynlighetsfordelingen til leddet.

Forventningen til leddet $pD(p)$ er ikke forventet $p$ multiplisert med forventet $D(p)$, men involverer også kovariansen mellom dem. Den generelle formelen er $E(XY) = EXEY + \text{cov}(X,Y)$. I vårt tilfelle:

$$E\left[pD(p)\right] = E[p]\cdot E\left[D(p)\right] + \text{cov}\left(p, D(p)\right)$$

Generelt kan kovariansen her ikke beregnes annet enn ved simulering. Men hvis vi kan anta at $p$ er normalfordelt med forventning $\mu$ og standardavvik $\sigma$, kan vi bruke Steins lemma, som i dette tilfellet gir:

$$E\left[pD(G)\right] = \mu\cdot E\left[D(G)\right] + \sigma^2 E\left[D'(G)\right]$$

Dette hjelper et stykke på veg, men forventet etterspørsel behøver ikke være etterspørselen vi får når vi setter inn forventet pris, så det trengs fremdeles simuleringer.

Når det gjelder usikkerhetsanalysen, dvs. hva som vil skje hvis variablene ikke har sine forventningsverdier, er det sannsynligvis simuleringer vi må stole på.

Kanskje kan en også ha bruk for en mer generell versjon av Steins lemma. La for eksempel $Y$ være en usikker enhetspris på miljø som man mener vil utvikle seg på en måte som avhenger av de samme faktorene som også påvirker en av de usikre parametrene Z i generaliserte kostnader. Generelt: La $Z$ og $Y$ være normalfordelt og $h$(Z) enn begrenset funksjon. Da er:

$$E\left[Yh(Z)\right] = E\left[Y\right]\cdot E\left[h(Z)\right] + \text{cov}(Y,Z)\cdot E\left[h'(Z)\right]$$

Imidlertid skal vi prøve å lure oss rundt mesteparten av problemet med ledd av forma $pD(p)$, der $p$ er en pris eller en parameter som inngår i prisen, og $D(p)$ er etterspørselsfunksjonen. Vi gjør det ved å addere konsumentoverskuddet og produsentoverskuddet:

$$\begin{aligned}
&\left(B^1 - B^0\right) + \left(P^1 - P^0\right) = \\
&\frac{1}{2}\left(p^0 - p^1\right)\left(X^0 + X^1\right) + \frac{1}{2}v\left(t^0 - t^1\right)\left(X^0 + X^1\right) + \left(p^1X^1 - p^0X^0\right) - \left(C^1 - C^0\right) = \\
&\frac{1}{2}v\left(t^0 - t^1\right)\left(X^0 + X^1\right) + \frac{1}{2}\left(p^0X^0 + p^0X^1 - p^1X^0 - p^1X^1 + 2p^1X^1 - 2p^0X^0\right) - \left(C^1 - C^0\right) = \\
&\frac{1}{2}v\left(t^0 - t^1\right)\left(X^0 + X^1\right) + \frac{1}{2}\left(p^0 + p^1\right)\left(X^1 - X^0\right) - \left(C^1 - C^0\right)
\end{aligned}$$

Av siste linje her ser vi at vi nå fortsatt har et problem med tidsdelen av konsumentoverskuddet, men prisdelen pluss operatørinntekten blir forvandlet til et usikkert gjennomsnitt av prisen før og etter, multiplisert med en prisendring som vi regner som sikker, fordi den enten er en direkte konsekvens av tiltaket eller et utslag som er beregnet i en kvalitetssikret modell.

Her har vi antatt at prisen i sin helhet er en overføring fra trafikanten til et operatørselskap. Hva skjer hvis bare en andel $\alpha$ av prisen trafikanten betaler, er en overføring til andre (operatører eller det offentlige)? Litt regning gir at

$$\left(B^1 - B^0\right) + \left(P^1 - P^0\right) =$$
$$\frac{1}{2}v\left(t^0 - t^1\right)\left(X^0 + X^1\right) + \frac{1}{2}\left(p^0 - (1-2\alpha)\,p^1\right)\left(X^1 - X^0\right) - \left(C^1 - C^0\right)$$

Vi ser at også usikkerheten i dette tilfellet er forenklet til noe som bare berører en av faktorene som inngår i den aktuelle delen av nyttekostnadsregnestykket. (Parameteren $\alpha$ må naturligvis anses kjent.) $\alpha$ kan for eksempel være billett, bompenger eller skatter og avgifter, multiplisert med skattefaktoren dersom det dreier seg om en overføring til det offentlige.

## 7 Konklusjon

Vi har undersøkt problemer knyttet til usikkerhetsanalyser i transportsektoren. Disse problemene er ignorert og ikke forstått av de som har laget de offisielle veglederne i samfunns-økonomisk analyse og usikkerhetsanalyse. Problemene knytter seg til at sjøl om det er mulig å tilordne subjektive sannsynligheter til fremtidige scenarioer, er det betydelig vanskeligere å knytte sannsynligheter til måten transportmodellen fungerer på og til det som kommer ut av modellen. Det er også tungvint eller umulig å endre noen av modellparametrene uten at modellen mister sine opprinnelige egenskaper. Problemene er tilnærmet løst ved en trinnvis framgangsmåte der vi

- anser modellusikkerheten som eliminert på forhåndgjennom systematisk validering,
- analyserer den systematiske risikoen på forhånd og inkluderer den i kalkulasjonsrenta eller enkle oppskrifter for beregning av sikkerhetsekvivalenter,
- grupperer leddene i nytteregnestykket slik at det fortrinnsvis bare er en faktor i leddet som betraktes som usikker,
- angir formler som kan forenkle eventuelle påkrevde simuleringer,
- åpner for å drive følsomhetsanalyse direkte på nytteregnestykket, uten å endre noe i modellen og uten å angi subjektive sannsynligheter.

Ved slike metoder kan vi komme lengre enn dagens praksis på transportområdet. Vi anbefaler at risikopremien i kalkulasjonsrenta estimeres på nytt, at man vurderer å endre modellen som blei brukt forrige gang slik at enhetsverdien ikke knyttes så nært opp til timelønna, og at det etableres en organisasjon som regelmessig ser gjennom og kvalitetssikrer modeller og verktøy og oppdaterer parametre som brukes i regnestykkene.

# Litteratur

Finansdepartementet (2005) Veileder i samfunnsøkonomiske analyser.

Finansdepartementet (2011) Rammeavtale om konsulenttjenester vedrørende kvalitetssikring av konseptvalg, samt styringsunderlag og kostnadsanslag for valgt konseptalternativ.[27]

Lund, D. (1987) Investing in non-marketable assets. Memorandum no. 2, February 19th, 1987. Department of Economics, University of Oslo.

Lund, D. (1993) Usikre investeringer under begrenset diversifisering. Beta 2/1993.

Minken, H. (2005) Nyttekostnadsanalyse i samferdselssektoren: Risikotillegget i kalkulasjonsrenta. TØI-rapport 796/2005.

Minken, H. (2009) Rammeverk for nyttekostnadsanalyse og finansieringsanalyse. Arbeidsdokument ØL/2156/2009.
(Inntatt som vedlegg 4 i Minken m.fl. (2009) Konseptvalgsutredninger og samfunnsøkonomiske analyser, TØI-rapport 1011/2009. Se også første artikkel i denne rapporten.

Minken, H. (2011) Tidsverdienes inntektsavhengighet og velferdsfunksjonens form. Arbeidsdokument ØL/2294/2011, TØI. (Andre artikkel i neste kapittel av denne rapporten.)

SSØ (2006) Behandling av usikkerhet i samfunnsøkonomiske analyser. www.sfso.no

SSØ (2010) Håndbok for samfunnsøkonomiske analyser. www.sfso.no

[27] Tilgjengeligheten av dette dokumentet er uklar.

## 2.5 Drivstoffavgift, bompenger og skattefinansiering[28]

[28] Dette upubliserte notatet er datert august 2008. Enkelte endringer er åpenbart gjort i 2016 eller 2017.

# Drivstoffavgift, bompenger og skattefinansiering


## Sammendrag

Problemet vi stiller oss her, er hvordan vi bør bruke prisvirkemidler som drivstoffavgifter, bompenger eller kilometerbasert vegprising dersom målet er å få best mulig samfunnsøkonomi i vegtransporten, og i hvilken grad et samfunnsøkonomisk riktig prissystem kan kombineres med finansiering av nye vegprosjekter.

For å besvare spørsmålene har vi formulert et enkelt matematisk optimeringsproblem. Ved hjelp av den har vi utviklet et kriterium som kan brukes til å avgjøre om det i det hele tatt er aktuelt å bruke bompenger. Deretter har vi utviklet en formel for hvordan drivstoffavgiften skal settes. På samme måte har vi også utviklet en formel for optimal bomavgift, gitt at drivstoffavgifta i utgangspunkt er satt optimalt. Samme formel, omformet til å gjelde pr. kilometer kjørt, kan brukes til å sette en optimal kilometerbasert vegpris. Vi drøfter i hvilken grad disse avgiftene kan differensieres etter kjøretøy, tid og sted.

Hvorvidt det finnes lokale investeringsprosjekter som trenger finansiering på de stedene der bommene står, er irrelevant for den optimale bomavgifta. Gitt at satsene skal settes samfunnsøkonomisk optimalt, blir det derfor meningsløst å fastsette på forhånd hvor stor del av det lokale prosjektet som skal finansieres med bompenger. Lovbestemmelsene om bompenger etter veglova bør flyttes til vegtrafikklova og samordnes med bestemmelsene om vegprising der.


# Drivstoffavgift, bompenger eller skattefinansiering?

## 1 Innledning

I skrivende stund er det 11 bomringsystemer i drift i Norge: Oslo, Bergen, Trondheim, Stavanger, Kristiansand, Namsos, Haugalandet, Grenland, Førde, Harstad og Bodø. Fem av dem (Trondheim, Kristiansand, Bergen, Grenland og Namsos) har tidsdifferensierte satser i en eller annen form.

I tillegg er det rundt 50 andre steder der bompenger finansierer større eller mindre deler av pågående eller fullførte vegprosjekter. 10-12 av disse er pakker med flere prosjekter. Forhåndsinnkreving av bompenger på ferja pågår på ytterligere 9 samband, og dessuten finansieres lokale samferdselstiltak i Tromsø med et lokalt tillegg i drivstoffavgifta.[29]

Drivstoffavgifter, bompenger og generelle skatteinntekter kan alle brukes til tre slags formål: reint fiskale formål, finansiering av konkrete investeringsprogrammer, og internalisering av ulike eksternaliteter.[30] Hva bør arbeidsdelingen mellom dem være? Dvs. hvordan skal hver av de tre formålene telle med når avgiftsbeløpet fastlegges? Det er dette vi vil drøfte i denne artikkelen. Vi utvikler kriterier for bruken av bomavgifter som finansieringsinstrument og viser

[29] For en oversikt over anleggene, se http://www.autopass.no/Betaling/veger-med-betaling-av-bompenger. Bomringsystemet i Oslo omfatter også bommer i Bærum, og systemet i Stavanger omfatter også Sandnes og andre kommuner på Nord-Jæren. Av pakker har vi bare tatt med de som kaller seg det.

[30] Vi bruker 'drivstoffavgift' som samlebetegnelse på mineraloljeavgift, dieselavgift og $CO_2$-avgift, som alle skrives ut med et visst kronebeløp pr. liter. Momselementet i drivstoffavgiftene er ikke medregnet.

hva som bør være rollefordelingen mellom drivstoffavgift på den ene sida og bomavgifter og kilometerbasert vegprising på den andre sida. Det vil vise seg at når skattekroner er dyre (eller når det eksisterer et samferdselsbudsjett med høyere skyggepris enn den generelle skyggeprisen på offentlige midler) skal både drivstoffavgift og bompenger, hver på sin måte, brukes til å internalisere eksternaliteter i transport. Hvorvidt det finnes lokale investeringsprosjekter som trenger finansiering på de stedene der bommene står, er derimot irrelevant for den optimale bomavgifta.

# 2 Litteratur

Den norske litteraturen på området har spesielt dreid seg om valget mellom bompenger og skattefinansiering når et vegprosjekt eller en prosjektpakke skal finansieres. Larsen (1986) er en nøkkelreferanse. Amdal m.fl. (2007) sier at bomfinansiering er å foretrekke framfor skattefinansiering dersom innkrevingskostnadene og avvisningskostnadene utgjør en mindre andel av netto bominntekter enn 20 prosent.[31] En tilsvarende tanke finns også i Welde (2005). Ramjerdi (1995) utvider problemstillingen fra å finne den mest kostnadseffektive finansieringsformen til også å inkludere internalisering av eksterne kostnader.

Lånekostnader har ikke vært trukket inn i denne litteraturen. Det skal heller ikke gjøres. I det offentlige ordskiftet er argumentet mot bompengefinansiering ofte at det medfører låneopptak til høy rente, mens det offentlige kan gi lavere rente. Men fra et samfunnsøkonomisk synspunkt koster kapitalen det samme uansett hvem som yter lånet. Om byggekostnaden, risikoen i prosjektet, bompengesatsene og innkrevingsperioden ikke påvirkes av hvem som yter lånet, er det altså irrelevant om det ytes av staten eller en privat finansinstitusjon, de samfunnsøkonomiske kostnadene bli de samme uansett (Eriksen m.fl. 2007, vedlegg 10). Det virkelige og viktige valget står mellom bompenger og skattefinansiering, ikke mellom ulike lånekilder.

Med en liten korreksjon kommer også vi fram til et tilsvarende resultat. Men dette resultatet utgjør bare et nødvendig, men ikke et tilstrekkelig krav til når det er fornuftig å bruke bompenger. Og sjøl om dette kravet er oppfylt, er det ikke dermed sagt at den satsen som velges, er den best mulige. Våre resultater tilsier jo at bompenger best kan brukes som en grov form for vegprising. Dermed sier vi i praksis at ingen av de rundt 50 enkeltstående bomprosjektene burde hatt krav på seg om å finansiere en bestemt del av et vegprosjekt eller en prosjektpakke på denne måten. Det samme gjelder nok de fleste bomringene. Finansieringskravet ødelegger for samfunnsøkonomien.

Resultatene tilsier også at lovbestemmelsene om bompenger etter veglova burde vært samordnet med bestemmelsen om vegprising i vegtrafikklova til én lov.

Bompengeprosjekter skal vedtas av Stortinget. I den forbindelsen gjelder det såkalte nytteprinsippet, som sier at de som betaler bompenger, også i størst mulig grad skal være de som får nytte av prosjektene, og at ingen som ikke får nytte av prosjektet, skal måtte betale for det. Våre resultater er ikke forenlig med nytteprinsippet i streng forstand. En mulig måte å bøte på dette, kan være å delfinansiere investeringsprosjekter med lokale skatter som ikke påvirker atferden i transportsystemet.

[31] 20 øre er den samfunnsøkonomiske kostnaden ved å bruke en skattekrone i Norge (FIN 2014).

## 3 Forutsetninger, notasjon og relasjoner

En grunnforutsetning i modellene nedafor er at trafikantene er like, i den forstand at de verdsetter spart reisetid likt. Denne forutsetningen står jo dessverre i grell kontrast til virkeligheten, men det er den som gjør det mulig å regne sammen pengeutlegg, kjørekostnader og tidsbruk på en entydig måte til en generalisert reisekostnad $g$, slik det er vanlig i transportøkonomi. Den aggregerte etterspørselen $x$ etter reiser med bil på strekningen vi betrakter, er en funksjon av generaliserte kostnader, $x = x(g)$.

Problemet med riktig bomavgift eller vegpris på en enkelt veg der det finnes omkjøringsmuligheter eller gode alternative reisemåter, stiller seg på en annen måte om vi opphever denne forutsetningen og antar at det finnes grupper av trafikanter med ulik tidsverdi. Det blir da et spørsmål om å gjennomføre tredjegrads prisdiskriminering på en måte som gjør velferden størst mulig (Verhoef og Small 1999, 2005, Chakirov 2016). Vi kommenterer ikke dette problemet her, men konsentrerer oss om enklere tilfeller.

Vi forutsetter også at alle kjøretøy er like, dvs. at de har samme drivstofforbruk pr. kilometer, samme utslipp pr. liter, samme ulykkesfrekvens, samt at de krever samme plass i trafikken og derfor har samme tendens til å forårsake kø for andre. Når det gjelder bompenger og kilometerbasert vegprising er dette ikke noen begrensende forutsetning. Vi kan jo i prinsipp beregne den optimale bompengesatsen for de ulike typene av kjøretøy hver for seg. Da kan vi også ta hensyn til at biler med ulik størrelse har ulik "køskapende evne" ved å regne om bilens størrelse til et antall personbilekvivalenter, og differensiere bomavgiftene etter antall personbilekvivalenter.

Når det gjelder drivstoffavgifter er det litt verre. Mange eksterne miljøkostnader (spesielt klimagassutslipp og luftforurensning) vil riktignok ha en nær sammenheng med drivstofforbruket, slik at de kan bli internalisert på en enhetlig måte gjennom drivstoffavgifta sjøl om kjøretøyene er ulike (Santos 2017). Men eksterne *køkostnader* har mindre med drivstoffforbruket å gjøre. En svak sammenheng er det likevel, siden større kjøretøy både skaper mer kø og bruker mer drivstoff.

Andre eksterne kostnader, som vegslitasje, ulykker og støy, påløper pr. kilometer, men har enda mindre med drivstofforbruket å gjøre, og vil derfor vil ikke kunne internaliseres gjennom drivstoffavgifter. Spesielt når det gjelder vegslitasje er bilenes størrelse eller vekt av største betydning.

Den helt vesentlige forskjellen mellom elbiler og hybridbiler på den ene sida og diesel- og bensinbiler på den andre, må håndteres med avgiftene på kjøp og eie av bil, og hører heller ikke med til reguleringen av *bruken* av bil, som er det vi behandler her.

I tillegg til å anta at alle biler er like, forutsetter vi også at alle turer er like lange, nemlig $a$ kilometer. Når det gjelder drivstoffavgifta og eventuelle kilometerbaserte avgifter, er problemet med det bare at trafikanter på langtur kan ha andre vurderinger og tilpasninger til avgiftene enn trafikanter på småturer, slik at den felles aggregerte etterspørselsfunksjonen $x(g)$ blir litt for enkel. Når det gjelder bomavgift stikker problemet dypere. I forhold til perfekt vegprising vil enhver bom eller bomring være en veldig grov måte å internalisere eksterne kostnader på, i det korte turer med små eksterne kostnader vil betale det samme som lange turer med store eksterne kostnader. Det er ikke en svakhet ved vår modellformulering, men ved virkemidlet.

Vi er nå klar til å formulere sammenhengene i modellen. Tabell 1 oppsummerer notasjonen.

*Tabell 1 Notasjon*

| | |
|---|---|
| $a$ | turlengde (km) |
| $b$ | bomavgift (kr) |
| $c$ | samlede bominnkrevingskostnader (kr) |
| $e$ | ekstern miljø- og ulykkeskostnad og vegslitasje (kr/km) |
| $f$ | den inverse av drivstoffeffektivitet (liter/km) |
| $g$ | generalisert reisekostnad (kr) |
| $I$ | infrastrukturinvestering (kr) |
| $m$ | momssatsen på drivstoff |
| $m_0$ | gjennomsnittlig momssats på alt forbruk |
| $p$ | drivstoffpris (kr/liter) |
| $r$ | budsjettvirkning, drivstoffavgift (kr) |
| $s$ | drivstoffavgift (kr) |
| $S$ | overføring fra det offentlige til bomselskapet (kr) |
| $t$ | reisetid pr. kilometer (timer) |
| $v$ | tidsverdi (kr/time) |
| $W$ | velferd (kr) |
| $x$ | etterspørsel (turer/time) |
| $\lambda$ | skyggepris på offentlige midler |

Literprisen på drivstoff, $p$, består av produksjonsprisen $p_0$ pluss drivstoffavgifta $s$, alt multiplisert med momsen:

(1) $$p = p(s) = (1+m)(p_0+s)$$

Vi forutsetter at bilistene tar hensyn til drivstofforbruket, men ikke de andre kilometerbaserte kjørekostnadene, når de bestemmer seg for å reise. Dette anses som en realistisk forutsetning. De øvrige kilometerbaserte kostnadene kan da behandles som eksterne kostnader.[32] De "atferdsrelevante" kjørekostnadene for en tur kan følgelig skrives *paf*. Tilsvarende kan tidskostnaden på en tur skrives *vat* og eventuelle direkte utlegg *b*. Generaliserte kostnader er da

(2) $$g = paf + vat + b$$

Når det verken finnes kø eller fartsgrenser, vil drivstofforbruket pr. kilometer være en u-formet funksjon av farta. Ved lav fart kreves det likevel drivstoff til å holde motoren i gang og til andre

[32] Virkelige kostnader som trafikantene ikke reagerer på, kan behandles som eksterne kostnader. Men siden trafikantene likevel før eller siden må betale disse kostnadene, har vi å gjøre med manglende rasjonalitet, ikke egentlig manglende kontroll over noen av kostnadene. Å forsøke å internalisere slike kostnader gjennom avgifter vil ikke løse problemet. Det eneste som kunne løse det, var om kostnadene (dekkslitasje, oljeskift, reparasjoner og bruksrelatert reduksjon av kjøretøyets salgsverdi, blant annet) måtte betales, eller i det minste synliggjøres, umiddelbart etter hvert som de oppstår. For enkelhets skyld har vi utelatt slike kostnader fullstendig fra modellen.

formål, som lys, oppvarming og aircondition. Drivstofforbruket pr. kilometer er da høyt, men synker raskt når farta øker. Minimum nås ved ca. 60 km/t for personbiler. Deretter vil luftmotstanden og friksjonen mellom dekk og underlag medføre en svak, men etter hvert raskere stigende *f*. Dersom det er kø i området vi betrakter, vil imidlertid farta ikke kunne velges fritt. Stadig oppbremsing og akselerasjon vil gi den u-formede kurva et skift oppover, og minimums-punktet vil flyttes mot høyre. I praksis vil farta under køforhold alltid være lavere enn den som gir minimum drivstofforbruk. Som funksjon av *trafikkvolumet* vil derfor *f* være tiltakende og konveks. Det samme gjelder reisetid pr. kilometer:

(3)
$$
\begin{aligned}
f &= f(x), \quad f' \geq 0,\ f'' \geq 0 \quad f(0) = f_0 \\
t &= t(x), \quad t' \geq 0, t'' \geq 0 \quad t(0) = t_0
\end{aligned}
$$

De aller fleste kjørte kilometer i Norge foregår på strekninger uten kø, der kjøretid og drivstofforbruk er bestemt av fartsgrensene og uavhengig av trafikkvolumet. Vi skriver *f* og *t* i det tilfellet som $f_0$ og $t_0$. Disse verdiene avhenger naturligvis av egenskaper ved kjøretøyet og vegen.

Det offentlige får inntekter av drivstoffavgifta, men samtidig vil penger brukt på bensin eller diesel innebære at det offentlige mister momsinntektene på forbruk som fortrenges. Det staten tjener på en liter bensin eller diesel er $mp_0 + (1+m)s$, og det den taper i form av momsinntekter på fortrengt forbruk er $m_0(1+m_0)^{-1}(1+m)(p_0+s)$. Budsjettvirkningen *r* for det offentlige blir da

(4)
$$r = r(s) = \frac{1}{1+m_0}\left[(1+m)s + (m-m_0)\,p_0\right]$$

Bortsett fra en litt annen notasjon er denne formelen for *r* identisk med tilsvarende formel i Minken og Samstad (2005).

De eksterne kostnadene pr. kilometer, *e*, vil bestå av en del som er proporsjonal med driv-stofforbruket ($CO_2$, $NO_x$), en del (ulykkeskostnader og støy) som avhenger på kompliserte måter av både trafikkvolumet *x* og farta (eller den inverse av farta, *t*), og en del som kan anses som konstant (vegslitasje). I tillegg kommer partikler fra diesel, som avhenger av drivstof-forbruket, og partikler fra oppvirvlet støv, som avhenger av farta. For vårt formål er det tilstrekkelig å skrive alt på redusert form:

(5)
$$e = e(x) = \psi(x, f(x), t(x)) \quad \left(\psi(0, f(0), t(0)) = e_0\right)$$

Parametrene i denne funksjonen vil avhenge av kjøretøyteknologien, vegenes standard og befolkningstettheten rundt. Funksjonen er derfor områdespesifikk.

Dersom vi anvender bompenger, påløper det innkrevingskostnader. Innkrevingskostnadene *c* består av en fast systemkostnad og en kostnad pr. passering:

$$c = c_0 + c_1 x$$

Samfunnsøkonomisk effektivitet, målt med en velferdsfunksjon, er kriteriet vi bruker for å avgjøre hva som er riktig måte å innkreve skatter og avgifter på, og hvordan vi bør finansiere et infrastrukturprosjekt. Velferdsfunksjonen vil måtte ta hensyn til kostnader og nytte for fire ”sektorer” – trafikantene, bomselskapet, det offentlige og samfunnet for øvrig, som jo utsettes for de eksterne kostnadene $e$.[33] Trafikantnytten er arealet under etterspørselskurven ned til prislinja (her: linja for generaliserte kostnader).[34] Hvis vi ser på en situasjon der en infrastrukturinvestering $I$ er gjennomført og et bomsystem som skal finansiere en andel $\alpha I$ av investeringen er etablert, vil bomselskapets overskudd være $P = (b - c_1)x - (c_0 + \alpha I) + S$, der $S$ er tilskudd fra det offentlige. (Når det gjelder bomselskaper er $S$ sikkert ikke positiv.) Det offentliges overskudd er $r(s)afx - S - (1-\alpha)I$. Vi kan trygt gå ut fra bomselskapet verken vil gå med underskudd eller overskudd, altså $P$ = 0. Når vi beregner S ut fra dette og setter inn i det offentliges budsjett, innebærer det at det offentliges budsjett blir $\left(r(s)af + b - c_1\right)x - (c_0 + I)$. Siden skattekroner er dyre skal det offentliges budsjett multipliseres med skyggeprisen på offentlige midler, 1 + $\lambda$, der $\lambda$ = 0,2 i følge retningslinjene fra Finansdepartementet. Velferdsfunksjonen blir derfor

(6) $$W = \int_g^\infty x(u)\,du + (1+\lambda)\left[\left(r(s)af + b - c_1\right)x - (c_0 + I)\right] - eax$$

Vår framgangsmåte for å finne optimal bruk av virkemidlene $s$ og $b$, og dermed indirekte også omfanget av skattefinansiering, er å maksimere $W$ gitt (1), (2) og (4), samt eventuelle andre bibetingelser. Dersom $b$ = 0, faller naturligvis innkrevingskostnadene bort. Alt etter omstendighetene kan vi også ha $I$ = 0, og $f$, $t$ og $e$ kan være funksjoner av $x$ eller konstanter, i henhold til (3) og (5). Virkemiddelet $b$ kan enten brukes i henhold til bestemmelsene om vegprising eller bestemmelsene om bompenger. I det sistnevnte tilfellet kommer det på et par ekstra restriksjoner som vi skal komme tilbake til.

# 4 Nødvendig betingelse for bruk av bompenger med en gitt sats

Uansett hvilket formål vi har med bompengeinnkrevingen og hvilket prinsipp vi bruker for å fastsette bomsatsen, må vi sikre oss at velferden er større ved bruk av bompenger med denne satsen enn ved å sette $b$ = 0. Som et minstekrav må vi altså sørge for at $W(b) \geq W(0)$. La $B(s)$ være mengden av de bomsatser $b$ som tilfredsstiller dette kravet ved drivstoffavgift $s$.

La $g_0$ og $x_0$ være generalisert kostnad og etterspørsel i situasjonen *uten* bompenger. I situasjonen *med* bompenger bruker vi ingen fotskrift på disse variablene. Vi har $g = g_0 + b$ og

$$W(0) = \int_{g_0}^\infty x(u)\,du + (1+\lambda)\left(rafx_0 - I\right) - eax_0$$

$$W(b) = \int_g^\infty x(u)\,du + (1+\lambda)\left(\left(raf + b - c_1\right)x - (c_0 + I)\right) - eax$$

Vi regner ut $W(b) - W(0)$, og får:

[33] Se Minken og Samstad (2005), eller eventuelt artikkel 2.1 i denne rapporten.

[34] Dette er den indirekte nyttefunksjonen til en representativ trafikant, gitt at alle trafikanter har indirekte nyttefunksjoner av forma $V(g) + R$, der $R$ er inntekt. Se Varian (1992), kapittel 7, eller første artikkel i TØI-rapport 1934/2023.

$$B(s) = \left\{ b \geq 0 \,\middle|\, W(b) - W(0) \geq 0 \right\}$$

(7)

$$= \left\{ b \geq 0 \,\middle|\, \int_{g_0}^{g} x(u)\,du + (1+\lambda)(c_0 + c_1 x) \leq (1+\lambda) bx + \left(ea - (1+\lambda) raf\right)(x_0 - x) \right\}$$

Integralet på venstresida av ulikheten er nyttetapet for trafikantene. Ulikheten sier altså at kostnadene ved bompenger, i form av redusert nytte for trafikantene og den samfunnsøkonomiske verdien av innkrevingskostnadene, er mindre enn den samfunnsøkonomiske verdien av bominntektene pluss de sparte eksterne kostnadene, i den grad de ikke allerede er dekket av den samfunnsøkonomiske verdien av inntektene fra drivstoffavgifta.

**Setning 1**

Samfunnsøkonomisk bruk av bompenger krever at $b \notin B(s) \Rightarrow b = 0$.

*Bevis:* Hvis $b \notin B(s)$ så er $W(0) > W(b)$.

Det er viktig å være oppmerksom på at setning 1 er et minstekrav, men ikke en tilstrekkelig betingelse for samfunnsøkonomisk bompengebruk. Sjøl om $b \in B(s)$ er det på ingen måte sikkert at vi har funnet den samfunnsøkonomisk beste bompengesatsen. Det er heller ikke sagt at det ikke finnes bedre finansieringsformer enn bompenger, sjøl om kravet er oppfylt.

Ulikheten i (7) kan bearbeides videre. Hvis vi bruker trapesformelen som en tilnærming til integralet på venstre side, har vi

$$\int_{g_0}^{g} x(u)\,du \approx \frac{1}{2} b(x_0 + x) = bx + \frac{1}{2} b(x_0 - x)$$

Dermed har vi det tilnærmede kriteriet

(8)

$$b \in B(s) \Leftrightarrow$$

$$\frac{1}{2} b(x_0 - x) + (c_0 + c_1 x) \leq \lambda\left((b - c_1)x - c_0\right) + \left(ea - (1+\lambda) raf\right)(x_0 - x)$$

Det første leddet på venstresida er det såkalte dødvektstapet. Det tilnærmede kriteriet er direkte sammenliknbart med et tilsvarende kriterium i Amdal m.fl. (2007). Forskjellen er det siste leddet på høyresida, dvs. at Amdal m.fl. ikke tar hensyn til eksterne kostnader og det offentliges inntekter av drivstoffavgifta.

Når vi nå skal finne optimale avgifter, må vi kreve at enten er $b \in B(s)$, eller så er $b$ = 0. Dette er ekvivalent med å inkludere følgende bibetingelse i optimeringsproblemene:

(9)

$$b \cdot \left[ \int_{g_0}^{g} x(u)\,du - (1+\lambda)\left((b - c_1)x - c_0\right) - \left(ea - (1+\lambda) raf\right)\left(x(g_0) - x(g)\right) \right] \leq 0$$

# 5 Optimal drivstoffavgift

Unntatt i spesielle tilfeller, som i Tromsø, er det vanskelig gjennomførbart å bruke en geografisk differensiert drivstoffavgift. Drivstoffavgift er i sin natur et nasjonalt virkemiddel, som bør tilpasses gjennomsnittlige forhold. Vi antar at vi kan sette *f*, *t* og *e* til $f_0$, $t_0$ og $e_0$ i den forbindelsen. Av grunner som snart vil bli klare, maksimerer vi *W* både med hensyn på *s* og *b*. Både *s* og *b* må naturligvis være større eller lik null. Dersom *b* = 0, faller innkrevingskostnadene vekk, og det markerer vi ved å sette inn kroneckerdeltaet $\delta$ foran disse kostnadene, og definere $\delta$ til å være null når *b* = 0 og 1 ellers. Problemet blir da:

$$\max_{g,s,b} W = \int_g^\infty x(u)\,du + (1+\lambda)\Big[\big(r(s)af_0 + b - \delta c_1\big)x - \big(\delta c_0 + I\big)\Big] - e_0 ax$$

(10)

$$\text{gitt} \quad p(s)af_0 + vat_0 + b = g \quad (\mu)$$
$$g > 0,\ s \geq 0,\ b \in B(s) \cup \{0\}$$
$$b = 0 \Leftrightarrow \delta = 0$$

Parameteren $\mu$ er Lagrangeparameteren tilknyttet den første bibetingelsen. *r*(*s*) er gitt ved (4) og *p*(*s*) ved (1). Vi danner Lagrangefunksjonen *L*. Kuhn-Tuckerbetingelsene for løsningen i dette tilfellet med positive valgvariable blir:

(11) $$\frac{\partial L}{\partial s} = (1+\lambda)\frac{1+m}{1+m_0}af_0 x - \mu(1+m)af_0 \leq 0 \quad (= 0 \text{ for } s > 0)$$

(12) $$\frac{\partial L}{\partial b} = (1+\lambda)x - \mu \leq 0 \quad (= 0 \text{ for } b > 0)$$

(13) $$\frac{\partial L}{\partial g} = -x + (1+\lambda)\big(r(s)af_0 + b - \delta c_1\big)\frac{\partial x}{\partial g} - e_0 a\frac{\partial x}{\partial g} + \mu = 0$$

I utgangspunktet finnes det fire muligheter: Både *s* og *b* kan være større enn null, en av dem kan være større enn null, eller ingen av dem.

Hvis både *s* og *b* er større enn null, gjelder likhetstegnet i både (11) og (12). Men det gir opphav til en motsigelse ($\mu$ kan ikke anta to ulike verdier samtidig). Og hele poenget med å prøve om det var bruk for begge virkemidlene var nettopp dette: å vise at det umulig kan gi en optimal løsning. Til å internalisere de eksterne kostnadene som bilkjøring medfører uansett hvor og når den finner sted, trenger vi i høyden ett virkemiddel – det mest effektive av de to.

Så hva er best av *s* og *b*? Ved å sammenlikne verdien av *W* når *s* brukes optimalt med verdien av *W* når *b* brukes optimalt kan vi sikkert utlede en regel for dette. Men det er neppe nødvendig. Det er utenkelig at en kostnadsfri innkreving pr. kilometer (dvs. drivstoffavgift) skal være dårligere enn et dyrt og mindre finmasket system av bommer over hele landet. Vi konsentrerer oss derfor om å finne den beste måten å bruke *s* på, og ser derfor nå bort fra relasjon (12). Samtidig kan vi anta at $\delta = 0$, og utelate *b*, $c_1$ og $c_0$ fra likningene for *W* og *g* og fra likning (13).

Vi har da to tilfeller igjen å se på, nemlig *b* = 0 og *s* > 0 og *b* = *s* = 0. Når *s* > 0 har vi likhet i (11), og finner $\mu = (1+\lambda)(1+m_0)^{-1}x$. Innsatt i (13) gir det

(14) $$(1+\lambda)r(s)af_0 = ea + \left(1 - \frac{1+\lambda}{1+m_0}\right)\frac{x}{\partial x/\partial g} = ea + \left(1 - \frac{1+\lambda}{1+m_0}\right)\frac{g}{El_g x}$$

Likning (14) balanserer tre effekter av en marginal økning av *s* mot hverandre – verdien for samfunnet av de økte avgiftsinntektene, nyttetapet for trafikantene, og de unngåtte eksterne kostnadene. Den sier at vi skal øke *s* inntil det punkt hvor den marginale nytten for samfunnet av økte avgiftsinntekter motsvarer marginale eksterne kostnader pluss det marginale nyttetapet for trafikantene. Dette framgår imidlertid ikke helt klart av likningen slik den står.

Multipliserer vi begge sider av likningen med den deriverte av *x*, og flytter de eksterne kostnadene over på venstresida, framkommer et klarere bilde. Venstresida består da av virkningene som skyldes endringer i trafikkvolum, mens høyresida er de direkte virkningene av å øke generaliserte kostnader med én krone. Parentesen på høyresida viser at det gir et nyttetap for trafikantene på én krone pr. reise, men dette motvirkes av en inntektsøkning for det offentlige av omtrent samme størrelse (avhenger av parametrene $m_0$ og $\lambda$). Venstresida, på sin side, inneholder nå den samfunnsøkonomiske verdien av tapet av avgiftsinntekter på grunn av at reiser som ville ha vært gjennomført ved lavere avgift, innstilles, minus gevinsten som reisenedgangen gir for miljø og ulykker. I optimum skal disse to virkningene stå i et forhold til hverandre som bestemmes av forholdet mellom de to direkte virkningene på høyresida, dvs. av parametrene $m_0$ og $\lambda$.

Hvis $m_0 = \lambda$, blir høyresida lik null, og likningen sier at den samfunnsøkonomiske verdien av avgiftsinntektene pr. reise skal tilsvare de eksterne kostnadene pr. reise. Dette er en variant av prinsippet pris lik grensekostnad, men ”prisen” skal altså vurderes etter at det er tatt hensyn til skyggeprisen på offentlige midler og til at avgifter i transport fortrenger annet, momsbelagt forbruk.

Likning (14) har gitt et utgangspunkt for drøfting av hvilke faktorer som spiller inn for riktig fastsettelse av drivstoffavgifta. Vi kan imidlertid omforme den til en mer eksplisitt formel for *s*. Det gjør vi ved å sette inn det eksplisitte uttrykket for *g* i likning (2), bruke (1) og (4) og ordne. Vi finner:

$$s = \frac{\frac{1+m_0}{1+\lambda}\cdot\frac{e}{f_0} - \frac{m_0-\lambda}{1+\lambda}\frac{1}{El_g x}\cdot\frac{vt_0}{f_0}}{\left(1+m\right)\left(1-\frac{m_0-\lambda}{1+\lambda}\frac{1}{El_g x}\right)} - \frac{\frac{m-m_0}{1+m} - \frac{m_0-\lambda}{1+\lambda}\frac{1}{El_g x}}{1-\frac{m_0-\lambda}{1+\lambda}\frac{1}{El_g x}}p_0 \tag{15}$$

Denne formelen ser komplisert ut, men forenkles drastisk dersom en kan sette $m_0 = \lambda$. Vi ser for øvrig at når produksjonsprisen på bensin eller diesel øker, skal drivstoffavgifta reduseres, men med en relativt liten andel av produksjonsprisøkningen. Den skal også øke med de eksterne kostnadene og reduseres når drivstofforbruket pr. kilometer øker. Disse to justeringene skal være relativt kraftige.

Det gjenstår å undersøke om kanskje ikke bare *b*, men også *s*, skal settes til null i optimum. Av (8) og (9) kan vi da utlede

$$\mu \geq \max\left[\left(1+\lambda\right)x, \left(1+\lambda\right)\frac{1+m}{1+m_0}x\right] = \left(1+\lambda\right)\frac{1+m}{1+m_0}x$$

Bruker vi denne ulikheten på (10), kan vi finne følgende vilkår for at *s* skal være 0 i optimum:

$$\left(1+\lambda\right)r\left(0\right)af_0 - e_0 a \geq -\left(\left(1+\lambda\right)\frac{1+m}{1+m_0} - 1\right)\frac{g}{El_g x} \tag{16}$$

Setter vi inn realistiske verdier i (16), vil vi finne at dette vilkåret langt fra kan oppfylles under normale norske forhold. *Vi kan altså konkludere med at drivstoffavgifta skal settes i henhold til formel (15). Vi vil heretter kalle s som er satt på denne måten for* $s^*$.

# 6 Optimal tidsdifferensiert bomavgift (vegprising på bomring)

Med drivstoffavgifta fastlagt for å internalisere eksterne kostnader i et normaltilfelle på nasjonalt nivå, har vi framleis et virkemiddel *b* som vi kan bruke når de lokale forholdene avviker fra dette i form av kø eller uvanlig store miljø- og ulykkeskostnader, eller for å finansiere en lokal investering. Problemet er:

$$\underset{g,b}{Max}\, W = \int_g^{\infty} x(u)\,du + (1+\lambda)\Big[\big(r(s^*)af(x) + b - c_1\big)x - (c_0 + I)\Big] - e(x)ax$$

(17) gitt $p(s^*)af(x) + vat(x) + b = g \quad (\mu)$

$g > 0,\ b \geq 0$

Vi antar indre løsning. Førsteordensbetingelsene for maksimum er:

(18) $$\frac{\partial L}{\partial b} = (1+\lambda)x - \mu = 0$$

(19) $$\frac{\partial L}{\partial g} = -x + \Big[(1+\lambda)\big(raf + raxf' + b - c_1\big) - \big(ea + e'ax\big)\Big]\frac{\partial x}{\partial g} - \mu\big(paf' + vat'\big)\frac{\partial x}{\partial g} + \mu = 0$$

Vi bruker (18) til å eliminere $\mu$ fra (19). Deretter deler vi alle ledd med $(1+\lambda)\frac{\partial x}{\partial g}$ og multipliserer med *g* over og under brøkstreken i det ene leddet som nå inneholder $\frac{\partial x}{\partial g}$. Så danner vi elastisiteten av etterspørselen med hensyn på generaliserte kostnader som en faktor i dette leddet, og skriver den gjenværende faktoren *g* eksplisitt ut, slik som i likning (2). Til slutt ordner vi slik at leddet som inneholder *b* kommer på den ene sida og alle andre ledd på den andre, og omfatter noen flere ledd til elastisiteter. Den optimale bomavgifta kan da skrives:

(20) $$b^* = \frac{c_1 + \dfrac{ea(1+El_x e)}{1+\lambda} + \big(paf' + vat'\big)x - raf\,(1+El_x f) - \dfrac{\lambda}{1+\lambda}\dfrac{paf + vat}{El_g x}}{1 + \dfrac{\lambda}{1+\lambda}\dfrac{1}{El_g x}}$$

I (20) er nevneren mindre enn 1, så alle elementene over brøkstreken vil bli blåst opp med en viss faktor. Leddene over brøkstreken er i rekkefølge:

(1) de marginale innkrevingskostnadene,
(2) de marginale eksterne kostnadene som påføres miljøet, de som rammes av trafikkulykker og aktører utafor transportsektoren,
(3) de marginale køkostnadene som påføres trafikantene i form av økt reisetid og økte kjørekostnader,
(4) de økte avgiftsinntektene som følger av en marginal økning av drivstofforbruket pr. km, og
(5) et ledd som internaliserer et nyttetap knyttet til skyggeprisen på offentlige midler.

En liten omorganisering av de midtre leddene i (20) bringer fram et ytterligere poeng:

(21) $$b^* = \frac{c_1 + \frac{ea\left(1+El_x e\right)}{1+\lambda} + \left(p-r\right)axf' + vaxt' - raf - \frac{\lambda}{1+\lambda}\frac{paf+vat}{El_g x}}{1+\frac{\lambda}{1+\lambda}\frac{1}{El_g x}}$$

Her ser vi (i tredje ledd) at internaliseringen av den marginale økningen i drivstoffkostnadene som påføres de andre trafikantene, delvis allerede har skjedd i form av avgiftsinntektene *r* som er innbakt i drivstoffprisen. Derfor skal ikke hele den marginale økningen i drivstoffutgiftene for andre trafikanter som følger med de økte køene, internaliseres på nytt gjennom bomavgifta *b*. På samme måte skal også provenyet av drivstofforbruket, *raf*, regnes med som dekning av marginale eksterne kostnader, slik at ikke bomavgifta behøver å gjøre hele denne jobben aleine.

Siden køforholdene endrer seg over døgnet og er ulike fra sted til sted, og siden uvanlig høye miljøkostnader også bare forekommer enkelte steder, vil bompenger satt etter (21) også variere med tid og sted. Vi snakker altså her om en form for vegprising – mer eller mindre grov og imperfekt alt etter bommenes antall og plassering. Åpenbart vil bompengene kunne finansiere en del av investeringsprogrammet *I*, men siden vi ikke har satt noe krav til finansieringsgraden, vil det være tilfeldig hvor stor denne delen er.

Vi legger merke til at verken investeringen *I* eller den faste kostnaden ved etablering av bomsystemet, $c_0$, inngår i (21). Det spiller altså ingen rolle for den optimale bomavgifta hvor stor investering som skal finansieres. Heller ikke den faste kostnaden ved å etablere bomsystemet har noe å si. Men dersom den sistnevnte kostnaden er stor nok, er det mulig at det ikke lønner seg å etablere bomsystemet i det hele tatt, sjøl ikke ved optimale satser. Før vi i det hele tatt etablerer bomsystemet, må vi altså bruke kriteriet i Setning nr.1 for å undersøke dette.

# 7 Optimal kilometerbasert vegprising

Det har nå blitt teknologisk mulig ved hjelp av GPS å drive vegprising i form av en avgift pr. kjørte kilometer innafor et vilkårlig avgrenset geografisk område i et vilkårlig avgrenset tidsrom. Vi kan anta at den marginale kostnaden ved å kreve inn avgifta, $c_1$, er null, dvs. at når systemet er etablert og den enkelte er registrert som bruker, vil det ikke koste mer å kreve inn avgifta for mange kilometer pr. måned enn for få.

Vi kan bruke formel (21) til å beregne riktig kilometersats også. Kall kilometersatsen *k* og sett $b = ka$. Vi ser at alle ledd i (21) er proporsjonale med *a*. Derfor vil den optimale kilometerbaserte vegprisen $k^*$ framkomme ved å dele med *a* på begge sider i (21):

(22) $$k^* = \frac{\frac{e\left(1+El_x e\right)}{1+\lambda} + \left(p-r\right)xf' + vxt' - rf - \frac{\lambda}{1+\lambda}\frac{pf+vt}{El_g x}}{1+\frac{\lambda}{1+\lambda}\frac{1}{El_g x}}$$

Fordelen med kilometerbasert avgift er for det første at alle som bidrar til kø og eksterne kostnader innafor et definert område, vil måtte betale. Det er altså ikke slik som ved en bomring, at store deler av trafikken slipper avgiften fordi de ikke krysser bomringen. Dessuten betaler de etter hvor mye de kjører, hvilket både er mer rettferdig og mer teoretisk riktig. Det blir ikke gratis å fortsette å påføre andre eksterne kostnader, slik det er for alle når de først har krysset bomringen.

Vi kan altså differensiere avgifta etter kjørelengde. Vi kan også differensiere etter kjøretøytype med samme formel, uten andre problemer enn at vi trenger å kunne identifisere kjøretøytypen for hver GPS-registering og kjenne kjøretøytypens drivstofforbruk, køskapende evne osv. Opplagt kan vi også differensiere etter tid på dagen.

Likevel finnes det et par problemer med å bruke en så enkel formel. Et av dem er at området vi avgrenser og tilordner et bestemt avgiftsopplegg, etter alt å dømme ikke vil være så homogent som vi antar. Det vil finnes veger med stor belastning og veger med liten belastning, veger i tett bebodde deler av området og veger i deler av området uten bebyggelse, og det vil finnes veger der trafikken skaper spesielle problemer, enten i form av barrierevirkninger, spesielt sårbare omgivelser eller på andre måter. Dette kan vi ikke komme helt vekk fra, for om vi definerer området for trangt, blir det for komplisert for trafikantene å tilpasse seg. Dessuten impliserer vår formel at vi ser på hele turer – dette er implisitt i etterspørselsfunksjonen $x(g)$ og i integralet som skal uttrykke trafikantnytten. Da bør området være stort nok til å romme de fleste hele turer.

Et annet, men beslektet problem med formel (22) er hvordan vi skal fastlegge $x$ som skal brukes i formelen. Transportarbeid eller trafikkarbeid, eller kanskje trafikkvolum på utvalgte vegstrekninger? Det er rimelig å forstå elastisiteten av $x$ med hensyn på $g$ som et spørsmål om hvordan antall turer endrer seg når generaliserte reisekostnader pr. tur endrer seg. Dette bør være en gjennomsnittlig sammenheng som er uavhengig av hvilken biltype som er brukt (så lenge det dreier seg om personbil). Men når det dreier seg om drivstofforbruk, køvirkning og miljøpåvirkning kan vi være interessert i å skille mellom biltyper, og operere med egne deriverte av $e$, $f$ og $t$ for hver biltype. Likevel er det den totale trafikken av alle biltyper som må utgjøre den $x$ som inngår i $f$ og $t$, i alle fall. Det er totaltrafikken, eventuelt vektet med biltype, som bestemmer farta, og dermed drivstofforbruket og reisetida. Vi antar at det ikke er for vanskelig å finne praktiske løsninger på dette problemet.

# 8 Tre metoder som kan kvalitetssikre hverandre

Det finnes to måter å beregne optimal kilometerbasert vegpris på. I tillegg til å estimere en funksjon som (22), kan vi også gjennomføre en serie av transportmodellkjøringer med tilhørende beregning av samfunnsøkonomisk lønnsomhet. Ved å gjøre dette på en systematisk måte vil det ikke være altfor komplisert eller ressurskrevende å finne en tilnærmet optimal løsning. Vi foreslår at man bruker begge metoder samtidig, i alle fall til å begynne med. Det er nemlig slik at transportmodell-tilnærmingen kan si oss noe om elastisitetene og funksjonsformene i likning (22) er fornuftige, mens omvendt formlene kan gi en kontroll på om modellen er fornuftig.

I tillegg er naturligvis ingenting så nyttig som virkelige eksperimenter. Minken 2015, kapittel 8 skisserer hvordan optimal avgift enklest kan undersøkes på denne måten.

# 9 Konklusjon

Problemet vi stiller oss i denne rapporten, er hvordan vi bør bruke prisvirkemidler som drivstoffavgifter, bompenger eller kilometerbasert vegprising dersom målet er å få best mulig samfunnsøkonomi i vegtransporten, og i hvilken grad et samfunnsøkonomisk riktig prissystem kan kombineres med finansiering av nye vegprosjekter. For å besvare spørsmålene har vi formulert et enkelt matematisk optimeringsproblem.

Vi finner at med mindre forholdene varierer fra det ene stedet til det andre, eller fra et tidspunkt til et annet, kan det aldri være riktig å bruke både drivstoffavgifter og bompenger. Av disse to vil det være drivstoffavgifta som er det mest effektive virkemidlet. Det bør derfor utformes for å internalisere de eksterne virkningene som eksisterer under normale omstendigheter i landet som helhet. Bompenger eller kilometerbasert vegprising kan da settes inn for å internalisere mer lokalt forekommende eksterne virkninger, som kø, særlig stor ulykkesrisiko eller særlig store støyproblemer eller miljøproblemer lokalt.

Det følger av dette at det ikke er fornuftig å kreve at bompengene på et bestemt sted skal finansiere en på forhånd satt andel av et investeringsprosjekt på samme sted. Om det ansees nødvendig av rettferdighetsgrunner at de som ikke får nytte av prosjektet heller ikke skal måtte betale for det (det såkalte nytteprinsippet), eller at de som får nytte av prosjektet også betaler noe for det, må man i stedet utforme lokale skatter eller avgifter som ikke påvirker atferden i transportsystemet. Det følger også at det ikke vil ha noen hensikt å ha to forskjellige lover for bompenger og vegprising. Alle bomprosjekter bør gå etter bestemmelsene i vegtrafikklova.

I rapporten har vi utviklet et kriterium som kan brukes til å avgjøre om det i det hele tatt er aktuelt å bruke bompenger. Dette er setning 1 i tilknytning til likning (7). Deretter har vi utviklet en formel for hvordan drivstoffavgiften skal settes. Dette er likning (15). På samme måte har vi utviklet en formel for optimal bomavgift, gitt at drivstoffavgifta i utgangspunkt er satt optimalt. Dette er likning (20). Samme formel, omformet til å gjelde pr. kilometer kjørt, kan brukes til å sette en optimal kilometerbasert vegpris. Dette er likning (22). Vi drøfter i hvilken grad disse avgiftene kan differensieres etter kjøretøy, tid og sted.

I praksis kan det være hensiktsmessig å bruke tre ulike metoder til å fastsette riktig vegpris, nemlig våre formler, optimering ved hjelp av en transportmodell, og virkelige eksperimenter. På den måten vil feil og begrensninger i hver av metodene kunne korrigeres.

# 10 Avsluttende betraktninger

### Drivstoffavgifta

Drivstoffavgifta, slik vi har definert den, består av to deler, mineraloljeavgift og $CO_2$-avgift, eventuelt dieselavgift og $CO_2$-avgift. Den første delen gis ofte en fiskal begrunnelse, mens den andre tilsynelatende er satt for å internalisere en av de eksterne kostnadene. Men navnet på avgiftene spiller naturligvis ingen rolle for hvordan de virker og om de er satt riktig. Det viktige er om de to avgiftene til sammen internaliserer de eksterne kostnadene som bilkjøring medfører under gjennomsnittlige norske forhold. Da må vi først vite hvor store de eksterne kostnadene er, og dernest må det finnes en politisk vilje til å innrette drivstoffavgifta etter det, og ikke etter tilfeldige fiskale hensyn eller politiske kjepphester.

Det er stor usikkerhet om kostnaden på $CO_2$-utslipp. Formodentlig bør den settes som kostnaden ved å oppnå målsetningen i klimameldingen, men denne målsetningen er foreløpig uklar, og det mangler studier av hvilken skyggepris på $CO_2$ den vil implisere. Når det gjelder de andre eksterne kostnadene, er det for øyeblikket Thune-Larsen m.fl. (2016) som har de mest oppdaterte estimatene, uttrykt som gjennomsnittssatser pr. kjøretøygruppe, og oppdelt på henholdsvis storby, tettbygd strøk og spredtbygde strøk der det er nødvendig. Men utlysningen av ny verdsettingsundersøkelse er på trappene. Men vi kan ikke vente at den nye undersøkelsen vil ta hensyn til skyggeprisen på offentlige midler og virkningene på offentlige budsjetter av at når husholdningene bruker mer penger på drivstoff, fortrenger det annet forbruk, slik vi gjør i den foreliggende rapporten. Derfor vil verdsettingsundersøkelsen måtte suppleres med funnene her når en skal utforme avgifter.

Vi ser at skyggeprisen på offentlige midler spiller en rolle både når det gjelder hvilken vekt som skal legges på de marginale eksterne miljø- og ulykkeskostnadene, og når det gjelder å avveie avvisningseffekten av høyere avgifter i transport mot effektivitetsvirkningen i resten av økonomien av å unngå dyrere og mer vridende former for beskatning. Men skyggeprisen på offentlige midler er i seg sjøl en svært usikker størrelse, som kunne fortjene nye studier. Spesielt ville det være fint om en i den forbindelsen kunne inkludere drivstoffavgift og bompenger blant de vridende skattene, ettersom det er åpenbart at transportkostnadene har noe å si for effektiviteten i bl.a. arbeidsmarkedet.

Endelig må det altså finnes en politisk vilje til å bruke drivstoffavgifta slik som våre formler tilsier. Det kan holde hardt.

**Bompenger**

Vi bør skille mellom to situasjoner, nemlig situasjonen hvor det er sikkert at en investering skal gjennomføres, men usikkert hvordan den skal finansieres, og situasjonen hvor en kan velge mellom å gjennomføre en investering ved hjelp av bompenger eller å la være å gjennomføre investeringen. Når bompenger har blitt så utbredt, er det fordi beslutningstakerne har opplevd at de var i den sistnevnte situasjonen.

Hva som er riktig å gjøre i den sistnevnte situasjonen, bør avgjøres av en nyttekostnadsanalyse. Det som avgjør resultatet, er da ikke bare størrelsen på bomavgifta og innkrevingskostnadene, men også størrelsen på investeringen og forbedringene den gir. Men metoden for en slik analyse kan ikke være vegvesenets tradisjonelle. Normalt regner nemlig vegvesenet med at trafikken vil være den samme med og uten tiltaket. Dermed ser man bort fra avvisningseffekten av bompengene, som kan være stor eller liten avhengig av om det finnes omkjøringsmuligheter eller andre gode alternativer til å bruke bil over bomsnittet. Følgelig trenges det en elastisk etterspørselsmodell og god rutevalgsmodell. Åpenbart må man også endre på den uheldige praksisen at man skal se bort fra innkrevingsmåten når man vurderer prosjekter i NTP-sammenheng. Slik denne praksisen har vært har vi fått for liten kunnskap om den virkelige lønnsomheten av bomprosjektene utenfor de store byene.

Kriteriet som fremmes i Amdal m.fl. (2007) er naturligvis ikke ment å avgjøre om et bomprosjekt er lønnsomt eller ikke, men om et vedtatt prosjekt bør finansieres av bompenger eller over skatteseddelen. Vi har utledet et kriterium for det samme som ser helt annerledes ut, hvilket må bety at kriteriet til Amdal m.fl. er feil. Ved nærmere ettersyn ser vi hvorfor. Den viktigste grunnen er at Amdal m.fl. tar ikke hensyn til den potensielt viktigste positive virkningen av bompenger, nemlig at de kan brukes til å internalisere eksterne køkostnader og mer enn gjennomsnittlig høye miljø- og ulykkeskostnader.

Vi for vår del finner at det eneste som kan forsvare bruk av bompenger framfor skattefinansiering, er nettopp eksistensen av køkostnader og over gjennomsnittlig høye miljø- og ulykkeskostnader. I det tilfellet skal bomavgifta brukes som en grov form for vegprising, og bomsystemet skal utformes slik at vegprisingen blir mest mulig effektiv (tidsdifferensierte satser, flest mulig av de som bidrar til kostnadene, og færrest mulig av de som ikke bidrar, skal betale, osv.). Hvor mye av en investering som kan finansieres ved dette, er uten betydning.

Dersom det allerede i utgangspunktet er bestemt at et visst beløp skal tas inn gjennom bompenger og brukes på et infrastrukturprosjekt, er det naturligvis irrelevant å stille spørsmålet om det ikke hadde vært bedre med skattefinansiering. Verken Amdals eller våre formler er relevante. Spørsmålet er utelukkende om prosjektet er lønnsomt under den gitte forutsetningen om bomfinansiering. Men vi må være klar over at med mindre satsene er nettopp slik at de eksterne kostnadene internaliseres, oppstår det et tap i forhold til om bomfinansieringskravet ikke hadde eksistert. Derfor bør praksisen med bomprosjekter etter veglova opphøre, og alle bomprosjekter bør gå etter bestemmelsene om vegprising i vegtrafikklova.

## Referanser


Amdal, E., G. Bårdsen, K. Johansen og M. Welde (2007) Operating costs in Norwegian toll companies: A panel data analysis. *Transportation* **34**(6), 681-695.

Chakirov, A. (2016) Urban mobility pricing with heterogeneous users. Doctoral Thesis at ETH Zurich, Diss. Eth. No 23 174.

Eriksen, K.S., Minken, H., Stenberg, G., Sunde, T. og Hagen, K.-E. (2007) Evaluering av OPS i vegsektoren. TØI-rapport 890/2007.

Finansdepartementet (2014) Prinsipper og krav ved utarbeidelsen av samfunnsøkonomiske analyser mv. Rundskriv R-109/2014.

Larsen, O.I. (1986) Bompenger som finansieringsform. *Sosialøkonomen* **40**(4), 9-11.

Minken, H. (2015) Samfunnsøkonomisk ineffektivitet i samferdselssektoren. TØI-rapport 1444/2015.

Minken, H. (2006) A note on the effects of the marginal cost of funds and the indirect tax correction factor on optimal user charges in transport, with a caution. TØI Working Paper TØ/1898/2006.

Minken, H. og H. Samstad (2005) Nyttekostnadsanalyser i transportsektoren: rammeverk for beregningene. TØI-rapport 798/2005.

Ramjerdi, F. (1995) *Road Pricing and Toll Financing*. Essay no 3. Ph.D dissertation at the Royal Institute of Technology, Stockholm.

Samstad, H., M. Killi og R. Hagman (2005) Nyttekostnadsanalyser i transportsektoren: parametre, enhetskostnader og indekser. TØI-rapport 797/2005.

Thune-Larsen, H., Veisten, K., Rødseth, K.L., Klæboe, R. (2016) Marginale eksterne kostnader ved veitrafikk – med korrigerte ulykkeskostnader, TØI-rapport 1307/2014 (revidert 2016).

Varian, H. (1992) Microeconomic Analysis. Third Edition, W.W. Norton & Company, New York.

Verhoef, E.T. and K.A. Small (1999) Product differentiation on roads: Second-best congestion pricing with heterogeneity under public and private ownership. Irvine Economics Paper 99-00-01, University of California at Irvine.

Verhoef, E.T. and K.A. Small (1999) Product differentiation on roads. Journal of Transport Economics and Policy 38(1), 127-156.

Welde, M. (2005) Bompengefinansiering – innkrevingskostnadene avhenger av mange forhold. *Økonomisk Forum* 1/2005.

## 2.6 Supplerende prosjektvalgsregel[35]

[35] Dette er arbeidsdokument 50314 fra 2016.

# Valg av prosjektalternativer til en pakke. Supplerende regel for tilfellet der noen av alternativene er sjølfinansierende

I et foreløpig utkast til artikkel om nyttekostnadsberegning av tiltak for å oppnå nullvekstmålet (Minken 2018) sier jeg at hvis bypakka som nullvekstmålet skal gjelde for, har en gjennomsnittlig økonomisk ramme pr år, vil årlig netto nytte pr budsjettkrone (NNB) være det riktige lønnsomhetsmålet. I en kommentar til utkastet reiser Vidar Rugset fra SVV innvendingen at dersom eventuelle restriktive tiltak medfører en så stor innbetaling til det offentlige at det offentlige går i pluss (investeringer medregnet) gir ikke NNB mening.

Jeg takker Rugset for kommentaren. I dette arbeidsdokumentet presiserer jeg først hvordan man skal velge mellom sjølfinansierende prosjekter generelt, og utleder deretter det riktige kriteriet for valg av prosjektalternativ dersom ett eller flere av alternativene er sjølfinansierende, mens andre krever budsjettmidler over offentlige kasser.

*

Jeg ser på et enkeltprosjekt som enten bare foreligger i ett alternativ, eller som foreligger i flere alternativer som er formulert slik at alle alternativene er gjensidig utelukkende – dvs. at i høyden ett av dem kan velges. Slike alternativer kan besluttes separat fra resten av bypakka dersom ingen av dem belaster budsjettet (Minken 2016 side 12). Det vil si:

- Velg hvert eneste uavhengige prosjekt som har positiv netto nytte, dersom det ikke foreligger negative ikke-prissatte konsekvenser, begrensninger for hvor mange prosjekter som kan iverksettes samtidig, eller andre betenkeligheter.
- Velg alternativet med høyest netto nytte, dersom det foreligger flere gjensidig utelukkende alternativer med positiv netto nytte.
- Velg ingen alternativer, dersom alle alternativer er ulønnsomme.

Men hva om noen av de gjensidig utelukkende alternativene krever budsjettmidler og andre ikke?

Anta det finnes to alternativer – et reint avgiftsalternativ som ikke belaster offentlige budsjetter, og nøyaktig det samme avgiftssystemet, men kombinert med en investering, la oss si i sykkelveger eller kollektivtransport. Velger man det ene alternativet, har man pr definisjon valgt bort det andre. Anta videre at investeringen koster mer enn det man tar inn i form av avgifter.

Ved å velge å kombinere avgiftssystemet med en investering, har man ikke bare valgt bort det reine avgiftsalternativet, men i tillegg har man også valgt bort noen andre prosjekter, nemlig en del av de som da blir trengt ut av bypakka fordi investeringen ikke er sjølfinansierende.

Vi må derfor stille krav om at investeringsalternativet ikke bare er en forbedring i forhold til avgiftsalternativet, men *en så stor forbedring* at det i tillegg kompenserer for den tapte netto nytten av de prosjektene som er fortrengt fra pakka.

Jeg viser nå til Minken 2016, «Project selection with sets of mutually exclusive alternatives”, side 14, formel 5. La prosjektet vi ser på være $j$, og la alternativet som ikke belaster ramma være $i=1$, mens det som belaster ramma er $i=2$. Skal målfunksjonen $V$ øke om vi velger alternativ 1, må følgende ulikhet gjelde:

$$b_{1j} > b_{2j} - (1 + \lambda + k)c_{2j}$$

Eller med andre ord: Velg alternativ 2 framfor det reine avgiftsalternativet 1 hvis og bare hvis

(23) $$b_{2j} - (1 + \lambda)c_{2j} > b_{1j} + kc_{2j}$$

I ulikheten (1) er venstresida lik netto nytte av det sammensatte avgifts- og investeringsalternativet 2, mens høyresida er netto nytte av avgiftsalternativet pluss et tillegg, som er verdien av den marginale nytten vi går glipp av ved å bruke en del av investeringsramma til prosjekt 2, når vi kunne ha valgt et alternativ som ikke belaster investeringsramma i det hele tatt. Vi må altså stille høyere krav til et alternativ som belaster investeringsramma enn til et alternativ som ikke gjør det.

I dette tilfellet fantes det bare ett prosjekt, som kunne gjennomføres i to gjensidig utelukkende varianter – med og uten en investering som belaster den gitte investeringsramma. Men regelen vi har utledet (ulikheten (1)), er like brukbar på et hvilket som helst antall prosjekter eller gjensidig utelukkende prosjektkombinasjoner som finnes i en variant som er sjølfinansierende og en variant som krever midler fra et offentlig budsjett eller en bypakke.

## Referanser


Minken, H. (2018) Slik gjør vi: samfunnsøkonomisk analyse av nullvekstmålet. Arbeidsdokument 51287 (UTKAST).

Minken, H. (2016) Project selection with sets of mutually exclusive alternatives, *Economics of Transportation* 6(2016), 11-17.

## 2.7 Optimalt prosjektvalg[36]

[36] Denne teksten gjengir hovedtrekkene i en systematiske metode for å velge prosjekter til en transportplan på, gitt at hvert prosjekt eventuelt kan realiseres på flere ulike måter. Detaljene finnes i Minken, H. (2016) Project selection with sets of mutually exclusive alternatives, *Economics of Transportation* 6 (2016), 11-17.

# Optimalt prosjektvalg med gjensidig utelukkende prosjektalternativer

Vi ser på valget av investeringsprosjekter i en transportplan. La $P$ være mengden av prosjektene som er kandidater til å tas med i planen, og indekser prosjektene med $j \in P$. Kall mengden av gjensidig utelukkende alternativer i prosjekt $j$ for $A_j$, og indekser disse mengdene med $i \in A_j$. Som i hovedteksten[37] definerer vi nyttekostnadsbrøken som

$$(b_{ij} - (1 + \lambda)c_{ij})/c_{ij},$$

der $b_{ij}$ er neddiskontert netto nytte for trafikanter, operatører og tredjepart i tilknytning til den alternative utformingen $i$ av prosjekt $j$, mens $c_{ij}$ er neddiskontert verdi av inn- og utbetalinger over offentlige kasser, og $\lambda$ er den marginale kostnaden ved skattefinansiering.

Vi antar uendelig delelige prosjekter, og formulerer det lineære programmeringsproblemet å finne den sammensetningen av prosjektalternativer som maksimerer samfunnsøkonomisk nytte på følgende måte:

$$\text{(LP3)} \qquad \underset{\mathbf{x}}{\text{maks}} \sum_{j \in P} \sum_{i \in A_j} \left[ b_{ij} - (1+\lambda) c_{ij} \right] x_{ij} \quad \text{gitt} \quad \sum_{j \in P} \sum_{i \in A_j} c_{ij} x_{ij} \le a$$

$$\sum_{i \in A_j} x_{ij} \le 1, \quad j = 1, 2, \ldots, |P|$$

$$x_{ij} \in [0,1] \quad \forall j \in P \text{ og } i \in A_j$$

Her er $a$ budsjettet og $|P|$ er antall prosjekter. Den første bibetingelsen sier altså at vi må holde oss innafor budsjettet, mens den andre betingelsen (eller rettere sagt, de $|P|$ neste betingelsene) sier at andelene som kan realiseres av hvert alternativ, må summere seg til 1. Som vi skal se, vil det i praksis bety at i høyden ett alternativ kan velges.

For å løse problemet formulerer vi et liknende problem, som under visse betingelser gir samme løsning som LP3, men er lettere å løse, i alle fall tilnærmet. Det vi gjør, er å fjerne budsjettbetingelsen som bibetingelse, og ta den inn i målfunksjonen i stedet, multiplisert med en parameter som vi her skal kalle $k$. Den nye målfunksjonen blir:

$$\sum_{j \in P} \sum_{i \in A_j} \left[ b_{ij} - (1+\lambda) c_{ij} \right] x_{ij} - k \left( \sum_{j \in P} \sum_{i \in A_j} c_{ij} x_{ij} - a \right)$$

Litt omforming av målfunksjonen gir:

$$ka + \sum_{j \in P} \sum_{i \in A_j} \left[ \left( b_{ij} - (1+\lambda) c_{ij} \right) - k c_{ij} \right] x_{ij} = ka + \sum_{j \in P} \sum_{i \in A_j} \left( \frac{b_{ij} - (1+\lambda) c_{ij}}{c_{ij}} - k \right) c_{ij} x_{ij}$$

$$= ka + \sum_{j \in P} \sum_{i \in A_j} \left( h_{ij} - k \right) c_{ij} x_{ij}$$

[37] Det er uklart hvilken hovedtekst det her refereres til. Men det kan godt være «Project selection with sets of mutually exclusive alternatives», Economics of Transportation 2016 (6).

Her er $h_{ij}$ nyttekostnadsbrøken, på samme måte som i hovedteksten.

Denne typen omforming av målfunksjonen kalles Lagrange-relaksasjon. Det er ofte en enkel måte å finne en tilnærmet løsning på. Her, hvor vi opererer med kontinuerlige variable, kan vi faktisk finne en eksakt løsning også. Det gjøres ved å minimere målfunksjonen med hensyn på Lagrangeparameteren $k$. Vi ser av siste linje at dette minimum oppstår når summen av kostnadene for de utvalgte prosjektene er akkurat lik budsjettet $a$. Vi må altså finne den $k$ som gjør dette mulig. Men samtidig skal vi også ta hensyn til de gjenværende bibetingelsene, nemlig betingelsene som sier at summen av utvalgte alternativer for hvert av prosjektene skal være høyst 1.

Vi kan nå se hvorfor denne omformingen var lur: For enhver gitt $k$ vil det nemlig være enkelt å finne de alternativene som gir størst målfunksjon. Alle alternativer som skal bidra til å øke målfunksjonen må for det første ha nyttekostnadsbrøk $h_{ij}$ større en $k$. Vi finner lett ut hvilke prosjekter det er. For det andre: For prosjekter som tilfredsstiller det minimumskravet, blir det lett å plukke ut det alternativet som bidrar mest til målfunksjonen, og vi skjønner at det skal inngå i sin helhet, siden det bare gjør målfunksjonen dårligere om vi blander inn andeler av andre alternativer. Bare det aller dårligste av prosjektene som bidrar positivt, vil inngå med en mindre andel.

Hadde vi ikke gått fram på denne måten, ville vi kanskje måtte vurdere milliarder av mulige prosjektkombinasjoner. Det eneste som nå gjenstår, er å eksperimentere med $k$ til vi finner en verdi som gjør at alle utvalgte prosjekter får plass i budsjettet. Som sagt i teksten, kan det gjøres raskt og med god tilnærming ved en søkealgoritme som det gyldne snitt.

Vi har altså vist at framgangsmåten fra hovedteksten gir en samfunnsøkonomisk optimal plan med god tilnærming. Og det gjelder faktisk også om prosjektalternativene ikke er delelige. Det eneste som kan være usikkert, er om vi har iterert mange nok ganger til at den gjenstående avstanden mellom den $k$ vi bruker og den som fyller budsjettet helt nøyaktig, ikke spiller noen rolle, samt hva vi skal gjøre med midlene som blir tilovers etter at siste hele prosjektalternativet har fått plass i planen.

*

## Heltallsprogrammering av totrinns prosjektvalg med begrensede ressurser

Her gir vi et eksempel på det matematiske problemet som må løses hvis det finnes *flere* begrensede ressurser, prosjektalternativene utelukker hverandre gjensidig, og prosjektene og alternativene *ikke* kan antas å være delelige. Det vi mener med at det finns flere begrensede ressurser, er at i tillegg til budsjettet finnes det for eksempel et klimautslippsmål for planen som helhet, et mål om at antall ulykker ikke skal overskride et visst tall, eller en kvantifisert bestemmelse om hvor mye dyrket mark som planen kan legge beslag på. Mengden av slike mål og restriksjoner kaller vi $S$, og indekserer den med $s$. Nytte, kostnad, utslipp, ulykker osv. for hvert alternativ antas å være regnet ut på forhånd. Med for øvrig samme notasjon som i artikkelen ellers har vi dette heltallsproblemet:

$$\text{(IP1)} \qquad \underset{\mathbf{x}}{\text{maks}} \sum_{j\in P}\sum_{i\in A_j} b_{ij}x_{ij} \quad \text{gitt} \quad \begin{aligned} &\sum_{j\in P}\sum_{i\in A_j} c_{ij}x_{ij} \le a \\ &\sum_{j\in P}\sum_{i\in A_j} e_{ijs}x_{ij} \le m_s, \ s = 1,2,\ldots,|S| \\ &\sum_{i\in A_j} x_{ij} \le 1, \quad j = 1,2,\ldots,|P| \\ &x_{ij} \in \{0,1\} \quad \forall j \in P \text{ og } i \in A_j \end{aligned}$$

Den første bibetingelsen er naturligvis budsjettbetingelsen. De |S| neste gjelder klimagass-restriksjonen og andre mål som stilles til planen som helhet. Deretter har vi et sett av betingelser som sier at fra hvert prosjekt må ikke mer enn ett alternativ (inkludert nullalternativet) velges. Siste linje i problemet sier at (i motsetning til LP-problemene) må valgvariablene være enten 0 eller 1, ikke noen mellomting.

Framgangsmåten for å løse problemet i vedlegg 1 vil være mye vanskeligere å anvende på problemet (IP1). Problemet har en såkalt blokkangulær struktur og likner på lineære problemer i litteraturen om dekomponerbare systemer og desentralisert planlegging (Williams 1990). Nøyaktig hva slags framgangsmåte som er best egnet til å løse problemet, har vi ikke sett på.

# 3 Teoretisk grunnlag

## 3.1 Byvekst og transport[38]

[38] Med noen få språklige endringer er denne artikkelen tidligere offentliggjort som kapittel 6 i Bråthen m.fl. (2003).

# Byvekst og transport

Hva er de indre kreftene som styrer og påvirker de store byenes vekst? Og hvilken rolle spiller transporten i byene i denne sammenhengen? Det er det vi skal prøve å belyse i dette kapitlet.

Vi bygger framstillingen på læreboka Fujita (1989) og en del artikler. En av dem er av John Quigley og handler om mangfoldet i byer og hva den har å si for økonomisk vekst (Quigley 1998). En annen er en økonometrisk undersøkelse av Vernon Henderson som påviser at det finns et punkt der konsentrasjonen av befolkningen i den største eller de aller største byene har gått for langt og hemmer landets økonomiske vekst (Henderson 2000).

## Urban economics

Grovt sett kan man skille mellom to teoritradisjoner innen økonomisk teori om byutvikling. Den første, urban economics, blomstret i 60- og 70-årene (sjøl om det framleis kan gjøres verdifullt arbeid med dette utgangspunktet). Alonso (1964) regnes som det grunnleggende verket i denne tradisjonen. En moderne lærebok er Fujita (1989). Denne teorien tar eksistensen av byen og den økonomiske virksomheten i den for gitt. Spesielt tar den for gitt at det finnes et sentrum i byen der den viktigste økonomiske virksomheten foregår. I den enkleste versjonen foregår all produksjon i sentrum, og innbyggerne må reise dit for å arbeide. Transporten koster mer jo lenger ut man bor. Husholdningene etterspør tomteareal til boligformål og et sammensatt forbruksgode. Det teorien tar sikte på å forklare, er lokaliseringsmønster og botetthet innad i byen, byens geografiske utstrekning, eventuelt utnyttingsgraden på tomtene (høyden på husene), osv. Alt dette varierer med transportkostnader, inntektsnivå, prisen på jord ved bygrensa, og sjølsagt med innbyggertallet.

Det antas at byen ligger på en helt flat slette og er utbygd likt i alle retninger. Den er altså sirkelformet. Denne forenklingen gjør de matematiske beregningene enklere. Men sjøl i den enkleste versjonen er denne teorien i stand til å forklare mange empirisk observerbare trekk ved byens indre struktur og ved eiendomsmarkedet og boligmarkedet, kanskje spesielt i amerikanske byer.[39]

Teorien kan bygges ut i mange varianter: Vi kan ha åpne eller lukkede byer (dvs. at innflytting kan være mulig eller umulig), jordrenta kan tilegnes av tomteeiere som bor i eller utenfor byen, alle innbyggere kan være like eller oppdelt i ulike klasser, tjenesteyting kan foregå lokalt, produksjon kan foregå utenfor sentrum, deler av byen kan ha bedre miljø eller bedre tilbud av offentlige goder, ulike folkegrupper kan foretrekke å bo fra hverandre, og vi kan ha kø og trengsel. Det sistnevnte er naturligvis eksempler på negative eksterne effekter i boligmarkedet eller transportmarkedet. I fravær av slike effekter vil markedsløsningene være kostnadseffektive, og en optimal (kostnadseffektiv) bystruktur med ønsket fordelingsvirkning kan alltid oppnås gjennom en markedsløsning med lump-sum skatter. De fundamentale velferdsteoremene gjelder altså i en slik by. Løser vi på forutsetningen om et bysentrum der den økonomiske virksomheten foregår, stiller dette seg annerledes.

[39] I hvert fall hvis ikke byens form og utstrekning bevisst blir bevart, slik som i noen europeiske byer.

## Hva hvis vi ikke forutsetter et bysentrum?

Hvorfor skulle det eksistere et sentrum der den økonomiske aktiviteten samler seg? En opplagt grunn kunne være at der fantes det et råstoff for produksjonen som ikke så lett lot seg transportere andre steder (en gruveby, en fiskerihavn). Andre muligheter er at historiske tilfeldigheter (fra et økonomisk synspunkt) har skapt et krystalliseringspunkt for økonomisk aktivitet akkurat her (en garnisonsby, en fyrsteresidens), eller at stedet er en utskipingshavn eller et transportknutepunkt generelt. For øvrig må svaret være at det eksisterer en eller annen form for stordriftsfordeler eller samlokaliseringsfordeler i produksjonen. Hvis all produksjon, uansett hvor liten, foregikk med konstant skalautbytte, ville byer ikke ha noen hensikt i det hele tatt. Ethvert behov kunne tilfredstilles med lokal produksjon, og dermed sparte man transportkostnader.[40]

La oss se på hvordan transportkostnadene innad i en by kan frambringe et bysentrum, eller i hvert fall en ”indre by”. Borukhov og Hochman (1977) tar vekk den sentrale forutsetningen i urban economics-litteraturen om at det finnes et sentrum dit alle reiser går. I stedet antar de at enhver innbygger (husholdning eller bedrift) har behov for å reise til alle andre innbyggere i byen. De skal velge lokalisering på bakgrunn av dette. Dette kan virke like urealistisk, men det kan tolkes som at de må velge lokalisering før de veit hvor de får arbeid (eller, hvis de er bedrifter, før de veit hvor kundene vil være). Denne besøks- eller kontaktvirksomheten pluss bygging av boliger (næringsbygg) er den eneste økonomiske aktiviteten i byen. I utgangspunktet vil folk lokalisere seg mest mulig sammen for å spare transportkostnader. Men det krever mer kapital pr. kvadratmeter tomteareal og dermed høyere husleie. Individene vil minimere summen av husleie og transportkostnader for den gitte besøksaktiviteten. Borukhov og Hochman viser at det da i markedslikevekt vil dannes et sentrum med høy tetthet, og avtakende tetthet ut til bygrensa.

Denne markedsløsningen er imidlertid ikke samfunnsøkonomisk optimal. En planlagt løsning, der planleggeren minimerer kostnadene i byen, vil også gi et bysentrum. Men i forhold til markedsløsningen vil den optimalt planlagte byen ha mindre utstrekning, og spesielt ha høyere arealutnyttelse nær bygrensa. Det vi finner i denne modellen, er at det eksisterer en form for positiv ekstern virkning av samlokalisering (en agglomerasjonsfordel). Når hver enkelt bestemmer sin lokalisering, tar han nemlig hensyn til sine egne reisekostnader, men ikke til at alle de andre som skal besøke ham, kunne spare reisekostnader dersom han hadde bosatt seg litt mer sentralt. De eksterne virkningene kan internaliseres gjennom en eiendomsskatt i ytre by og boligsubsidiering i indre by.

Vi er da halvveis over i den andre teoritradisjonen, som prøver å besvare spørsmålet om hvorfor byer oppstår og vokser, hva som setter grenser for veksten, og hvordan byvekst fremmer eller hemmer økonomisk vekst. Sentralt i denne tradisjonen står begrepet eksterne stordriftsfordeler, som er kostnadsbesparelser eller økt nytte for bedrifter og konsumenter ved at det finnes mange andre virksomheter og folk i nærheten.

[40] Dette poenget, som framhever stordriftsfordeler og transportkostnader som de fundamentale årsakene til at det oppstår et mønster av byer med omliggende distrikter, gjøres mange steder, bl.a. Krugman (1998), Quigley (1999).

## Gir mangfoldet og valgmulighetene i byen økonomisk vekst?

Den andre teoritradisjonen har dominert den økonomiske tenkingen om byer i de siste 15-20 årene. Delvis kan vi si at dette er «den nye økonomiske geografien» i den grad den er anvendt på byer. Den er altså opptatt av dynamiske prosesser som økonomisk vekst og den geografiske fordelingen av denne veksten. Men dette er bare delvis riktig, fordi den nye økonomiske geografien i streng forstand framleis opererer med et modellapparat som blir litt for spesielt og litt for vanskelig å teste empirisk og trekke praktiske konklusjoner av. Mye arbeid foregår derfor i en statisk ramme.[41] Den nye teoritradisjonen er imidlertid opptatt av alle former for agglomerasjonsfordeler, og spesielt eksterne stordriftsfordeler. Eksterne stordriftsfordeler, og innsikten om at forbrukerne har nytte av bredden og variasjonen i utvalget av goder, og produsentene har nytte av at det finns mange typer av produksjonsutstyr og andre innsatsfaktorer, er felles trekk som det meste av moderne økonomisk forskning på byer deler med den nye økonomiske geografien.

Vi kan definere agglomerasjonsfordeler som reduksjon i gjennomsnittlige kostnader når produksjonen samles i et gitt geografisk område. Aggglomerasjonsfordeler omfatter dermed de naturgitte eller historisk oppståtte gunstige betingelsene som vi har vært inne på, tilgangen på et offentlig gode som ikke finnes andre steder, stordriftsfordeler i den enkelte bedrift og eksterne stordriftsfordeler. Eksterne stordriftsfordeler kan deles i lokaliseringsfordeler og urbaniseringsfordeler. Lokaliseringsfordeler er at det gir kostnadsbesparelser om det finnes mange bedrifter innen samme bransje i samme distrikt, mens urbaniseringsfordeler er at det gir kostnadsbesparelser om det finns mye annen virksomhet i distriktet, uansett bransje. Tilsvarende fordeler vil også kunne finnes i forbruket.

Quigley (1998) behandler fire slags agglomerasjonsfordeler.

1. Stordriftsfordeler i produksjonen (større fabrikkanlegg) og forbruket (offentlige goder som parker, idrettsanlegg m.m.).
2. Felles innsatsfaktorer. I produksjonen innebærer det for eksempel at arbeidskraft innen vedlikehold, regnskapsføring, juridiske tjenester m.m. kan utnyttes bedre og dermed bli billigere eller kunne bli mer spesialisert. I forbruket kan kulturtilbudet, tilbudet av restauranter m.m. bli mer variert eller billigere fordi det finns et større kundegrunnlag.
3. Lavere transaksjonskostnader. Søkekostnadene på arbeidsmarkedet reduseres, og spesialiserte handledistrikter gir lavere søkekostnader for de handlende.
4. Statistiske stordriftsfordeler. Eksistensen av mange bransjer gjør det mulig å få arbeid i en bransje som går godt dersom en mister arbeidet i en bransje som går dårlig. Bruktmarkeder for alle slags maskiner og utstyr er mer effektivt, og i forbruket er det mulig å finne gode substitutter for goder en ikke får tak i.

[41] Krugman (1996) deler det vi her kaller den nye teoritradisjonen i to: Nyklassisk bysystemteori og teori som bygger på monopolistisk konkurranse. Den nye økonomiske geografien er den sistnevnte. Som et fortrinn for denne teorien framhever han at den er eksplisitt om hvilke mekanismer som fører til de eksterne stordriftsfordelene. En ulempe er at denne konkretiseringen kan ta fatt i de mindre viktige virkningene og overse de viktige.

De statistiske stordriftsfordelene er det grunn til å utdype. Delvis skyldes de at variasjoner i etterspørselen blir mindre i følge de store talls lov. Men det er også viktig at det trenges et relativt mindre sikkerhetslager av innsatsfaktorer når produksjonen eller salgsvolumet er høyt. For eksempel vil et busselskap som driver med tre busser bli nødt til å ha en buss i reserve. Et busselskap med hundre busser trenger kanskje bare ti busser i reserve.

Av alle disse grunnene kan en større by gi grunnlag for flere varianter av varer og tjenester og/eller billigere varer og tjenester enn en mindre. Dette øker produktiviteten i store byer og nytten til de som bor der. I tillegg til dette kommer den raskere kunnskapsoverføringen og oppbyggingen av kunnskapskapital som følger med de hyppigere kontaktene i byene, og genereringen av ny kunnskap og nye ideer som det medfører. Ifølge endogen vekstteori er dette prosesser som kan gi grunnlag for en vedvarende økonomisk vekst.

Empiriske undersøkelser gir belegg for at disse prosessene finns, sjøl om det er vanskelig å identifisere dem nøyaktig. For eksempel er det vist at patenter hyppigere blir anvendt i den geografiske nærheten av det stedet de ble tatt ut. Dette tilsier at fysisk nærhet og kontakt er viktig for kunnskapsspredning.

Når det gjelder kunnskapsspredning, ser det ellers ut til at hard konkurranse mellom mange bedrifter gir en raskere spredning av kunnskapen og større eksterne stordriftsfordeler enn monopol. Derimot er det usikkert om det er lokaliseringsvirkninger eller urbaniseringsvirkninger som betyr mest. Glaeser et al (1992) finner at urbaniseringsvirkningen er viktigst, og tilbakeviser dermed bl.a. Porters teori om betydningen av klustre. Glaeser (1999), derimot, ser ut til å komme til motsatt resultat.

Men mens byvekst gir eksterne stordriftsfordeler, skaper det også ”sentrifugale” krefter, som kø, trengsel og andre ulemper, som kan tilta til et punkt hvor de koster så mye at de bremser og motvirker ikke bare veksten av den enkelte byen, men også den økonomiske veksten som byene understøtter.

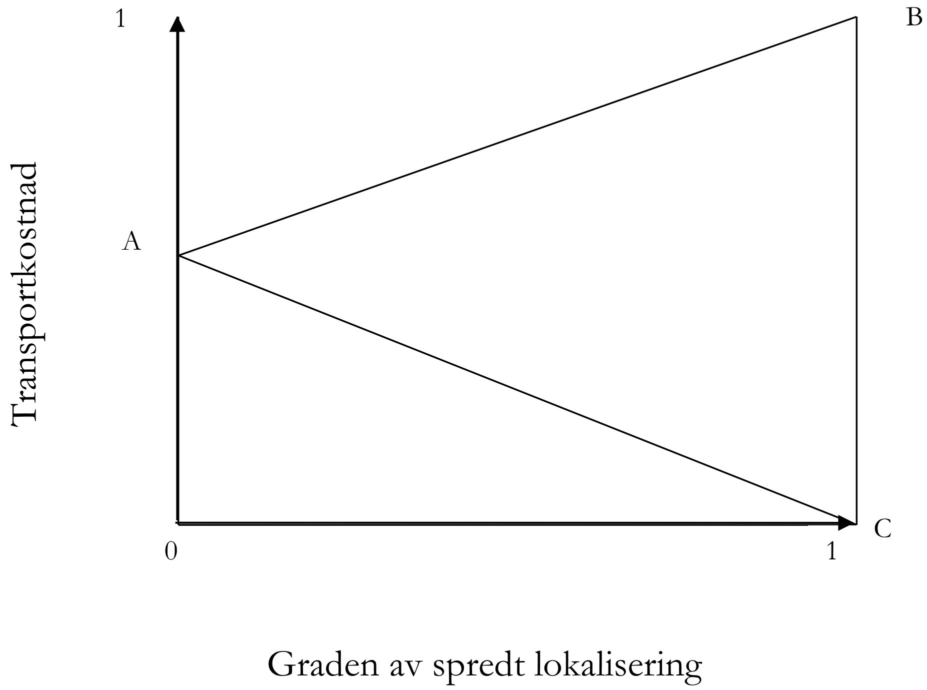
## Entropi er nyttig

Med entropi mener vi tendensen til å variere reisemål. Høy entropi betyr at reisene fra ethvert sted i byen fordeler seg jamt utover hele byen. Dette står i motsetning til å la reisekostnadene bestemme bestemmelsesstedet, noe som naturligvis fører til at man velger det nærmeste stedet der reisehensikten kan bli oppfylt. I modellen til Borukhov og Hochman var entropien maksimal. Siden innbyggerne allerede i utgangspunktet hadde bestemt seg for å besøke alle, måtte de bare ta med på kjøpet de høye transportkostnadene som det medførte. I stedet prøvde de å avveie transportkostnadene mot husleia ved å lokalisere seg kostnadseffektivt. Tar vi derimot lokaliseringen for gitt, må vi regne med at folk avveier behovet for variasjon mot behovet for å redusere transportkostnadene. De vil prøve å maksimere entropien for et gitt reisebudsjett, eller minimere reisekostnadene for en viss entropi, eller de prøver å finne fram til den beste kombinasjonen av de to hensynene samtidig.[42] Hvis vi altså tar folks preferanser alvorlig, vil det ikke være noe mål i seg sjøl å minimere reisekostnadene i byen.

Tar vi i tillegg hensyn til at kontakt mellom folk kan skape positive eksterne virkninger, eller at variasjonsmulighetene i byen er et gode både i produksjon og forbruk, så skal vi gjerne få i

[42] En annen tolkning av folks reiseadferd som gir mye av den samme typen transportmodeller, er at den enkelte veit hvor han vil og holder seg til det dag ut og dag inn, men at folk foretrekker forskjellige reisemål, uten at vi helt klarer å observere forskjellen nøyaktig.

stand litt mer entropi enn det folk velger på egen hånd. Dette betyr ikke nødvendigvis at folk reiser til hverandre enkeltvis. De kan jo alternativt alle reise til samme sted og møtes der. Vi kan illustrere de ulike mulighetene med ”Brotchie-trianglet” (Brotchie 1984), som er gjengitt i forenklet form i figur 1.




**Figur 1** Brotchietrianglet

Langs den vertikale aksen måler vi transportkostnad pr. reise. Langs den horisontale aksen måler vi graden av spredt lokalisering av reisemålene. Lokaliseringen av de som skal reise, må antas gitt.

Anta vi ser på innkjøpsreiser. Punkt A er da åpenbart tilfellet hvor alle butikkene er i sentrum. Det er urealistisk å regne med at alle boligene også er det, så punktet 0 er umulig i praksis. Transportkostnaden eller reiselengden i gjennomsnitt er i stedet A. Når boliglokaliseringen er gitt, er A en gitt størrelse. Hele linja BC representerer tilfellet der butikkene ligger spredt ut i boligområdene. Her har vi imidlertid ”fritt valg” langs hele linja BC. Punktet C representerer ekstremtilfellet der alle handler i sin lokale butikk. Punktet B representerer tilfellet der alle drar overalt for å handle, omtrent som hos Borukhov og Hochman. For enhver reisehensikt (eller for bedriftenes transporter) kan en gitt by karakteriseres ved et punkt i trianglet ABC i figuren. Linja AC representerer en transportkostnadsminimerende by med butikker som må ha mer og mer landhandelpreg når vi går fra A til C. Linja AB representerer tilfellet hvor hver har sine spesialiserte butikker som de drar til uansett kostnad, eller hvor alle hele tida prøver alt.

Jo lenger vi går oppover i diagrammet, jo større entropi, og dermed større samlet reisekostnad. Hvis vi venter oss at det er kundene som må ta initiativet hvis vi skal oppnå de positive eksterne virkningene av variasjon og økte kontakter, så må vi på en eller annen måte få beveget byen litt oppover i trianglet. Hvis vi derimot venter oss at dette kan skje ved at butikkene utnytter stordrifts- og samlokaliseringsfordeler, så skal vi dytte på litt til venstre. Hvilke virkemidler vi har, kan imidlertid diskuteres.

Legg merke til at byer mot venstre vil kunne utnytte stordriftsfordeler i transport i form av kollektivtransport, store godsbiler og konsentrasjon av veginvesteringene til store innfartsårer, mens byene til høyre vil være privatbilbyer, i hvert fall litt opp i diagrammet. Transporttiltak i byen vil kunne lette en bevegelse oppover i diagrammet, og dermed eventuelt være et nestbeste-tiltak for å internalisere en bestemt form for agglomerasjonseksternaliteter. Transporttiltakene kan imidlertid enten føre byen til høyre eller venstre.

## Byvekst og økonomisk vekst

Vi må skille mellom urbaniseringsgraden og graden av konsentrasjon av bybefolkningen til noen få store eller en enkelt stor by. Det er det sistnevnte som angår oss her. Siden det finnes krefter som tilsier at byene bør bli større og krefter som tilsier det motsatte, bør det kunne finnes en optimal grad av konsentrasjon av bybefolkningen til de største byene. Mange har også antatt at de største byene i flere tredje verden-land har blitt for store, og hemmer utviklingen i landet. Krugman (1996) og Henderson (2000) er blant de som drøfter dette.

Henderson studerer spørsmålet i en økonometrisk modell med data fra 80-100 land fra 1960 til 1995. Vi har ikke funnet ut om Norge er med i dette materialet. Han finner at det finns en optimal konsentrasjonsgrad. Den øker med BNP pr. innbygger opp til et visst punkt og avtar deretter langsomt. Den optimale konsentrasjonsgraden er mindre i små land. Siden Norge ikke er nevnt i undersøkelsen, kan det antas at konsentrasjonsgraden er omtrent optimal dersom Norge er med. (Landene med for liten eller for stor konsentrasjon er nemlig listet opp).

Det er rimelig, hevder Henderson, at land på et lavt inntektsnivå bør tilstrebe høy konsentrasjonsgrad, siden det sparer infrastruktur og utnytter muligheten til kunnskapsoverføringer i en situasjon hvor kunnskapen er konsentrert hos noen få. Mer utviklede land har større muligheter til å bygge ut infrastrukturen i hele landet og spre veksten. Tapene ved for stor konsentrasjon, uttrykt som tapt årlig BNP-vekst, er store, og større jo mer utviklet landet er. De kan gå opp i 1,5%.

Blant de faktorene som kan motvirke for høy konsentrasjon er det bare infrastrukturbygging i distriktene som slår kraftig ut i Hendersons undersøkelse. Dette virkemidlet virker kraftigere jo høyere nasjonalinntekten er. For et land med høy inntekt, men for stor konsentrasjon til noen få byer, vil derfor infrastruktur i distriktene kunne øke den økonomiske veksten betydelig.

## Konklusjon

Det er grunnlag for å hevde at de grunnleggende drivkreftene bak byutviklingen er stordriftsfordeler og samlokaliseringsfordeler i produksjonen, sammen med behov for variasjon og spesialisering i forbruket. Byenes størrelse og vekst – og dermed til en viss grad «sentraliseringssuget» i form av flytting fra distriktene til byen – avgjøres i samspillet mellom slike agglomerasjonsfordeler på den ene sida, og trengselskostnader i form av kø i transporten, høye husleier og andre ubehagelige sider ved tette ansamlinger av folk på den andre sida. Transporttiltak i byene kan derfor gi grunnlag for større byer. Hvorvidt dette fører til økt økonomisk vekst kommer an på om det framleis er agglomerasjonsfordeler å hente ut. (Om den eventuelt økte veksten kommer distriktene til gode eller skjer på bekostning av distriktene tar vi ikke stilling til her.)

Vi har antydet at så lenge det finns betydelige agglomerasjonsfordeler, dvs. eksterne stordrifts- og samlokaliseringsfordeler, vil byen utvikle seg mer spredt og med færre innbyggere enn det som er samfunnsøkonomisk optimalt. Samtidig vil også uprisede negative eksterne virkninger i transportsystemet og arealbruken medføre at byene blir større og kanskje mer spredt enn det som er samfunnsøkonomisk forsvarlig. Riktige priser vil være viktig, ikke bare for transportmarkedet og boligmarkedet isolert sett, men for hele byutviklingen. I transportmarkedet dreier det seg om vegprising, i boligmarkedet er kanskje litt vanskeligere. På frunnlag av riktige priser kan en deretter vurdere om det finnes samfunnsøkonomisk lønnsomme transporttiltak og byutviklingstiltak.

For å målrette transporttiltakene i byene bedre med sikte på å ta ut agglomerasjonsfordeler og oppnå økonomisk vekst, burde man vite mer om hva slags virkninger som konkret gjør seg gjeldende. Er det virkninger av kontakter og kunnskapsoverføring eller andre virkninger? Kommer virkningene i stand ved at bedrifter lokaliserer seg på samme sted i byen eller ved at kunder og leverandører lett kan ta kontakt med den enkelte bedrift uansett hvor i byen den ligger?

Dersom konsentrasjonen av bybefolkningen til en eller flere store byer er gått for langt, vil transporttiltak i distriktene være et effektivt tiltak for å motvirke og rette på det.

## Litteraturliste


Alonso W (1964). *Location and Land Use. Toward a General Theory of Land Rent.* Harvard University Press, Cambridge, Mass.

Borukhov E and O Hochman (1977). Optimum and market equilibrium of a city without a predetermined center. *Environment and Planning A*, **9**: 849-856.

Brotchie J F (1984). Technological change and urban form. *Environment and Planning A*, **16**: 583-596.

Bråthen, S., K.S. Eriksen, H. Minken, F. Ohr og I. Thorsen (2003) Virkninger av tiltak innen transportsektoren. En kunnskapsoversikt. Rapport til Effektutvalget. http://www.regjeringen.no/upload/kilde/krd/rap/2003/0006/ddd/pdfv/184613-kunnskapsoversikt.pdf

Fujita M (1989). *Urban Economic Theory*. *Land Use and City Size*. Cambridge University Press, Cambridge.

Glaeser E L (1999). Learning in Cities. *Journal of Urban Economics* **46**: 254-277.

Glaeser E L, H D Kallal, J A Scheinkman and A Schleifer (1992). Growth in Cities. *Journal of Political Economy* **100** 6: 1126-1152.

Henderson V (2000). *How Urban Concentration Affects Economic Growth.* World Bank Policy Research Working Paper 2326, World Bank, Washington DC. Finnes på nettet på www.worldbank.org/research/workingpapers.

Krugman P (1996). Urban Concentration: The role of Increasing Returns and Transport Costs. *International Regional Science Review* **19** (1 &2), 5-30.

Krugman P (1998). Space: the final frontier. *Journal of Economic Perspectives* 12, 161-174.

Quigley J M (1998). Urban Diversity and Economic Growth. *Journal of Economic Perspectives* **12** 2: 127-138.

## 3.2 Tidsverdienes inntektsavhengighet og velferdfunksjonens form[43]

[43] Dette er et arbeidsdokument fra 2011, ment som et innspill til ekspertutvalget som da skulle se på retningslinjene for samfunnsøkonomiske analyser. Det er tidligere inntatt i TØI-rapport 1198/2012.

# Tidsverdienes inntektsavhengighet og velferdsfunksjonens form

## 1 Innledning

Det har vært framholdt at man her i Norge bør gjøre som i andre land når det gjelder de samfunnsøkonomiske analysene av infrastrukturinvesteringer, og øke tidsverdiene år for år i takt med forventet inntektsutvikling. Dette vil øke den beregnede lønnsomheten av prosjektene, spesielt i samferdselssektoren.[44]

Spørsmålet er ett av flere som skal utredes av en ekspertgruppe som er satt ned av Finansdepartementet for å forbedre våre nyttekostnadsanalyser. Det er imidlertid ikke så enkelt som det ofte framstilles, at vi i Norge har blitt hengende etter, og nå må se å innføre de metoder som man for lengst har blitt enige om i utlandet. Tvert imot har det nylig oppstått faglig usikkerhet om hva som er riktig på dette området, parallelt med at vi har fått nye og bedre undersøkelser som kan kaste lys over spørsmålet.

Spørsmålet deler seg i to: Er det noen sammenheng mellom inntektsutviklingen i samfunnet og verdien av spart tid i ulike anvendelser? Og hvis det er slik, hvilke konsekvenser bør det ha for måten vi gjør samfunnsøkonomiske analyser på?

## 2 Tidsverdiforskningen

Vi ser først på hva forskningen sier om sammenhengen mellom inntektsutviklingen og tidsverdiene.

I stadig større grad er det hypotetiske spørreundersøkelser (SP-undersøkelser) som brukes til å anslå tidsverdier. Litt enkelt sagt stilles respondentene overfor valg hvor de må avveie spart reisetid mot økte kostnader eller omvendt, og svarene brukes til å estimere en diskret valg-modell. Tidsverdien er forholdet mellom den estimerte parameteren til reisetida og den esti-merte parameteren til reisekostnaden, eller grensenytten av tid delt på grensenytten av penger, om man vil.[45] Det er gode grunner til å anta at grensenytten av penger er en avtakende funksjon av inntekt, og det vil isolert sett tilsi at tidsverdien øker med inntekta. Men det er større usikkerhet om hvordan grensenytten av tid vil utvikle seg.

Tre typer av undersøkelser, type A, B og C, vil kunne kaste lys over hvordan tidsverdiene påvirkes av inntektsutviklingen. *Type A* studerer inntektsavhengigheten til tidsverdiene i enkelt-stående tidsverdiundersøkelser, dvs. i tverrsnittsdata. Det gjøres typisk ved å spørre etter inntekt samtidig som man samler inn verdsettingsdata, og enten beregne tidsverdien separat for hver

[44] Det vil derimot i liten grad endre den relative lønnsomheten til prosjektene. Et prosjekt som er bedre enn et annet vil normalt også være det etter en slik justering av regnereglene. Det vil sannsynligvis heller ikke kreves økte budsjetter for å få plass i planen til alle de prosjektene som da blir lønnsomme. Det kan ordnes ved å kaste ut prosjekter som er ulønnsomme også etter justeringen.

[45] Det forutsettes at en liten økning i reisekostnaden har samme nyttevirkning som en liten reduksjon i inntekt. Forutsetningen kan svikte på grunn av at folk ikke er helt rasjonelle eller ikke har god nok informasjon om kostnadene, eller at det finns budsjettskranker i tillegg til budsjettskranken på høyeste nivå.

inntektsgruppe og gjøre en enkel regresjonsanalyse på resultatene, eller spesifisere tids- og kostnadsparametrene direkte som funksjoner av inntekt i den diskrete valgmodellen som skal estimeres.

Den aller første nasjonale tidsverdiundersøkelsen, den britiske fra midt i 80-åra, fant at tidsverdiene for både private reiser og tjenestereiser hadde en elastisitet med hensyn på inntekt på omtrent 1. Dette tok man som en bekreftelse på gjeldende britisk praksis fra gammelt av, nemlig å oppjustere tidsverdiene år for år med forventet inntektsvekst i de samfunnsøkonomiske regnestykkene. Seinere undersøkelser, både i Storbritannia og andre land, har imidlertid for det meste funnet elastisiteter rundt 0,5.

Den første norske tidsverdiundersøkelsen (Ramjerdi m.fl. 1997) beregnet ikke tidsverdiens inntektselastisitet, men Figur 10.2 i rapporten viser tydelig at den må være mindre enn 1. Dessuten er den mindre for husholdningsinntekt enn for personlig inntekt. Den siste norske studien (Ramjerdi m.fl. 2010) finner elastisiteter for de ulike transportmåtene på mellom 0,25 og 0,6 og karakteriserer det som lavt. Det er imidlertid ikke veldig lavt for undersøkelser av type A. Et unntak er den danske tidsverdiundersøkelsen, som finner en elastisitet på 0,9 og ikke med sikkerhet kan si at den er forskjellig fra 1 (Fosgerau 2005).

Det gjør en merkbar forskjell om man regner med personlig inntekt eller husholdningsinntekt. Det viser seg ikke bare i Ramjerdi m.fl. (1997), men helt generelt. Funnene på rundt 1 i den første britiske undersøkelsen gjaldt personlig inntekt. Elastisitetene med hensyn på husholdningsinntekt var lavere og mindre tydelige. Og dette mønsteret har vært konsistent siden da (Wardman 2001).

Metodemessig har de empiriske undersøkelsene utviklet seg fra binomisk logit til mixed logit eller semiparametriske metoder, uten at det har gitt fundamentalt nye konklusjoner om tidsverdiens inntektselastisitet.

En vektig innvending mot undersøkelser av type A er at de besvarer et annet spørsmål enn det man er interessert i, dersom det er oppjustering av tidsverdien med tida som er formålet. Ulikheter mellom inntektsgrupper på et bestemt tidspunkt er jo slett ikke det samme som hvordan gjennomsnittet over inntektsgruppene beveger seg over tid.

Undersøkelser av *type B* sammenlikner tidsverdiene fra mer eller mindre like tidsverdiundersøkelser som er gjennomført i samme land eller område med noen års mellomrom. Det dreier seg altså om tidsserieanalyser. Man har bl.a. sammenliknet de nasjonale tidsverdiundersøkelsene i Nederland i 1988 og 1997, i Storbritannia i 1985 og 1999 og i Sverige i 1994 og 2007. Med unntak av den svenske undersøkelsen, som vi skal komme tilbake til, finner man da elastisiteter på null eller mindre.[46]

Den tredje typen av undersøkelser, *type C*, er metaundersøkelser. Disse kan benytte data fra langt flere kilder enn de nasjonale tidsverdiundersøkelsene, og fra mange år, ikke bare to eller ett. Også metaundersøkelsene kan gjøres på mange måter, med data som varierer fra tidsverdier som er segmentert langs alle slags akser til verdier som ikke er segmentert i det hele tatt. I det sistnevnte tilfellet er det vanskelig å skille ut virkningen av inntekt fra andre virkninger og ulikheter mellom primærundersøkelsene. Mark Wardman har gjennomført en serie metaundersøkelser på britiske data (Wardman 2001, Wardman 2004, Abrantes og Wardman 2011). En av

[46] Det er ikke gjort noen seriøs vurdering av hvordan inntektsutviklingen har påvirket den norske tidsverdien fra 1996 til 2009. En første antydning kan man vel få ved å se på figur 5.3 i Ramjerdi m.fl. (2010). Bortsett fra korte bilreiser ser det ikke ut til at tidsverdien har økt vesentlig, målt i faste kroner. Realinntektsøkningen i sammen tidsrom er vel omtrent 30-35 prosent.

forskjellene mellom dem er tidsrommet som data hentes fra. Elastisiteten har økt fra undersøkelse til undersøkelse, fra 0,5 til 0,75 til 0,9 i den siste. I metaundersøkelser kan inntektsbegrepet være BNP pr. innbygger, disponibel personlig inntekt eller disponibel husholdningsinntekt. Det kan diskuteres hva som er det rette, men det er i alle fall slik at de to førstnevnte gir høyere elastisitet enn den sistnevnte.

De varierende resultatene fra empiriske undersøkelser ga et behov for å vurdere situasjonen og komme med anbefalinger om hvilke verdier som skal brukes. For Storbritannia er det gjort i Wardman (2001), Mackie m.fl. (2001) og Mackie m.fl. (2003). I tråd med anbefalingen i den sistnevnte artikkelen har britisk praksis blitt endret fra å oppjustere tidsverdiene med veksten i BNP/capita til å bruke 0,8*BNP/capita for private reiser, mens tjenestereiser framleis oppjusteres som før.

Som et ledd i den nye svenske tidsverdiundersøkelsen (WSP Analyse & Strategi 2010) blei den gamle undersøkelsen fra 1994 gjentatt på akkurat samme måte i 2007, nettopp for å finne ut mer om tidsverdienes utvikling over tid. Man finner at for lavinntektsgruppene er det omtrent ingen utvikling i tidsverdiene, målt i faste kroner, på disse 13 årene, mens derimot høyinntektsgruppenes tidsverdier har en elastisitet med hensyn på inntekt på nær 1. Følgelig vil tidsutviklingen til den gjennomsnittlige tidsverdien være avhengig av utviklingen av inntektsfordelingen.

Börjesson m.fl. (2009) og Börjesson (2010a og b) har analysert de svenske funnene nærmere og fremmer et prinsipielt synspunkt om at tidsverdiene *ikke* skal justeres med økende inntektsnivå. Grunnlaget for synspunktet er påpekningen i Mackie et al (2003) om at de individuelle subjektive tidsverdiene som vi finner i hypotetiske spørreundersøkelser (SP-undersøkelser) ikke kan adderes til samfunnets verdsetting uten en oppfatning om hvilken vekt som skal tillegges hvert individ. En ikke urimelig måte å gjøre det på er å sørge for at fra samfunnets side skal en krone ekstra tillegges samme vekt enten den tilfaller Per eller Pål. Folks marginale nytte av inntekt er forskjellig, men ved å vekte hvert individ med den inverse av den marginale nytten av inntekt, vil de alle telle likt i velferdsfunksjonen. Gjennomføres dette synspunktet på tvers av generasjonene, innebærer det at tidsverdien skal holdes konstant.

Den enkle antakelsen om at vi i Norge må oppjustere tidsverdiene med BNP per innbygger, slik det gjøres i Sverige, Danmark og (inntil for 5-6 år siden) Storbritannia, viser seg altså å være omstridt. Empirien er uklar når det gjelder hva slags elastisitet man skal bruke, og de færreste undersøkelsene tilsier elastisitet lik 1.

Nylig har COWI kommet med anbefalinger om tidsjustering av tidsverdiene og andre parametre som bygger på betalingsvillighet, som ulykkeskostnader, støykostnader og utslippskostnader (COWI 2010). Anbefalingene bygger på britiske anbefalinger, mens de siste svenske resultatene og anbefalingene i mindre grad er tillagt vekt. Det er fullt ut mulig å leve med COWIs anbefalinger en stund, men etter vårt syn er problemet er såpass komplisert at det bør følges opp i et særskilt forskningsprosjekt.

Vi skal nå formulere en modell som vil bekrefte at tidsverdien er en litt mer komplisert størrelse enn det den enkle antakelsen skulle tyde på. Samtidig formulerer vi modellen slik at det ikke er nødvendig å vekte individene med deres marginale nytte av inntekt. Dette gjør vi fordi det er slik de store etterspørselsmodellene i transport fungerer. Et vanskelig spørsmål blir da om tidsverdien, som er en parameter innafor en slik modellkontekst, skal tillates å utvikle seg med tida på en måte som er konsistent med våre transportmodeller, eller om vi skal foreta nytteberegning på en måte som står i strid med den underliggende etterspørselsmodellen vi bruker i beregningene – men som kanskje samsvarer bedre med virkeligheten på langt sikt.

# 3 En modell av en modell

Modellen vi formulerer har noen trekk felles med et vanlig transportmodellsystem, i første rekke at etterspørselsfunksjonene er funksjoner av generaliserte kostnader og ikke er funksjoner av inntekt. Dette er altså en liten modell av en større modell, nemlig av en transportmodell av den typen som er vanlig i samferdselssektoren. Hensikten er å kunne drøfte hva det er som påvirker tidsverdiene, hvilket ikke er mulig å se så tydelig i et virkelig transportmodellsystem. Materialet fra den norske tidsverdiundersøkelsen gir oss mulighet til å studere hvordan tidsverdiene avhenger av objektive kjennetegn, men ikke av subjektive holdninger og individuelle omstendigheter.

Vår modell kan sees som en variant av de Serpa (1971), som fremdeles brukes til teoretisk drøfting av tidsverdier. Mens de Serpa definerer nyttefunksjonen over ulike anvendelser av tida, dvs. aktiviteter som en kan bruke større eller mindre tid på, vil vi anta at aktivitetene tar en gitt tid pr. gang, men kan gjennomføres et valgfritt antall ganger. Vi kan betrakte dette som reiser med en på forhånd kjent og fast reisetid. En annen forskjell er at vi utelater reine fritidsaktiviteter fra nyttefunksjonen og bibetingelsene. Det spiller en rolle, men ikke for det vi er opptatt av her.

La $z$ være generelt forbruk, $\mathbf{x}$ en vektor av reiser og $t_w$ tid brukt i lønnsarbeid. Nyttefunksjonen er $U = z + v(\mathbf{x}, t_w)$. Prisen pr. reise av type $i$ er $p_i$, timelønna er $w$, arbeidsfri inntekt er $R$ og reisetida for reise i er $t_i$. Vi skal maksimere $U$ gitt et pengebudsjett, et tidsbudsjett og et krav om et minste antall timer brukt på arbeid og på andre aktiviteter (reiser) av hver type.

$$
\begin{aligned}
&\underset{z,\mathbf{x},t_w}{Maks}\, U = z + v(\mathbf{x}, t_w) \\
\text{gitt}\quad & z + \mathbf{px} \le R + wt_w && (\lambda) \\
& t_w + \mathbf{tx} \le T && (\mu) \\
& -t_w \le -b_w && (\varphi_w) \\
& -t_i x_i \le -b_i, \quad i = 1,\ldots,n && (\varphi_i)
\end{aligned}
$$

Vi danner Lagrangefunksjonen $\mathcal{L}$ og deriverer. Kuhn-Tuckerbetingelsene for maksimum kan skrives

$$(1)\quad \frac{\partial \mathcal{L}}{\partial z} = 1 - \lambda \le 0 \quad (= 0 \text{ for } z > 0)$$

$$(2)\quad \frac{\partial \mathcal{L}}{\partial t_w} = \frac{\partial v}{\partial t_w} + \lambda w - \mu + \varphi_w \le 0 \quad (= 0 \text{ for } t_w > 0)$$

$$(3)\quad \frac{\partial \mathcal{L}}{\partial x_i} = \frac{\partial v}{\partial x_i} - \lambda p_i - \mu t_i + \varphi_i t_i \le 0 \quad (= 0 \text{ for } x_i > 0) \quad i = 1,\ldots,n$$

$$
(4)\quad
\begin{aligned}
&\lambda \ge 0 \quad (= 0 \text{ for } z + \mathbf{px} < R + wt_w) \\
&\mu \ge 0 \quad (= 0 \text{ for } t_w + \mathbf{tx} < T) \\
&\varphi_w \ge 0 \quad (= 0 \text{ for } t_w > b_w) \\
&\varphi_i \ge 0 \quad (= 0 \text{ for } t_i > b_i) \quad i = 1,\ldots,n
\end{aligned}
$$

Anta $z > 0$ og $t_w > 0$. Vi har da likhet i (1) og (2) og følgelig $\lambda = 1$ og $\mu = w + \frac{\partial v}{\partial t_w} + \varphi_w$.

Setter vi dette inn i (3), har vi, for *i* = 1,…, n,

$$(5)\quad \frac{\partial v}{\partial x_i} \leq p_i + (\mu_i - \varphi_i)t_i = p_i + \left( w + \frac{\partial v}{\partial t_w} + (\varphi_w - \varphi_i) \right) t_i = p_i + \omega t_i = g_i$$

der det nest siste likhetstegnet definerer tidsverdien $\omega$ og det siste definerer generalisert kostnad $g_i$.

Vi ser at arbeidsfri inntekt ikke inngår i tidsverdien. Timelønn inngår på en avgjørende måte, men tidsverdien er ikke proporsjonal med timelønna. Behaget eller ubehaget ved å være på arbeid inngår som et tillegg eller fratrekk til timelønna. Til slutt er det også avgjørende om individet helst ville ha jobbet mindre enn det må, og om det helst ville brukt mindre tid under reise enn det må. Dvs. det er differansen mellom disse skyggeprisene som har betydning. Om minstekravet til å bruke en viss tid på arbeid er en viktigere hindring for nyttemaksimeringen enn kravet til å bruke tid på reise, er det siste leddet i tidsverdien positivt, i motsatt fall er det negativt.

# 4 Etterspørselsfunksjonenes form

Vi viser i vedlegget at budsjettbetingelsen ikke kommer inn i bildet ved løsningen av *x*-ene, bare ved fastleggingen av *z* til slutt. Etterspørselsfunksjonene **x** vil derfor være funksjoner av prisene **p** og timelønna *w*, men ikke av *R*. Vi kan skrive løsningen slik:

$$\mathbf{x} = \mathbf{D}(\mathbf{g})$$

der $\mathbf{x} = (x_1,\ldots,x_n)$, $\mathbf{D} = (D_1(\mathbf{g}),\ldots,D_n(\mathbf{g}))$, $\mathbf{g} = (g_1,\ldots,g_n)$ og $g_i = p_i + \omega t_i$ for alle *i*.

Dersom nyttefunksjonen ikke hadde vært kvasilineær[47], ville etterspørselsfunksjonene likevel være funksjoner av generaliserte kostnader, dvs. pris og reisetid vil inngå lineær slik som i (5). Det er de fire bibetingelsene som gir denne strukturen. Men i de generaliserte kostnadene ville pris og timelønn vært vektet med lagrangeparameteren $\lambda$, som da ikke ville være 1, og ikke engang en konstant. $\lambda$ vil være en funksjon av blant annet *R*, hvilket ville gjøre de generaliserte kostnadene til noe som varierte fra individ til individ. Etterspørselsfunksjonene ville dessuten trolig også være eksplisitte funksjoner av inntekt.

[47] Om kvasilinearitet, se under.

# 5 Implikasjoner for tidsverdiens utvikling over tid

Ut fra likning (5) kan vi trekke konklusjoner om tidsverdiens utvikling over tid. Anta at vi ønsker å oppjustere tidsverdiene med inntektsutviklingen. For det første ser vi at arbeidsfrie inntekter er irrelevante. Det betyr at inntekstbegrepet som vi bruker til oppjusteringen, ikke bør være BNP per capita eller disponibel husholdningsinntekt per capita, men *reallønn per time etter skatt*.

For det andre må vi innse at det finns grupper med et nyttemaksimeringsproblem der timelønna ikke spiller noen rolle, fordi de ikke jobber eller jobben spiller liten rolle for dem. Det gjelder blant annet trygdede og pensjonister, og i en viss grad studenter, skoleelever, hjemmeværende og andre som helt eller delvis blir forsørget av andre. For disse vil tidsverdien være definert ved $\mu_i - \varphi_i$, uten at det er mulig å knytte denne differansen mellom skyggepriser til timelønna i det hele tatt. Siden disse i høyeste grad er overrepresentert i gruppene med lavest inntekt, kan vi vente at tidsverdiene for grupper med lav inntekt i større grad utvikler seg uavhengig av inntektsutviklingen, enten vi bruker timelønn eller disponibel inntekt eller et annet inntektsmål. Og det er akkurat dette vi finner i den metodologisk mest tilfredsstillende undersøkelsen av dette spørsmålet (Börjesson 2010b).

For det tredje bør vi kunne gjøre oss tanker om hvordan andre sider av samfunnsutviklingen enn inntektsutviklingen påvirker tidsverdien. Sambandet mellom hvor mange timer vi jobber og hva vi har i inntekt er i ferd med å bli mer uklart. Det er færre som må stemple. Det er flere som har månedslønn, og sjøl om det forutsettes at man jobber et visst antall timer per måned og eventuelt at man må dokumentere det, er det en tendens til svakere kontroll med at det faktisk skjer. Det kan bety at en større del av inntekten *oppleves* som arbeidsfri i den forstand at den ikke er knyttet til et bestemt timeforbruk. Med andre ord: $R$ vokser på bekostning av $w$ i den subjektive betraktningen som vårt nyttemaksimeringsproblem skal være en modell av. Dette kan gi tidsverdien en lavere elastisitet med hensyn på offisielt registrert lønnsinntekt.

De to skyggeprisene i tidsverdien vil reduseres i samme grad som beskrankningene oppheves eller føles mindre ubehagelige. Hvis alle kunne velge arbeidstid fritt, ville $\varphi_w$ bli null. Det reduserer tidsverdien. Trolig går samfunnet i en retning der flere kan tilpasse arbeidstida fritt. Den siste beskrankningen er det kanskje vanskeligere å oppheve – det kommer an på hvordan vi tolker $x_i$. Hvis $x_i$ er en reise fra A til B, finns det en viss, men begrenset mulighet til å lette på restriksjonen ved å redusere $t_i$. En mye større mulighet vil ligge i å gjennomføre reisa sjeldnere, og en slik tendens finnes for arbeidsreiser. Videre: Hvis $x_i$ er en aktivitet som kan utføres på flere steder, kan $t_i$ tolkes som tida det tar å utføre aktiviteten, inkludert eventuell reisetid. Tendensen til å kunne drive flere aktiviteter hjemmefra vil da – hvis vi ikke utvider vårt repertoar av aktiviteter – kunne redusere eller oppheve beskrankningen. Den motsatte tendensen finnes også, idet vi reiser lengre av sted for å gjennomføre mange av våre aktiviteter nå til dags. Vi er ikke lenger fornøyd bare med tilbudet i nærområdet.

Alt i alt kan en kanskje gjette at $\varphi_w$ vil reduseres raskere enn $\varphi_i$, hvilket vil bidra til lavere tidsverdier.

Leddet $\partial v / \partial t_w$ er den marginale nytten av å utvide arbeidstida, helt bortsett fra hva det vil kunne bringe med seg av lønn og av kortere tid til andre aktiviteter. I optimum vil den være negativ og kanskje i tallverdi i samme størrelsesorden som $w$, i alle fall om vi kan se bort fra andre restriksjoner. Vi kan derfor vente at tidsverdien i de aller fleste aktiviteter er mye mindre enn timelønna. Det er også hva vi alltid finner. Men det er fullt ut mulig at det vil bli triveligere på jobben i framtida, slik at hvis vi ikke får dårligere tid fordi vi har viktigere ting å gjøre, vil dette leddet bidra til å trekke tidsverdien oppover i framtida. Men sikkert er det jo ikke.

Alt i alt må det være grunn til å anta at tidsverdien vil endre seg med inntektsutviklingen og med samfunnsforholdene forøvrig, men det er ingen grunn til at det vil skje proporsjonalt med inntekten, uansett hvordan den defineres. Sammensetningen av befolkningen på lønnstakere (og på timelønte lønnstakere) og andre vil ha stor betydning, som det også er påvist i den svenske undersøkelsen av dette.

# 6 Gormans polære form

Grunnlaget for nyttekostnadsanalyse er Kaldor-Hickskriteriet. Det sier at dersom de som vinner på et tiltak, kan holde taperne skadesløse uten at de sjøl dermed slutter å være vinnere, så skal tiltaket regnes som en gevinst for samfunnet. Siden kriteriet på ingen måte krever at taperne blir kompensert i praksis, innebærer det en form for likegyldighet til hvordan gevinsten faktisk fordeles. Så lenge det er en potensiell gevinst for samfunnet sett under ett, er kriteriet oppfylt. Dette er det svakest mulige kriteriet på samfunnsøkonomisk lønnsomhet, og det som er det enkleste å verifisere. En hvilken som helst interesse for hvordan godene er fordelt vil gjøre det betydelige mer komplisert å trekke slutninger om samfunnsøkonomisk lønnsomhet.

Det kan vises at et tiltak oppfyller Kaldor-Hickskriteriet (i en høvelig presisering) *hvis og bare* hvis det gir økning i en utilitaristisk velferdsfunksjon som har Gormans polære form (Chipman og Moore 1994). For at det skal være tilfelle må alle individer eller individuelle husholdninger også ha indirekte nyttefunksjoner av Gormans polære form, dvs. nyttefunksjonene har forma

$$V_h(\mathbf{p}, R) = a_h(\mathbf{p}) + b(\mathbf{p}) R_h,$$

der **p** er en vektor av priser og $R_h$ er disponibelt forbruksbudsjett for individ eller husholdning $h$. Funksjonen $a_h(.)$ er altså spesifikk for individ $h$, mens funksjonen b(**p**) er felles for alle individer. Ved summering over alle individer framkommer velferdsfunksjonen

$$V(\mathbf{p}, R) = \sum_h a_h(\mathbf{p}) + b(\mathbf{p}) \sum_h R_h$$

Vi ser at når individene har nyttefunksjoner av Gormans polære form, finns det en utilitaristisk velferdsfunksjon av samme form. Som sagt innebærer det at vi har en praktisk måte å beregne om Kaldor-Hickskriteriet er oppfylt på. Omvendt gjelder at hvis de individuelle nyttefunksjonene ikke har denne forma, vil summen av dem (eller en monoton transformasjon av summen) heller ikke ha det, og vi har da ingen velferdsfunksjon som er ekvivalent med Kaldor-Hickskriteriet.

Når velferdsfunksjonen har Gormans polære form, sier vi at det finnes en representativ konsument. Den indirekte nytten til den representative konsumenten er summen av individenes nytte for så vidt som den bare avhenger av prisene pluss summen av alle inntekter multiplisert med en faktor som bare avhenger av prisene. Vi ser umiddelbart at inntektsfordelingen ikke har noen betydning for velferden beregnet med denne funksjonen. Et annet karakteristisk trekk er at sjøl om individene har ulik inntekt og ulik smak, er de like i den forstand at om de får en krone til i inntekt, vil de alle sammen bruke den på samme måte. (Den deriverte av etterspørselsfunksjonene med hensyn på inntekt er lik for alle og uavhengig av inntekt.)

Som en vil skjønne, er det ikke sannsynlig at individene har slike nyttefunksjoner, og derfor heller ikke sannsynlig at det eksisterer en representativ konsument eller at individenes nytte kan legges sammen til en velferdsfunksjon som er ekvivalent med Kaldor-Hickskriteriet.

# 7 Kvasilineær nytte

Individ *h* har en kvasilineær nyttefunksjon hvis den indirekte nytten kan skrives

$$V_h(\mathbf{p}, R) = a_h(\mathbf{p}) + R_h$$

Åpenbart er dette et spesialtilfelle av Gormans polære form, og følgelig enda mindre realistisk. Vi ser at den indirekte nyttefunksjonen måler nytten i kroner. Grensenytten av inntekt er 1 for alle med slike nyttefunksjoner, og dermed også for den representative konsumenten, om hun finns.

Denne indirekte nyttefunksjonen er løsninga på problemet

$$\underset{z,\mathbf{x}}{Maks}\, U_h = z + v_h(\mathbf{x})$$
$$\text{gitt} \quad z + \mathbf{px} \leq R_h$$

eller et liknende problem, eventuelt med flere bibetingelser og/eller med arbeidstid og lønnsinntekt som variable. Det gir etterspørselsfunksjoner etter alle varer **x** som er uavhengig av budsjettet *R*. Bare etterspørselen etter den siste varen, z, er en funksjon av inntekt. Også i dette spesialtilfellet vil individene oppføre seg likt dersom de får en ekstra krone. I dette tilfellet vil ingen av dem endre etterspørsel etter **x**, og alle vil bruke de ekstra pengene på z.

La oss nå si at vi ikke er særlig interessert i etterspørselen etter z. Det er bare en indikator på forbruk av goder som ligger utenfor vårt primære studiefelt. Med andre ord, det er et uspesifisert sammensatt forbruk, mens **x** er de godene vi virkelig vil studere. La oss videre anta at vi er langt mer interessert i prisendringer enn i inntektsendringer. Så lenge inntekten er konstant, kan vi se bort fra den som et argument i etterspørselsfunksjonen. En annen måte å si det på, er at vi studerer virkninger av prisendringer på goder som tar en liten del av inntekten.

Under slike forutsetninger blir kvasilineær nytte en interessant modell. Nytten er umiddelbart uttrykt i kroner, det finns en opplagt velferdsfunksjon som er ekvivalent med Kaldor-Hicks-kriteriet, det er ingen forskjell på ekvivalent og kompenserende variasjon, osv.

# 8 Egenskaper ved vanlige transportmodeller

I en vanlig transportmodell er arealbruk, reisemålenes attraktivitet, befolkningene i ulike soner, sosioøkonomiske kjennetegn ved sonebefolkningene (som husholdsinntekt, førerkortinnehav og bilhold) alt sammen innputt sammen med egenskapene ved transporttilbudet (som drivstoffpriser, vegnettets kapasitet og standard, kollektivpris og kollektivfrekvens). De store beslutningene som virkelig gjør en forskjell for hva som er disponibelt til daglig forbruk – hvor skal jeg bo, hva slags bil skal jeg ha – er altså eksogene variable, eller *så godt som* eksogene variable. Det finnes for eksempel en førmodell for førerkortinnehav og bilhold til den nasjonale og de regionale persontransportmodellene, men i en konkret analyse holdes bilhold og arealbruk konstant over alle alternativer, og det foregår ingen nytteberegning av å anskaffe bil(er) eller omlokalisere aktiviteter.

Innafor en slik ramme er det ikke urimelig at etterspørselen etter reiser ikke er en funksjon av inntekt. Og det er den da heller ikke i transportmodellsystemene. Riktignok inngår inntekt som en sosioøkonomisk variabel som er med på å bestemme transportmiddelvalg, for eksempel, men inntekten stammer i det tilfellet ikke fra noen budsjettbetingelse, og må tolkes som en rein smaksvariabel. (Rike folk gjør andre valg enn fattige, men ikke nødvendigvis fordi de har mer penger.)

Vi har altså at transportmodellene fungerer som om de er resultatet av en kvasilinær nyttefunksjon. Dette har to viktige implikasjoner. For det første er det en måte å separere transportsektoren fra andre sektorer på. Med kvasilineær nytte er det helt i orden å gjennomføre en partiell analyse uten å trekke inn andre trekk ved resten av økonomien enn skattefaktoren, overføringer til og fra det offentlige og eksterne virkninger. For det er det slik at så lenge transportmodellene har denne karakteren, vil nytteberegningen kunne foretas uten å vurdere vekting av individene.

Når det gjelder inntektsutviklingen av tidsverdiene, reiser den kvasilineære karakteren av transportetterspørselssystemene som vi bruker, spørsmålet om ikke hele distinksjonen mellom individuelle og samfunnsmessige tidsverdier, slik det framstilles i Mackie m,fl. (2001), Börjesson (2010a) eller for den saks skyld Nyborg (2002), faller vekk. Det gjør den, med mindre vi skal bygge nyttekostnadsanalysen på andre prinsipper enn de som er nedfelt i modellsystemet. Ingen omregning, ingen vekting av individer i den statiske sammenhengen. Det logiske er da at Börjessons argument mot å inntektsjustere tidsverdiene også faller, fordi det bygger på at det er nødvendig å regne om slik at en time spart teller likt uansett inntekt.

# 9 Konklusjon

En litt for lang og springende drøfting av spørsmålet om tidsverdiene bør oppjusteres med framtidig forventet inntektsutvikling, har gitt følgende konklusjoner:

1. Oppjustering vil øke den beregnede lønnsomheten av alle infrastrukturprosjekter, men ikke forrykke forholdet mellom dem. Siden nåværende samferdselsbudsjetter vil kunne gi plass til de nye prosjektene som blir lønnsomme dersom de som framleis er ulønnsomme tas ut, har oppjusteringen derfor trolig mindre betydning for den praktiske politikken.

2. I konsumentteoretiske modeller vil tidsverdien både ha forbindelse med timelønna og med den inntekten som er disponibel for forbruk. Timelønnas innvirkning på tidsverdien vil imidlertid bli motvirket av en ulempe eller belastning ved å bruke tid på arbeid. Det er derfor grunn til å anta at tidsverdien tallmessig bare er en brøkdel av timelønna og at elastisiteten med hensyn på timelønn skal være mindre enn 1. Budsjettskrankens innvirkning avhenger av modellen. Med sterkt separable (kvasilineære) nyttefunksjoner vil budsjettskranken ikke ha noen betydning for tidsverdien. I motsatt fall vil skyggeprisen på budsjettskranken virke inn slik at jo høyere inntekt, jo høyere tidsverdi.

3. I diskrete valgmodeller er tidsverdien parameteren tilknyttet reisetida delt på parameteren tilknyttet reisekostnaden, eller grensenytten av tid delt på grensenytten av penger, som man sier. Disse parametrene kan igjen – eksplisitt eller implisitt – være funksjoner av inntekt. Ved tolkningen forutsetter man at grensenytten av inntekt er lik grensenytten av reisekostnadene med motsatt fortegn. Det er en tvilsom forutsetning hvis folk ikke er strengt rasjonelle og godt informert om sine utgifter, eller om det finns andre budsjettskranker enn totalskranken på forbruksbudsjettet.

4. I empiriske undersøkelser finner vi at vil individuelle tidsverdier på et gitt tidspunkt til en viss grad avhenger av respondentenes inntekt. Dette forklarer noe av de estimerte tidsverdiforskjellene mellom transportmåtene, men er antakelig ikke hovedgrunnen til disse forskjellene. Mesteparten av de gjennomførte studiene gir elastisiteter av tidsverdiene med hensyn på inntekt på rundt 0,5 eller litt mindre. Det gjelder også norske

tidsverdistudier. Metodemessig har de empiriske undersøkelsene utviklet seg fra binomisk logit til mixed logit eller semiparametriske metoder, uten at det har gitt fundamentalt nye konklusjoner om tidsverdiens inntektselastisitet.

5. Empiriske undersøkelser gir ikke noe entydig svar på hvordan tidsverdiene utvikler seg over tid når gjennomsnittlig inntekt øker. Alt av elastisiteter fra 0 til 1 er funnet, men med en viss tendens til høyere verdier når metodene har blitt bedre eller studiene som sammenliknes har blitt likere. Den svenske undersøkelsen av dette gjentar undersøkelsen fra 1994 i minste detalj, og kommer fram til at lavinntektsgruppene har elastisiteter nær null og høyinntektsgruppene har elastisiteter nær 1. Denne forskjellen virker nå godt etablert og kan forklares teoretisk med forskjeller i andelen av inntekten som er arbeidsinntekt. Tidsverdienes elastisitet med hensyn på gjennomsnittsinntekten i samfunnet vil dermed være mindre enn 1 og avhenge av inntektsfordelingen, og i bunn og grunn av hvor nært personlig inntekt er knyttet til timeforbruk.

6. Empiriske undersøkelser tyder videre på at det spiller en rolle hva slags inntektsbegrep man bruker. Elastisiteten med hensyn på disponibel personlig inntekt vil være høyere enn elastisiteten med hensyn på husholdningsinntekt eller med hensyn på brutto nasjonalprodukt pr. innbygger.

7. Når tidsverdiene inngår i prognosemodeller finns det ikke fordelingspolitiske argumenter for å endre de verdiene som er funnet empirisk. Når de inngår i samfunnsøkonomiske analyser basert på resultatene av prognosemodellene, derimot, hevdes det svært ofte at vi bør eliminere ulikheter i tidsverdiene som har sitt opphav i inntektsulikheter. En time spart skal telles likt uansett inntekten til den som får denne gevinsten. I mange land er dette grunnlaget for at man regner med samme tidsverdi på alle transportmåter.[48] Dette kan gjøres ved å beregne en vektet gjennomsnittlig tidsverdi, eller man kan bruke et prinsipp om at *samfunnets* grensenytte av tid skal være den samme uansett hvem som får tidsgevinsten. Individuelle tidsverdier eller gruppetidsverdier må da omregnes.

   Det kan hevdes at om man foretar en slik omregning for å oppnå likhet mellom inntektsgruppene nå, bør man også gjøre det i forholdet mellom de som lever nå og de som kommer seinere. Dette er essensen i Börjessons argument mot å oppjustere tidsverdiene med inntektsutviklingen (Börjesson 2010a og b).

8. I Norge går vi relativt langt i å beholde tidsverdiforskjellene mellom transportmåtene, uansett i hvilken grad de skyldes inntekt. Om vi skal være konsistente må vi da avvise Börjessons argument. Om vi derimot ønsker å rense tidsverdiene for inntektsforskjeller, vil det ha implikasjoner for oppjusteringene. Det tekniske problemet om hvordan vi skal rense for inntektsforskjeller, bør være løsbart, men er ikke løst.

9. COWI har kommet med anbefalinger om oppjusteringer av tidsverdier og andre parametre som avhenger av betalingsvillighet (COWI 2010). Det ser ut til at de i hovedsak har holdt seg til britiske resultater og anbefalinger, mens de siste svenske resultatene

[48] Argumentet står sterkere når man i tillegg finner så små forskjeller mellom transportslagene at det gjør liten forskjell om man setter samme verdi.

og anbefalingene er satt tilside. De underliggende fordelingspolitiske premissene for britiske anbefalinger er ikke klarlagt. Vi har ikke noe annet forslag, men trur at en bør se på saka en gang til. Det bør avklares om nytteberegningen skal være konsistent med den underliggende kvasilineære prognosemodellen eller om den skal bygge på andre prinsipper i tillegg. Det bør forsøkes på en systematisk sammenlikning av verdiene i Ramjerdi (1997) og Ramjerdi m.fl. (2010). Det finns også ikke analyserte data fra 2010 som kan gi nye estimater.

## Litteratur


Abrantes, P. and M. Wardman (2011) Meta-analysis of UK Values of time: an update, *Transportation Research A* **45** (1), 1-17.

Börjesson M (2010a) Inter-temporal variation in the marginal disutility of travel time and travel cost. Paper presented at 12th WCTR, Lisbon.

Börjesson M (2010b) Swedish values of travel time and their application in appraisal. Working paper, Centre of Transport Studies, KTH, Stockholm.

Börjesson M, M Fosgerau and S Algers (2009) The income elastic of the value of travel time is not one number. Presented at the 2009 European Transport Conference, Leiden, Netherlands.

Chipman JS and JC Moore (1994) The Measurement of Aggregate Welfare. In: Eichhorn, W. (Ed.), Models and Measurement of Welfare and Inequality. Springer-Verlag, Berlin.

COWI (2010) Realprisjustering av enhetskostnader over tid. Statens vegvesen, rapport.

de Serpa, A (1971) A theory of the economics of time. *The Economic Journal* **81**, 828-845.

Fosgerau, M. (2005) Unit income elasticity of the value of travel time savings, Urban/Regional 0508007, EconWPA.

Mackie PJ, S Jara-Diaz og AS Fowkes (2001) The value of travel time savings in evaluation. *Transportation Research E* **37**(2-3), 91-106,

Mackie, PJ, AS Fowkes, M Wardman, G Whelan, J Nellthorp and J Bates (2003) Value of travel time savings in the UK: Summary report. http://webarchive.nationalarchives.gov.uk/+/http:/www.dft.gov.uk/pgr/economics/rdg/valueoftraveltimesavingsinth3130

Nyborg K (2002) Miljø og nyttekostnadsanalyse. Noen prinsipielle vurderinger. Rapport 5/2002. Frischsenteret.

Ramjerdi F, IA Sætermo and K Sælensminde (1997) The Norwegian Value of Time Study. TØI Report 379/2007.

Ramjerdi F, S Flügel, M Killi og H Samstad (2010) Den norske verdsettingsstudien. Tid. TØI-rapport 1053b.

Wardman M (2001) Inter-temporal variations in the value of time. ITS Working Paper 566, Institute of Transport Studies, University of Leeds.

Wardman M (2004) Public transport values of time. *Transport Policy* **11**, 363-377.

WSP Analys & Strategi (2010) Trafikanters värdering av tid – den nationella tidsvärdestudien 2007-2008. WSP rapport 2010:11.

**VEDLEGG**

## Løsningen av optimeringsproblemet

Kuhn-Tuckerbetingelsene (1)-(4) gir oss ledetråder til hvordan nyttemaksimeringsproblemet skal løses i et praktisk tilfelle. Vi har antatt $z > 0$ og $t_w > 0$, slik at vi har likhet i (1) og (2). Anta nå først at også alle $x_i > 0$, slik at også alle de $n$ relasjonene (3) er likninger.

For alle $i$ der bibetingelsen $-t_i x_i \leq -b_i$ er bindende, er $x_i$ gitt av denne bibetingelsen. Vi veit ikke ennå hvilke $x$-er det er som er bestemt av en bindende bibetingelse, men vi kan anta at vi veit det. La oss si det er $k$ slike. Vi har da allerede funnet løsningen for disse, men mangler løsningen for de $n - k$ $x$-ene som ikke er bundet av bibetingelsen og de $k$ $\varphi$-ene som ikke er null. For de $n - k$ ukjente $x$-ene er $\varphi_i$ lik null, slik at vi har nøyaktig $n$ likninger til å beregne de $n - k$ ukjente $x$-ene og de $k$ $\varphi$-ene som ikke er null. Alle disse $n$ variablene og parametrene vil naturligvis fremdeles være funksjoner av den ukjente $\mu$.

Men vi har igjen bibetingelsen $t_w + \mathbf{tx} = T$ og likning (2). Nå må vi sette inn $x$-ene som funksjon av $\mu$ som vi alt har funnet i disse to likningene og bruke dem til å bestemme $t_w$ og $\mu$. Det vil si: Hvis $-t_w \leq -b_w$ er en bindende bibetingelse, er jo $t_w$ allerede bestemt, og likning (2) og $t_w + \mathbf{tx} = T$ vil gi oss $\mu$ og $\varphi_w$, men hvis arbeidstida er fri, vil $\varphi_w$ være null og de to likningene vil gi oss $t_w$ og $\mu$. Til slutt vil $z$ bli bestemt av den første bibetingelsen, budsjettbetingelsen.

Det store problemet er at vi ikke kan vite på forhånd hvilke bibetingelser som er bindende og hvilke som ikke er det. Vi må altså bruke en algoritme eller i verste fall teste alle mulige kombinasjoner av bindende og ikke bindende bibetingelser. Men dette behøver ikke bekymre oss nå, for poenget er at for hver kombinasjon vi tester, går vi fram på den rekursive måten som er beskrevet. Budsjettbetingelsen kommer ikke inn i bildet ved løsningen av $x$-ene, bare ved fastleggingen av $z$ til slutt. Det gjelder da naturligvis også den kombinasjonen som vil vise seg å være den beste. Det finns én eneste parameter som bare inngår i budsjettbetingelsen og ingen andre steder, og det er $R$. Etterspørselsfunksjonene $\mathbf{x}$ vil derfor være funksjoner av prisene $\mathbf{p}$ og timelønna $w$, men ikke av $R$. Vi kan skrive løsningen slik:

$$\mathbf{x} = \mathbf{D}(\mathbf{g})$$

der $\mathbf{x} = (x_1,\ldots,x_n)$, $\mathbf{D} = (D_1(\mathbf{g}),\ldots,D_n(\mathbf{g}))$, $\mathbf{g} = (g_1,\ldots,g_n)$ og $g_i = p_i + \omega t_i$ for alle $i$.

## 3.3 Nyttekostnadsanalyse og flertallsavgjørelser[49]

[49] Dette notatet sammenfatter en ide som opprinnelig er drøftet i artikkelen «Betydningen av samfunns-økonomisk lønnsomhet ved prioritering av prosjekter i Nasjonal transportplan», som er kapittel 5 i Odeck og Welde (red.) (2015): Ressursbruk i transportsektoren. Concept-rapport nr. 44, NTNU.

# Nyttekostnadsanalyse og flertallsavgjørelser

## 1 Tidsverdien er skjevfordelt

Et funn fra den norske tidsverdiundersøkelsen fra 2010 er at tidsverdien ikke er symmetrisk fordelt blant trafikantene. De fleste har mindre tidsverdi enn gjennomsnittet. Medianen er den observasjonen som ligger i midten, i den forstand at halvparten av de observerte tidsverdiene er lavere og halvparten er høyere. Det vi fant var altså at medianverdien var lavere enn den gjennomsnittlige tidsverdien. Faktisk var medianen bare rundt 0,7 ganger så stor som gjennomsnittsverdien for de fleste reisemåter, og bare rundt 0,6 for korte bilreiser (Ramjerdi m.fl. 2010, tabell 5.1).

## 2 Tolkning

Her er en tolkning av hva dette betyr: Anta at det finns et tiltak som vil gi en tidsbesparelse i samme størrelsesorden som den som ligger til grunn for tabell 5.1 i Ramjerdi m.fl. Anta videre at tiltaket skal finansieres med en avgift som er den samme for alle reisende som opplever tidsbesparelsen. Dersom den nødvendige avgiften er mindre enn verdien av tidsbesparelsen beregnet med mediantidsverdien, vil dette tiltaket få flertall i en avstemning der alle reisende og ingen andre er med. Men hvis den nødvendige avgiften er større enn dette, vil tiltaket bli nedstemt. I denne forstand gir mediantidsverdien et likevektspunkt – en såkalt Bowenlikevekt.

Hvis derimot tiltaket blir bedømt med den høyere *gjennomsnittstidsverdien*, vil det bli anbefalt gjennomført sjøl om avgiften er høyere enn det som skal til for at det får flertall i en avstemning, så lenge den ikke er høyere enn gjennomsnittsverdien. I samfunnsøkonomisk forstand er det gjennomsnittstidsverdien som er den rette, fordi gjennomsnittet multiplisert med antall reisende er lik den totale betalingsvilligheten for tiltaket. Hvis betalingsvilligheten er større enn kostnaden, er det et samfunnsøkonomisk lønnsomt prosjekt.[50]

Vi kan kalle beslutninger med medianverdien som tidsverdi for demokratisk funderte, og beslutninger med gjennomsnittsverdien som tidsverdi for samfunnsøkonomisk funderte. Demokratisk funderte beslutninger slipper altså gjennom færre prosjekter enn samfunnsøkonomisk funderte beslutninger. Faktisk vil det demokratisk funderte synspunktet innebære at en krone samfunnsøkonomisk nytte ikke skal regnes som mer verdt enn 70 øre. Av prosjekter som kan gi kortere reisetid med bil på korte turer, vil enda færre (i forhold til samfunnsøkonomisk riktig nivå) bli gjennomført, dersom tingene skal avgjøres med flertallsvedtak blant de som får gevinsten, og de som får gevinsten også er de som må betale.

## 3 Et paradoks?

Jeg vil understreke at dette er et resultat vi kan stole på. Det har framkommet med hjelp av de beste metoder, og det virker robust. Likevel vil det slå de fleste som absurd. Vi veit jo at

[50] At prosjektet blir bedømt med gjennomsnittsverdien, er ikke det samme som at gjennomsnittsverdien blir brukt som avgift. Siden vi ikke har sagt om det forekommer køer eller andre avgifter, er det trolig best i samfunnsøkonomisk forstand om det ikke blir innkrevd noen avgift.

politikerne med god samvittighet sier ja til en mengde prosjekter som har blitt beregnet til å være samfunnsøkonomisk ulønnsomme. Det ser ut til at demokratiske prosesser slipper gjennom flere prosjekter enn lønnsomhetsberegningene i dag vil gjøre. Og vi veit at det er et stort press for høyere tidsverdier og andre endringer i nyttekostnadsmetodikken som kan gjøre at mange flere prosjekter kan bli vurdert som lønnsomme.

# 4 Én forklaring …

Det finns to mulige forklaringer på dette paradokset.

Den ene er at det vi i Norge i dag oppfatter som demokratiske prosesser på samferdselsområdet, er meget langt fra den idealiserte prosessen jeg her har beskrevet, der de som berøres, sjøl tar avgjørelsen og deler på regninga. På en eller annen måte må folk oppfatte kostnadene som mindre enn de er. Dersom tiltakene ikke finansieres helt ut med bompenger eller trafikant-betaling, kan det skyldes at et tiltak i mitt fylke eller mitt distrikt i alt vesentlig betales over skatteseddelen av folk andre steder. De opplevde kostnadene ved å få et prosjekt i mitt fylke inn i planen, er derfor mye mindre enn de virkelige samfunnsøkonomiske kostnadene. Det vil derfor oppstå et spill mellom distriktene om å få inn sine prosjekter i planen. I dette spillet vil det måtte inngås flertallskoalisjoner av distrikter som støtter hverandres prosjekter uansett hvor små gevinster eller hvor høye kostnader de medfører. Det samlede resultatet blir at det blir flere prosjekter enn flertallsbeslutninger skulle tilsi (hvilket et stykke på veg kan være bra), og at mange dårlige prosjekter kommer inn i planen i stedet for noen som er bedre (og det er ikke så bra).

Høy bompengeandel vil utvilsom kunne motvirke denne tendensen til at samferdselsplan-leggingen blir et spill. Men sjøl i bompengeprosjektene er det muligheter for å få andre til å betale store deler av regninga. Bommene kan legges på grensa mot fylker og kommuner som ikke er med i beslutningsprosessen, slik at utenforstående får en høy del av regninga. Dette skjer ofte, til tross for det såkalte nytteprinsippet, som sier at de som ikke får noen gevinst, heller ikke skal måtte betale. Det kan også gis omfattende rabatter til lokalbefolkningen som bruker infrastrukturen daglig. Dermed blir forholdet mellom gevinster og kostnader annerledes for disse enn for gjennomsnittet. Dette kan gi flertall for prosjekter som i virkeligheten koster mer enn det medianvelgeren er villig til å betale. Endelig kan villigheten til å betale bompenger brukes i et spill mellom distriktene og et spill mot staten om å få prosjekter inn i planen. Når de vel er innafor, kan de arbeide videre for å få opp andelen som skal betales av staten.

For lave kostnadsanslag er en meget viktig form for undervurdering av de virkelige samfunns-økonomiske kostnadene. Ofte øker kostnadene mye fra planlegging starter til vedtak om byg-ging, og fra vedtak om bygging til ferdigstillelse. Dette gir en tendens til å vedta flere prosjekter enn vår stiliserte modell med flertallsavgjørelse blant de som skal betale skulle tilsi. Og det gir en mulighet til å kreve at overskridelsene skal tas ved ekstrabevilgninger fra staten i stedet for økte bompenger.

# 5 … og én til

Ut fra vår idealiserte demokratimodell var det et paradoks at det snarere vedtas for mange prosjekter enn for få, og at lønnsomhetskravet snarere settes for lavt enn for høyt. Den første mulige forklaringen på paradokset var at kostnadene undervurderes, eller at beslutningstakerne ikke behøver å ta alle kostnader i betraktning. Den andre mulige forklaringen er at nytten overvurderes. Det kan gjøres på to måter. Den vanligste er å vise til at samferdselsinfrastruktur er vesentlig for regional og nasjonal økonomisk vekst, og at dette ikke er med i de samfunnsøkonomiske kalkylene. Det er derfor liten grunn til å ta hensyn til regnestykkene som viser at et prosjekt er ulønnsomt. Det vil uansett være tvingende nødvendig å gjennomføre prosjektet om ikke regionen eller nasjonen skal komme i bakleksa.

Uansett hva visse pensjonerte økonomiprofessorer og NHO måtte mene, er det intet vitenskapelig belegg for at et prosjekt med beskjedne direkte nyttevirkninger (i form av tidsbesparelser, økt pålitelighet osv.) skulle være vesentlig for økonomisk vekst eller konkurranseevne i et land som Norge i dag. Dette er en rein myte, som opprettholdes fordi det finnes noen prosjekter i noen land som faktisk har avgjørende betydning for veksten. Men de prosjektene vil vel så godt som alltid også ha betydelige direkte nyttevirkninger.

Den andre måten å overvurdere nytten på er å legge inn for optimistiske anslag for trafikkveksten (det gjøres sjelden) eller å oppjustere enhetsverdier og parametre uten noe godt vitenskapelig belegg. Det sistnevnte gjøres stadig oftere. Man hauser opp metodisk svake og dårlig dokumenterte undersøkelser som gir høye tidsverdier, mens godt dokumenterte undersøkelser med de beste metoder i verden blir ignorert. Man bruker pensjonerte bankøkonomer som sannhetsvitner på at kalkulasjonsrenta må settes ned, og viser til masteroppgaver for å begrunne kritikk mot eksisterende veiledere. På denne måten får man beslutningstakerne til å tru at tvilsomme infrastrukturprosjekter kommer til å kaste mer av seg enn beregningene viser.

# 6 Drøfting

Beslutningssituasjonen vi har sett på er nokså endimensjonal. Det dreier seg om å vedta eller forkaste et prosjekt som bare gir nytte av ett slags, nemlig en tidsbesparelse som er den samme for alle trafikanter. Prosjektet skal finansieres gjennom en avgift på alle reiser. Trafikantene er ikke interessert i annet enn tid og penger. Deres verdsetting av tid i penger, dvs. deres tidsverdi, er kjent, slik at alle trafikantene kan ordnes etter tidsverdien fra den minste til den største. Tidsverdien fordeler seg slik blant trafikantene at medianen er mindre enn gjennomsnittsverdien. Dette er den sammenhengen vi har sett på.

I denne sammenhengen kan vi vise at om et prosjekt blir vedtatt gjennom flertallsvedtak blant de berørte, så er det bevist at det er lønnsomt. Hvis det ikke er lønnsomt, veit vi da at det ikke kommer til å bli vedtatt. Men det vil også finnes *lønnsomme* prosjekter som ikke blir vedtatt. Dette følger i grunnen direkte av situasjonsbeskrivelsen, men er vist mer formelt i vedlegget.

Hvis vi ikke trur at spillsituasjoner og uheldige insentiver har påvirket avgjørelsen, kan vi altså bruke en demokratisk avgjørelse om å gjennomføre et prosjekt som prøve på at vår lønnsomhetsberegning er riktig hvis den har vist lønnsomhet, og feil hvis den har vist ulønnsomhet. Men vi kan ikke ta det faktum at et prosjekt er forkastet som bevis på at det ikke er lønnsomt.

Når det er flere elementer enn tid og penger som har betydning i en samfunnsøkonomisk analyse, kan vi ikke lenger trekke slike konklusjoner – i alle fall ikke før vi veit fordelingen til betalingsvilligheten for disse elementene i den befolkningen som påvirkes av prosjektet. Enda mer komplisert blir situasjonen om vi også vil ta hensyn til ikke-prissatte elementer.[51] Dersom

[51] Den "offisielle" måten å ta hensyn til ikke-prissatte konsekvenser på, er den som beskrives i kapittel 6 i SVV(2006). Den gir imidlertid ingen garanti for at de prosjektene som velges ut til å gjennomføres, er de beste for samfunnet. Det er ikke engang sikkert at det er mulig å snakke om hva som er best for samfunnet i denne sammenhengen, for folk kan være svært uenige og legge vekt på helt forskjellige ting. Det beste vi kan håpe på,

vi bruker et bedømmelsesopplegg som baserer seg på karakterer for hver type av virkning eller på at alle alternativer kan ordnes etter hvor godt de gjør det på hver type av virkning, vil imidlertid gjennomsnittet igjen spille en rolle for prioriteringene. Men denne gangen dreier det seg om gjennomsnittsscore over ulike virkninger, ikke gjennomsnittlig betalingsvilje for en enkelt virkning.

Så hvilket prinsipp bør vi bruke til å velge prosjekter? Prinsippet om flertallsavgjørelser gjennom avstemning blant de berørte stammer jo fra Athen i antikken og er et grunnleggende demokratisk ideal. Effektivitetsprinsippet, dvs. at summen av individenes betalingsvilje skal være avgjørende, virker kanskje mindre godt begrunnet. Betalingsviljen avhenger jo av hvor mye penger man har. Men den avhenger også av hvor interessert man er i å kunne oppnå nyttevirkningen som prosjektet gir. Hvis alle folk hadde hatt like mye penger, ville det kunne betraktes som et framskritt om man ikke bare tok hensyn til om folk syntes virkningen var verdt prisen, men også hvor mye mer verdt enn prisen den var for den enkelte. Det vil si at samfunnet tok hensyn til at folk har ulike interesser. Det motsatte kan sees som et flertallsdiktatur.

En praktisk situasjon vil alltid være mer komplisert enn vårt stiliserte tilfelle, men det er muligens et ideal å få beslutningssituasjonen så enkel og direkte at valget står mellom flertallsavgjørelse eller effektivitet. De taktiske hensynene virker det som en god ting om vi kunne bli kvitt, sjøl om de kanskje delvis kan spille en positiv rolle ved å motvirke tendensen til at flertallet ønsker seg færre prosjekter enn det som er samfunnsøkonomisk effektivt.

## VEDLEGG

Vår modell er en omformulering og utdyping av modellen i Varian (1992, avsnitt 23.6). Vi skal gi Varians modell en tolkning som gjelder tidsgevinster og tidsverdier i det enkle tilfellet der alle gevinstene tilfaller de som skal treffe avgjørelsen om prosjektene, samtidig som også alle kostnader bæres av de som skal treffe avgjørelsen om prosjektene. Tidsgevinster er eneste nyttevirkning, og medianverdien til tidsverdien er mindre enn gjennomsnittsverdien.

Vi skal vise at prosjektene som vil bli vedtatt ved flertallsavstemning er en ekte delmengde av de samfunnsøkonomisk lønnsomme prosjektene. Dermed er et positivt flertallsvedtak bevis på lønnsomhet, men det vil også finnes lønnsomme prosjekter som ikke blir vedtatt.

Vi skal først finne samfunnsøkonomisk optimalt omfang på gjennomføringen av tidsbesparende tiltak når individene har nyttefunksjoner med et entydig maksimumspunkt, såkalte "single-peaked preferences". Tidsbesparelser kan gjennomføres i fritt valgt utstrekning til en fast samfunnskostnad av $c$ kroner pr. minutt. Individene er indeksert med $i$. Individ $i$ har nyttefunksjonen

$$v_i = u_i(g) - s_i \cdot g$$

der $g$ er tidsgevinsten og $s_i g$ er individets bidrag til finansieringen. Vi maksimerer summen av individenes nytte under en bibetingelse om fullfinansiering. Vi antar at funksjonen $u(.)$ er tiltakende og konkav. For enkelhets skyld forutsetter vi også at $g > 0$ og at ikke alle $s_i$ er null. Maksimeringsproblemet kan skrives:

er at vi kjenner styrkene og svakhetene ved den metoden vi bruker, og at folk ikke er håpløst uenige om de virkningene som skal bedømmes. Beslutningsteori og avstemningsteori er store og levende fagfelter, se for eksempel "voting theory" på Wikipedia, vedlegg 2 i Jordanger m,fl. (2007) og Minken m.fl. (2009).

$$Max_{g,s}\, W = \sum_i \left(u_i\left(g\right) - s_i g\right) \text{ gitt } \sum_i s_i g \geq cg \quad \left(\lambda\right)$$

$\lambda$ er Lagrangeparameteren til bibetingelsen. Kuhn-Tuckerbetingelsene for maksimum kan skrives

$$\frac{\partial \mathrm{L}}{\partial g} = \sum_i \left(\frac{\partial u_i}{\partial g} - s_i\right) + \lambda\left(\sum_i s_i - c\right) = 0$$

$$\frac{\partial \mathrm{L}}{\partial s_i} = -g + \lambda g \leq 0 \ \left(= 0 \text{ for } s_i > 0\right) \quad \left(\forall i\right)$$

$$\lambda \geq 0 \quad \left(= 0 \text{ for } \sum_i s_i g > cg\right)$$

Siden det finns minst en strengt positiv $s_i$ og $g > 0$, må minst en av ulikhetene på midterste rad gi resultatet $\lambda = 1$. Dermed veit vi av den tredje raden at bibetingelsen er oppfylt med likhet. Med $\lambda = 1$ finner vi også av første rad at

$$\sum_i \frac{\partial u_i}{\partial g} = c$$

I optimum skal altså summen av alle individenes marginale nytte av en ytterligere tidsbesparelse være lik kostnaden ved å frambringe den. Vårt første resultat framkommer ved å dele på antall individer, *n*:

$$\text{(1)} \qquad \frac{1}{n}\sum_i \frac{\partial u_i}{\partial g} = \frac{c}{n}$$

På venstresida i (1) står den gjennomsnittlige betalingsvilligheten for en liten forbedring av reisetida, eller den gjennomsnittlige tidsverdien, om man vil. Den skal i optimum være lik den gjennomsnittlige kostnaden pr. person for å forbedre reisetida med én enhet.

Dette resultatet sier hvor mye reisetidsreduksjon som det er samfunnsøkonomisk effektivt å gjennomføre, gitt kostnadene ved å produsere det. Men det sier ikke noe bestemt om hvordan kostnaden skal fordeles. Om vi ser tilbake på første linje i Kuhn-Tuckerbetingelsene for optimum, ser vi at den siste parentesen, $\Sigma s_i - c$, vil være null, fordi bibetingelsen i optimeringsproblemet er oppfylt med likhet i optimum. En måte å oppnå effektivitet på er derfor å sørge for at hvert individ tilpasser seg slik at tidsverdien blir lik $s_i$, slik at alle ledd i den første summen blir null. Og dette vil de gjøre av seg sjøl om de tillates å maksimere individuell nytte med de kostnadsandelene de har blitt tildelt. Men også andre tilpasninger vil kunne gi optimum, bare summen over alle individer av tidsverdien minus kostnadsandelen blir lik null. Og en hvilken som helst fordeling av kostnadsandelene vil gjøre samme nytte. For eksempel kunne alle andeler være like uten at det endrer effektiviteten. Dette er fordi nyttefunksjonene har en form som gjør at inntektsfordelingen ikke spiller noen rolle for effektiviteten.

Dette resultatet må nå sammenliknes med omfanget av tidsbesparelser som kan vinne flertall i avstemning blant de berørte. Vi ordner alle individene etter tidsverdien. Siden tidsverdien i vår modell er en funksjon av nivået på tidsbesparelsen, må vi tenke oss at denne ordningen skjer med utgangspunkt i et bestemt nivå på *g*. La oss si dette er det samfunnsøkonomisk optimale nivået, og la oss si at alle må betale samme kostnadsandel *s*. *Fullfinansiering* med *n* individer krever $sng \geq cg$ eller med andre ord $s \geq c/n$. *Flertall* krever at median-individet (som har like mange med lavere betalingsvilje enn seg sjøl som med høyere betalingsvilje) stemmer for. Kall dette individet *m*. Vi må altså ha:

$$\frac{\partial u_m}{\partial g} \geq s \geq \frac{c}{n} \qquad (2)$$

Men dette er umulig, for median-tidsverdien er mindre enn gjennomsnittsverdien. Gjennomsnittsverdien er *c/n* ifølge (1), så (2) motsier (1). Vi klarer altså ikke å få flertall for et forslag som ville kunne finansiere det samfunnsøkonomisk effektive nivået på tidsbesparelser. Det meste vi kan få flertall for, er et lavere bidrag $s_m$ som er det meste mediantrafikanten kunne tenke seg å betale for tidsgevinsten, dvs.

$$\frac{\partial u_m}{\partial g} = s_m \qquad (3)$$

Men dette krever et lavere nivå på *g*. (Ved et lavere nivå på *g* vil den deriverte av $u_m(g)$ øke hvis u er konkav, og det nødvendige nivået på hvert individuelle bidrag vil også kunne reduseres, sjøl om det ikke skjer i vår forenklede kostnadsmodell der enhetskostnaden er konstant.)

Resonnementet er ikke helt presist, men vi har vi i det minste sannsynliggjort at den demokratiske modellen gir mindre tidsbesparelser enn det som er samfunnsøkonomisk effektivt, og at det derfor vil finnes prosjekter som er lønnsomme men ikke blir vedtatt. Alle de som blir vedtatt vil imidlertid være lønnsomme.

## Litteratur

Jordanger, I., S. Malerud, H. Minken og A. Strand (2007) Flermålsanalyser i store statlige investeringsprosjekt. Concept-rapport 18. http://www.concept.ntnu.no.

Minken, H., O.I. Larsen, J.H. Braute, S. Berntsen og T. Sunde (2009) Konseptvalgsutredninger og samfunnsøkonomiske analyser. TØI-rapport 1011/2009.

Ramjerdi, F., S. Flügel, H. Samstad og M. Killi (2010) Den norske verdsettingsstudien. Tid. TØI-rapport 1053B, TØI.

SVV (2006) Konsekvensanalyser. Vegledning. Håndbok 140.

Varian, H.R. (1992) Microeconomic Analysis. Third Edition. W.W. Norton & Company, New York.

## 3.4 Merknader om mernytte[52]

[52] Dette er arbeidsdokument ØL/2333/2011, tidligere publisert i TØI-rapport 1198/2012

# Merknader om mernytte

## Innhold

# 1 Samfunnsøkonomiske analyser i transportsektoren i dag

Samfunnsøkonomiske analyser av infrastrukturtiltak i samferdselssektoren konsentrerer seg i dag om de direkte effektene i transportmarkedene:

- trafikantenes og godseiernes tids- og pålitelighetsgevinster og monetære kostnader (kjørekostnader, billettkostnader, bompengekostnader),
- overskuddet til kollektivselskapene og de andre selskapene i sektoren (for eksempel bompengeselskaper, parkeringsselskaper, private selskaper som bygger eller driver infrastruktur),
- budsjettvirkningene for det offentlige (kostnader vedrørende bygging og drift av infrastruktur, overføringer til og fra private selskaper i sektoren, budsjettvirkninger av endringer i inngangen av skatter og avgifter fra transportsektoren),
- ulykkeskostnader, støykostnader, kostnader ved utslipp av klimagasser og lokal luftforurensning.

Etterspørselen etter transport antas som regel å være en funksjon av generaliserte transportkostnader, der ulike former for tidskostnader og monetære kostnader inngår. Trafikantenes nytte uttrykkes generelt som konsumentoverskuddet, målt enten med trapesformelen eller ved en logsum. Virkningene for det offentlige multipliseres med en skattefaktor (skyggepris på offentlige midler) som skal fange opp kostnadene i økonomien som helhet ved å finansiere offentlig virksomhet over skatteseddelen.

Ved til slutt å summere over de fire gruppene – trafikanter, selskaper, det offentlige og samfunnet for øvrig – elimineres overføringer som billetter, bompenger, skatter og avgifter. På grunn av trafikantenes og selskapenes tilpasninger og skattefaktoren for det offentlige vil det ikke være riktig å eliminere overføringer før man har tatt hensyn til hvordan de påvirker atferden til trafikantene og selskapene og budsjettbalansen til det offentlige. Denne måten å føre nyttekostnadsregnestykket på kaller vi i Norge for bruttometoden.

Fra gammelt av er konsumentoverskuddet ved tiltak i vegsektoren beregnet som differansen mellom samlede generaliserte kostnader før og etter tiltak. Dersom det kan antas at ingen aktører endrer sin tilpasning som følge av tiltaket, er denne metoden bare et spesialtilfelle av den generelle metoden, men i motsatt fall gir den direkte feilaktige resultater.[53]

# 2 Hva er mernytte?

En mulig definisjon av mernytte er: *Mernytte er nyttevirkninger som ikke er inkludert i standard nyttekostnadsanalyser innen samferdsel i Norge i dag.* Med denne definisjonen omfatter

[53] Siden vi i Norge ikke har noen instans som treffer offisielle vedtak om hvilke forutsetninger, modeller, framgangsmåter og parametre som skal brukes i samfunnsøkonomiske analyser i transportsektoren, finnes det mange konsulentselskaper på markedet som ikke følger bruttometoden eller som ikke skjønner begrensningene i vegvesenets tradisjonelle metode. Trolig finnes også avvik mellom beskrivelsen av bruttometoden i for eksempel Minken (2009) og måten den er programmert på i etatenes beregningsverktøy.

begrepet også nyttevirkninger som finnes i transportsystemet, men som man ikke har klart å lage noen gode metoder for å beregne ennå. Den viktigste av slike virkninger er utvilsomt økt pålitelighet (redusert reisetidsvariabilitet), der vi veit mye i teorien, men mangler data om virkningen av ulike tiltak på reisetidsvariabiliteten. Noen har også behandlet emner som kalkulasjonsrenta, hyperbolsk diskontering, mer realistiske levetider, tidsverdienes utvikling over tid m.m. under overskrifta mernytte. Men dette er misvisende, for det dreier seg ikke om noe vi ikke inkluderer i dag, men mer om å skaffe riktigere anslag på parametre som brukes i dag eller effekter som er inkludert i dag.

Det finns altså det vi kunne kalle transportsektorintern mernytte. Dette notatet skal ikke handle om det, sjøl om det å få inkludert slike nyttekomponenter sannsynligvis betyr mer for å forbedre analysene enn noe annet. Det vi drøfter her, er nyttevirkninger av transportforbedringer som oppstår eller materialiserer seg i andre markeder enn transportmarkedene.

Det er allment kjent at dersom det hersker perfekt konkurranse i alle markeder i økonomien, med priser lik den samfunnsøkonomiske grensekostnaden, vil virkningene av et transporttiltak i andre deler av økonomien ikke være annet enn forvandlede former for den nytten for trafikantene som vi kan beregne i transportmarkedene. Vi kaller dette for reint pekuniære eksterne virkninger. At en ekstern virkning er pekuniær, betyr at den ikke endrer de økonomiske aktørenes nyttefunksjoner eller produktfunksjoner, men bare endrer de prisene de står overfor i markedet. De kan endre sitt forbruk eller sin produksjon som følge av slike prisendringer, men det er omfordelingseffekter og bidrar ikke med ny nytte ut over den som kan beregnes i det direkte berørte markedet.

Hvis det imidlertid finns imperfeksjoner i økonomien – og det gjør det jo – vil pekuniære eksterne effekter kunne ha reelle samfunnsøkonomiske virkninger.[54] Det kan tenkes at priser, markedsmakt, informasjonsforhold eller andre forhold er slik at det produseres mer eller mindre enn det burde i de bransjene som bruker transport, eller at husholdningene som bruker transport, konsumerer mer eller mindre enn de burde av visse varer og tjenester. Da kan transporttiltak bidra til helt eller delvis å motvirke underforbruket eller motvirke overforbruket, slik at den samfunnsøkonomiske effektiviteten i økonomien øker.

Dette er mernytte. Den kan være positiv eller negativ. I ekstreme tilfeller kan en transportforbedring frigjøre helt nye ressurser som ikke var tilgjengelige før, eller åpne opp for produksjonsmuligheter og konsummuligheter som ikke fantes før. Men for at vi skal plusse på våre regnestykker med mernytte av det ene eller andre slaget, holder det ikke bare å vise at økonomien ikke er perfekt. Et rimelig minstekrav må være at vi identifiserer mest mulig nøyaktig hvilke misforhold i økonomien som en konkret transportforbedring vil kunne motvirke eller korrigere, slik at vi kan skaffe oss et solid grunnlag for å estimere hvordan tiltaket vil virke og produsere mernytte.

Stikkordmessig: Det holder ikke å si at denne vegen vil øke eksporten. Du må også si hvorfor det ligger en samfunnsøkonomisk gevinst i å øke eksporten. Det holder ikke å si at denne jernbanen vil gi et nytt bosettingsmønster og pendlingsmønster. Du må også fortelle oss hva som er så gjevt med det.

[54] Den klassiske referansen som viser at det finnes mernytte av transportinvesteringer i imperfekte økonomier, men at det ikke finnes slik mernytte i perfekte økonomier, er Jara-Diaz (1986).

## 2.1 Former for mernytte

Her er en ufullstendig liste over forhold som kan gi opphav til mernytte:

1. Kortere ledetider og lavere sikkerhetslagre på grunn av raskere og mer pålitelig transport.

2. Stordriftsfordeler i produksjonen som kan utnyttes bedre på grunn av at et raskere, mer pålitelig eller mer høyfrekvent transporttilbud gir tilgang til et større marked og dermed færre, men større produksjonsanlegg.

3. Et mer effektivt arbeidsmarked på grunn av at lavere transportkostnader gjør det lønnsomt for flere å jobbe (utenfor hjemmet). Dette gir blant annet arbeidsgiverne flere jobbsøkere å velge blant og arbeiderne flere jobber å velge blant, og gir grunnlag for økt spesialisering av arbeidsstyrken, hvilket alt sammen gir økt produktivitet.

4. Lavere avgifter på arbeidsreiser eller kortere reisetid til jobb er en nestbesteløsning som kan redusere effektivitetstapet i arbeidsmarkedet på grunn av inntektsskatten.

5. Bedre transport gjør at butikker, tjenesteytende bedrifter og kulturtilbud har tilgang til et større kundegrunnlag, og dermed kan tilby et større og mer variert utvalg. Det kan antas at kundene på mange av disse områdene har nytte at større variasjonsbredde i tilbudet. Det er enten fordi de liker variasjon i seg sjøl, eller fordi et stort marked betyr at den enkelte butikk eller det enkelte tilbudet kan spesialisere seg. Kundene har ofte mer nytte av hvert sitt spesialiserte tilbud enn av et felles gjennomsnittstilbud.

6. Bedre transport gir bedrifter som er lokalisert på ulike steder, mulighet til å trenge inn i hverandres markeder. Slik kan det føre til at bedriftene mister markedsmakt og at prisene nærmer seg frikonkurransepriser, samtidig som produsert volum øker. Eller det kan føre til at noen blir utkonkurrert, hvilket på den ene sida kan gi økt effektivitet i produksjonen, men på den andre sida mer markedsmakt til de gjenværende bedriftene, høyere priser og lavere volumer.

7. Samlokalisering av bedrifter i samme bransje eller langs samme verdikjede kan gi gevinster i form av utveksling av ansatte, kunnskapsutveksling og gjensidig læring mellom de ansatte i bedriftene, et mer levedyktig og kompetent miljø, felles informasjonssystemer og annen infrastruktur, billigere innkjøp osv. osv. Dette kalles bransjevise agglomerasjonsfordeler. Det finns også byomfattende agglomerasjonsfordeler, som er gevinster ved å ha en stor eller tett befolket by, på tvers av bransjer. I den grad dette har med hyppigere og lettere kontakt mellom mennesker å gjøre, vil slik agglomerasjonsfordeler kunne styrkes ved bedre transport.
8. Endringer i transportsystemet fører til endringer i arealbruk og lokalisering.

"Imperfeksjonene" som fører til at pekuniære eksternaliteter har (positive) reelle virkninger. De sammenfattes ofte i følgende kategorier:[55]

[55] Small (1997) har en lengre liste: Hans sier pekuniære eksterne virkninger har virkelige økonomiske effekter om det finns imperfekt konkurranse, stordriftsfordeler, teknologiske eksterne virkninger, agglomerasjonsfordeler, skattevridninger, økende utbytte av tetthet eller økende utbytte av variasjon og mangfold.

- Stordriftsfordeler
- Agglomerasjonsfordeler
- Oppbryting av markedsmakt
- Mer effektivt arbeidsmarked

Agglomerasjonsfordeler er vel egentlig alt mulig som tilsier at produksjonen fungerer bedre eller at forbruket blir mer tilfredsstillende dersom flest mulig bedrifter eller folk samles tett sammen på et relativt lite område. Det dreier seg om slike ting som samarbeid og kontakt, gjensidig læring, spesialisering og større mangfold. Nytte av større mangfold og variasjon settes av og til opp som et eget punkt, men betingelsen for å kunne oppnå slik nytte er geografisk konsentrasjon av tilbudet.

Imidlertid finns det ikke bare mernytte på grunn av pekuniære eksternaliteter i imperfekte økonomier. Det finns også annen utelatt nytte som oppstår i markeder som står i nær forbindelse med transport.[56] Grunnen til det er rett og slett at vi har transportmodeller der ikke alle trafikantenes tilpasninger til endringer i systemet er med.[57] Hvis for eksempel vi hadde en del av transportmodellen som kunne si oss hvordan flere vil begynne å jobbe på grunn av at reisa til jobb blir enklere, så ville vi kunne måle nytten av det i transportmarkedet, med mindre det at flere begynte å jobbe hadde effekter for hvor produktive de som jobber er. Vi kan våge påstanden at hvis vi hadde perfekte modeller av transportmarkedene, ville ”markeds-imperfeksjoner” som stordriftsfordeler, agglomerasjonsfordeler, monopolmakt og preferanser for diversitet være eneste grunnlag for mernytte.

Det leder oss allerede til en viktig konklusjon: La oss være skeptiske mot å inkludere alle slags former for mernytte i transport, basert på dårlige anslag om virkninger som vi antar kan finnes, men egentlig veit svært lite om. La oss i stedet ta opp to andre former for arbeid. *For det første* kan vi identifisere de formene for ”imperfeksjoner” som vi mener er de viktigste i hvert tilfelle, og stille spørsmålet om hva som kan gjøres for å redusere dem. Det er slett ikke sikkert at det enkleste er transporttiltak – kanskje finns det enkle former for regulering av markedet. Når dette faktisk er gjort i praksis, kan vi evaluere transporttiltaket uten mernytte. *For det andre* kan vi systematisk forbedre vårt modellapparat, blant annet ved å prøve med modeller som kombinerer transport og arealbruk og modeller som kombinerer transport og arbeidsmarked. Vi bør også føre dette videre til geografisk oppdelte beregnbare generelle likevektsmodeller (SCGE-modeller). Det som er et problem med slike modeller når det gjelder evaluering av konkrete transporttiltak, er to ting: De inneholder for mange økonomiske sammenhenger av relativt liten betydning for den konkrete nytteberegningen, og dette er en uoversiktlig feilkilde. På den andre sida inneholder de for lite detaljer om transportsystemet. Det kan gjøre at tiltaket som skal

---

[56] Mernytte er nytte som ikke er inkludert i nyttekostnadsanalyser i transportsektoren med dagens beregningsverktøy, og som forårsakes av brudd på beregningsverktøyets forutsetning om fullkommen konkurranse og fravær av eksternaliteter i berørte markeder, samt av dets mangelfulle modellering av bedriftenes og husholdningenes tilpasninger til transportforbedringer på lang sikt.

(Hanne Samstad, Cowi, i en powerpoint fra 2010 funnet på nett)

[57] Punkt 1 og 8 på lista er vel eksempler på det. Men punkt 1 er faktisk dekket i Logistikkmodellen, den nye versjonen av den nasjonale godsmodellen. Punkt 4 kan studeres i generelle likevektsmodeller. Punkt 3 krever derimot en type modeller av arbeidsmarkedet som vi ikke har.

evalueres, ikke lar seg modellere i det hele tatt – det går så å si under radaren. Det er derfor viktig at SCGE-modellen utstyres med et realistisk og mest mulig korrekt transportnettverk.[58]

# 3 Mernyttediskusjonen i transportsektoren

## 3.1 Transport og eiendomspriser

Endringer i transportsystemet kan utløse endringer andre steder i økonomien. Fra 60-tallet av var man opptatt at hvordan transporttiltak kunne påvirke arealbruk, tomtepriser og eiendomspriser. Forskningsfeltet som i første rekke kunne se disse tingene i sammenheng, var økonomisk byteori (urban economics). Feltet fikk en teoretisk basis da Alonso (1964) tok von Thünens gamle teori om den isolerte staten (von Thünen 1826) og anvendte den på lokaliseringen av boliger og (etter hvert) næringsvirksomhet i byer.[59] En mye brukt lærebok er Fujita (1990).

I utgangspunktet studerte man monosentriske byer der alle arbeidsplasser er lokalisert i sentrum. Beboerne velger lokalisering og generelt forbruk, og resultatet er en likevekt der husleia synker og reisevegen øker med avstanden til sentrum, på en sånn måte at om alle beboere er like, vil de i likevekt også få samme nytte. Likevekta vil være samfunnsøkonomisk optimal, og dersom inn- og utflytting av byen var mulig i modellen, ville byen vokse til et optimalt folketall og en optimal geografisk størrelse.

Et resultat i denne litteraturen er at når transportforbedringene har ført til en ny likevekt i arealbruken og eiendomsmarkedene, vil det være de som eier land som sitter igjen med hele gevinsten (Mohring 1961, Wheaton 1977, Mohring 1993). Å måle gevinsten direkte i transportmarkedene og å måle den i form av prisøkninger i eiendomsmarkedet er derfor to alternative måter å beregne den samme nytten på, mens det å legge gevinster i eiendomsmarkedet til nytten som er beregnet i transporten, vil være dobbelttelling. Dette resultatet har utvilsomt lenge (kanskje litt for lenge) styrket transportøkonomene i trua på at en partiell likevektsmodell for transportsektoren er et tilstrekkelig verktøy for samfunnsøkonomiske analyser av transporttiltak.

Teorien blei raskt utvidet slik at ikke bare beboerne, men også bedriftene kunne velge hvor de ville slå seg ned. Eller med andre ord: Det gjaldt om å forklare hvorfor arbeidsplassene samlet seg i et sentrum, eller alternativt å vise at det var mulig med andre likevekter, med flere sentrale strøk eller et hierarki av næringsområder. Det viste seg at det da slett ikke var sikkert at det fantes noen likevekt, og om den fantes, behøvde den ikke være entydig.[60] Fujita (1986) kunne

[58] Tidligere TØI-medarbeider Olga Ivanova var trolig blant de aller første som utstyrte en SCGE-modell med et realistisk transportnettverk (Ivanova 2003). Hun bistår nå med råd ved oppbyggingen på TØI av en ny norsk SCGE-modell.

[59] von Thünen er i dag bare kjent som lokaliseringsteoretiker. Men i andre bind av boka om den isolerte staten, som kom ut nesten 30 år seinere, tok han opp et problem som var stilt av Hegel i rettsfilosofien, nemlig hvilken rolle kolonisering og imperialisme spiller i å stabilisere kapitalismen. Fra von Thünen går det derfor linjer videre til så ulike tenkere som Alfred Marshall og Karl Marx (Harvey 2001).

[60] Starrett (1978) viser at ved frikonkurranse og positive transportkostnader finns det ingen likevekts lokalisering. Bedrifter og kunder vil alltid ha et ønske om å flytte nærmere hverandre for å unngå transportkostnaden, slik at den eneste mulige likevekta blir at de holder til på samme sted. Men det er umulig hvis de i det hele tatt krever noe plass.

i følge Mori (2006) vise at betingelsen for at det oppsto noe sentrum i det hele tatt, var at minst ett av følgende vilkår var tilstede:

- Stedet har en fordel framfor andre steder i form av naturressurser, transportmuligheter eller allerede bygd infrastruktur,
- Det eksisterer eksterne virkninger i produksjon eller forbruk
- Det eksisterer markeder hvor det finnes en form for markedsmakt. Alle markeder kan altså ikke være preget av frikonkurranse.

Utviklingen etter dette er preget av modeller som forklarer framveksten av et bysentrum ved å innføre en eller flere av disse vilkårene i modeller der bedriftene kan velge lokalisering.[61,62] Et gjennomgående funn er at under slike vilkår gir markedslikevekten en by som har for stor utstrekning, og altså ikke er optimal.

Agglomerasjonsfordeler og stordriftsfordeler er ikke eksplisitt nevnt som egne vilkår for dannelsen av bysentra, men vi får anta at de er inkludert under andre og tredje kulepunkt. I alle fall er det jo slik at om det ikke fantes stordriftsfordeler eller agglomerasjonsfordeler, ville det mest effektive være at hver forbruker sjøl produserte alle ting i sin egen bakgård, slik at han slapp transportkostnaden. En annen årsak til bydannelse som nevnes i Fujita (1990) er ønsket om variasjon i forbruk eller produksjon. Å ha kontakt med mange ulike personer, å oppsøke mange ulike aktiviteter eller å kunne kjøpe mange ulike produktvarianter – kort sagt å ha mye å velge i – er intuitivt grunner for å bo i by, og dette kan også formaliseres, for eksempel slik som hos Borukhov og Hochman (se fotnote 8).

Mernytte er nyttevirkninger som ikke fanges opp i en nyttekostnadsanalyse i transportsektoren aleine. Vi så at ved et predeterminert senter og ingen imperfeksjoner for øvrig, er markedsløsningen optimal, og det eksisterer ingen mernytte av transportforbedringer. Vi så også i forrige avsnitt at om noen av de forholdene som vi nå har sett kan gi opphav til bydannelse, er tilstede, vil pekuniære eksternaliteter ha reelle økonomiske virkninger. Dermed bør det være normalt at det oppstår mernytte ved transportforbedringer i byområder. Fagfeltet urban economics ser ut til å bekrefte at så er tilfelle.

## 3.2 Omorganiseringsnytte i næringslivet (industrial reorganisation benefits)

Vi har snakket om at trafikantnytten (konsumentoverskuddet) kan ende opp som grunnrente til den produksjonsfaktoren som ikke er mobil, men foreligger i en gitt begrenset mengde på hvert sted, nemlig land. Et annet forhold som fører til at aktørene i økonomien tilpasser seg til forbedringer i transportsystemet på en slik måte at vi også får endringer utenfor transportsektoren, er stordriftsfordeler i produksjonen. Hvis det ikke fantes stordriftsfordeler i produksjonen, ville all transport være bortkastet, siden all produksjon ville kunne foregå i liten skala i umiddelbar nærhet til forbrukeren. Det er stordriftsfordeler i produksjonen som (sammen med

[61] Et spesialtilfelle er Borukhov og Hochman (1977) der det ikke finnes bedrifter, men alle innbyggere har et behov for å reise til alle andre innbyggere. Av dette oppstår det et sentrumsområde. Når folk tar kontakt med hverandre i den modellen, skapes det positive eksterne virkninger som de ikke sjøl tar hensyn til. Derfor blir byen mer utstrakt enn det som er optimalt.

[62] Se Mori (2006) for en oversikt.

behovet for variasjon i forbruket og behovet for menneskelig kontakt) skaper transportbehovet. Om transporten da blir raskere eller billigere eller mer pålitelig, vil det bli økonomisk lønnsomt å utnytte stordriftsfordelene enda mer. Et sannsynlig resultat av bedre transport er derfor større produksjonsanlegg som ligger lengre fra hverandre. Dette ble formalisert i modellen til Mohring og Williamson (1969).

Mohring og Williamson så for seg en bransje med produksjonssteder spredt helt jamt utover i planet. Befolkningstettheten, og dermed etterspørselen, var den samme overalt i planet. Alle produksjonssteder var eid av én og samme eier, og dette firmaet dreiv også transporten og bar kostnaden ved det. Alle kunder handlet på det nærmeste produksjonsstedet. Dermed oppstår sekskantede markedsområder rundt hvert produksjonssted. Ved en transportforbedring vil det oppstå en ny likevekt der produksjonsstedene blir færre og markedsområdene rundt dem vokser. Mohring og Williamson viste så at den samfunnsøkonomiske gevinsten av disse endringene kunne måles ved konsumentoverskuddet i transportmarkedet. Heller ikke i dette tilfellet fantes det altså noen mernytte.

Den meget anerkjente transportøkonomen Kenneth Small hevder at i motsetning til økt økonomisk aktivitet og økte eiendomspriser i nærheten av nybygd transportinfrastruktur, som stort sett enten er aktivitet overført fra andre steder eller trafikantnytten i en annen form, så er omorganiseringsnytte i form av bedre utnyttelse av stordriftsfordeler og samdriftsfordeler en viktig og muligens ganske stor effekt av bedre transport (Small 1997). Han godtar likevel Mohrings og Williamsons påvisning av at bedre utnytting av stordriftsfordelene fanges opp av en nyttekostnadsanalyse i transportsektoren. Om ikke alle effektene skulle tilfalle ett enkelt monopol, som i Mohring og Williamson, vil likevel mesteparten bestå og kunne beregnes i transportsektoren, later han til å mene. Han viser til Jara-Diaz (1986) for å sannsynliggjøre at heller ikke om Mohring og Williamson hadde forutsatt frikonkurranse, ville det ha eksistert noen mernytte som ikke var fanget opp i en partiell likevektsanalyse i transport. Dermed er det bare agglomerasjonsfordeler og reduksjon av monopolmakt som kan gi opphav til betydelig mernytte, synes han å mene.

I et upublisert arbeid som nå snart er 10 år gammelt, fant svenskene Jan-Owen Jansson og Rikard Wall på å endre på forutsetningene i Mohring og Williamsons modell (Jansson og Wall 2002, se også Wall 2002). I stedet for et monopol som eide alle bedriftene satte de frikonkurranse mellom dem, og de lot samtidig kundene ta på seg å hente varene på produksjonsstedet sjøl. Nå var det straks en meget betydelig mernytte ut over den som kunne beregnes i transportsektoren. I en kommentar til dette arbeidet (Minken 2011) pekte jeg på at dette hadde en enkel forklaring: Monopolet til Mohring og Williamson bar faktisk alle kostnader i samfunnet, og ville derfor sørge for kostnadseffektivitet i samfunnsøkonomisk forstand. Siden monopolet også tilranet seg hele konsumentoverskuddet og alle kunder ville bli betjent både før og etter tiltaket, er monopolløsningen ikke bare kostnadseffektiv, men også samfunnsøkonomisk optimal. I frikonkurranseløsningen, derimot, påfører bedriftene hverandre en negativ ekstern virkning gjennom at de innskrenker hverandres markedsområde.[63] Hver bedrift blir for liten, og gjennomsnittlige produksjonskostnader blir for store. Det finns derfor en mernytte i det tilfellet. Det gjaldt både om kundene bar transportkostnadene og om de ikke gjorde det, men mernytten var størst i det sistnevnte tilfellet.

Slik denne modellen var formulert, var pris slett ikke lik samfunnsøkonomisk grensekostnad, verken i frikonkurransetilfellet eller monopoltilfellet. Vi ledes til å anta at det som avgjør om

[63] Dette er en ekstern trengselskostnad. Rommet som bedriftene lokaliserer seg i, er en begrenset ressurs som blir overutnyttet ved fri etablering.

mernytte finnes eller ikke, dypest sett ikke er at prisene er lik grensekostnad, men om løsningen er samfunnsøkonomisk optimal eller ikke. Størrelsen på mernytten vil avhenge av graden av stordriftsfordeler, konkurranseformen og hvem som bærer transportkostnadene.

I modellene som studerer nytten av omorganisering i næringslivet, finnes det ingen eiendomsbesittere og ingen pris på arealer, slik som i Mohring (1961, 1993) og Wheaton (1977). Men i Mohring og Williamson (1969) finns det en annen aktør som ender opp med hele samfunnets overskudd, nemlig monopolbedriften. Det er interessant at det nettopp er i disse tilfellene at hele samfunnets nytte kan måles i transportsektoren, og at å legge til gevinstene til aktører i andre sektorer blir dobbeltregning.

## 3.3 Vilkår for mernytte under gitt arealbruk

I tilfellet vi nettopp har sett på, med omorganiseringsnytte i næringslivet, er forbrukerne stedfaste, men bedriftene kan flytte på seg uten noen som helst flyttekostnader. Vi skal snart se nærmere på hva som skjer om vi snur på forutsetningene om muligheten for å flytte på seg, dvs. at individene kan flytte, men bedriftene ligger fast. Men først må vi se på muligheten for mernytte om verken bedrifter eller folk kan endre lokalisering.

Det mest generelle resultatet i dette tilfellet er gitt av Jara-Diaz (1986). Det sier at hvis pris er lik samfunnsøkonomisk grensekostnad overalt i økonomien, vil det ikke finnes mernytte. Men hvis pris ikke er lik grensekostnad, kan det finnes effekter i økonomien som ikke fanges opp i en nyttekostnadsanalyse i transportsektoren. Virkningene kan være positive eller negative.

Sactra (1999) drøfter hvor store disse virkningene kan være og hvilken veg de vil kunne gå, og kommer fram til enkle tommelfingerregler som er samlet i en tabell som skiller mellom ni muligheter. Hver av de ni mulighetene er en kombinasjon av overprising eller underprising eller riktig pris i transportmarkedet med overprising eller underprising eller riktig pris i markedene som avtar transporttjenestene. På grunnlag av forsøk i en enkel generell likevektsmodell (Venables og Gasiorek 1999) har man videre laget en enkelt formel for hvor store effektene kan bli. Sactra regner ikke med at effektene kan utgjøre mer enn noen få titalls prosent påslag på nytten som er beregnet i transportmarkedet. Laird et al (2005) rapporterer for øvrig at annen forskning har nedjustert denne virkningen med en faktor på 10.

En oppdatert gjennomgang av de fleste former for mernytte og utfordringene med å beregne den er Lakshmanan (2010).

Hva enten det er feilprising i seg sjøl eller det at økonomien ikke er samfunnsøkonomisk optimal som gir opphav til mernytte, så er det i alle fall tydelig at eksistensen av mernytte bør begrunnes i hvert enkelt tilfelle ved å vise til hvilke mekanismer som har gjort at de mest relevante markedene ikke fungerer optimalt. Dermed oppstår også spørsmålet om vi ikke kan gjøre noe med disse markedene *før* vi eventuelt sier ja til dyre investeringer i transportinfrastruktur. Hvis det er realistisk, bør man så nytteberegne transporttiltaket uten mernytte, og først om det *ikke* anses realistisk bør man kunne regne med en mernytte av investeringen. Det blir i tilfelle å anse som en nestbesteløsning, og den kan derfor ikke bli bedre enn førstbesteløsningen, som er å rette opp prisene eller på andre måter bevege økonomien mot et optimum, og legge til investeringen om den skulle gi noe i tillegg til det.[64]

[64] I NOU 1997:27 gikk man inn for å nytteberegne trafikkinfrastruktur under en forutsetning om at prisene var riktige. For eksempel skulle man anta at det var innført køprising når man beregnet vegbygging i bystrøk. Dette

## 3.4 Mernytte når folk, men ikke bedrifter, kan flytte

I økonomisk byteori gjøres det et viktig skille mellom åpne og lukkede byer. En lukket by har en gitt befolkning, som imidlertid kan spres seg på et større eller mindre areal, avhengig av blant annet transportkostnadene. Normalt er det også antatt at alle arbeidsplasser (bedrifter) befinner seg i byens sentrum og ikke kan omlokaliseres. En åpen by beholder forutsetningen om at arbeidsplassene befinner seg i sentrum, men antar et uendelig stort omland av folk som kan flytte inn i byen dersom det gir større nytte enn å leve på landet.

Mens altså modellen til Mohring og Williamson ga muligheter for *bedriftene* til å endre geografisk plassering når transporten blei billigere, og dermed ta ut stordriftsfordeler som det ikke lønte seg å utnytte tidligere, vil en ”åpen by”-modell gi muligheter til teoretiske studier av agglomerasjonsfordeler ved at *folk* flytter på seg, slik at vi får en større bybefolkning. En større befolkning gir fordeler ved at bedriftene får mulighet til mer spesialisering av arbeidskrafta, næringslivet får muligheten til å opprettholde en mer variert produksjon eller et mer variert tjenestetilbud, vi får breiere kontakt mellom folk og dermed mer erfaringsutveksling og læring, m.m. Ingen av disse mekanismene er nærmere spesifisert i en enkel økonomisk bymodell, men om man kjenner mekanismen og hvordan den virker, kan man likevel modellere hvordan bedre transport vil øke byens befolkning og dermed produktiviteten eller overskuddet som blir produsert.

Venables (2007) antar en slik åpen by der folk har boliger som opptar et fast areal uansett hvor i byen de ligger. Alle arealer i byen utenom bysentrum er i bruk til boliger. Om folketallet øker, vil derfor byen dekke et større areal. I utgangspunktet er byen i likevekt, den har utnyttet de mulighetene den har under de gitte forholdene, og det har gitt bebyggelsen og folketallet en viss størrelse. Så får vi en transportforbedring. Det som skjer er helt i tråd med læreboka i økonomisk byteori – byen vokser. Bedre transport leder til innflytting, fordi det blir mulig å ta seg jobb i bysentrum og bo utafor det som var yttergrensa for byen før, og likevel oppnå en nytte som er større enn de som bor på landet.

Den viktige nye forutsetningen hos Venables er imidlertid at produksjonsnivået i byen er en tiltakende funksjon $f(N)$ av folketallet $N$. Innflyttingen øker $N$ og leder altså til økt produktivitet. Den økte produktiviteten fanges ikke opp i en nyttekostnadsanalyse i transportsektoren, og er en mernytte av potensielt betydelig størrelse. Inntekstskatt som skaper effektivitetstap i arbeidsmarkedet, fører til at mernytten blir større.

Det empiriske grunnlaget for å lage en modell som virker på denne måten, er observasjoner som tilsier at verdiskapningen i byer, målt ved lønnsnivået, er større jo større byen er. Det innrømmes at dette kan ha ulike andre årsaker som man prøver å korrigere for.

Det teoretiske grunnlaget er i utgangspunktet relativt generelt, med en by inndelt i soner som både kan inneholde boliger og arbeidsplasser, og med en produktfunksjon for hver sone $s$ der produksjonsnivået avhenger av en vektet sum av innbyggerne i sonene, med mindre vekt til soner som ligger lenger vekk fra $s$. Men på vegen fram mot formler for størrelsen av mernytten gjøres det spesielle forutsetninger som nok gjør regningen enklere, men byen mindre lik virkelige byer og produktfunksjonen mer urealistisk. Byen får alle arbeidsplassene i sentrum, og produksjonen blir en funksjon $f(N)$ av det totale innbyggertallet $N$, som nevnt. Dermed blir avstanden mellom folk uten betydning, bare de befinner seg innafor bygrensa.

var på et tidspunkt hvor det var helt klart at køprising ikke var politisk aktuelt. Utvalgets synspunkt på dette blei da heller ikke videreført i seinere veiledere.

Men Venables går videre fra formlene i den enkle matematiske modellen til en beregnbar modell. I den er all produksjon samlet i byens sentrum, men produksjonen tar nå plass, slik at sentrum blir større jo mer produksjon som skal foregå der. Det som *nå* gir økt produktivitet, er en vektet sum av de ansatte i sentrumssonene, der bedrifter teller mindre jo lenge ut de ligger. I denne siste modellen er det altså arbeid i selve sentrum som skaper mernytte, og mernytten som den enkelte ansatte skaper, avtar med avstanden mellom arbeidsplassen hans og rådhuset, for å si det slik. I ingen av modellvariantene spiller kø noen rolle.

Venables' arbeid kan illustrere noen viktige sammenhenger på en skissemessig måte, men jeg syns ikke det kan brukes til beregninger av mernytte i et aktuelt tillfelle. Til det vil virkelige byer ha en beliggenhet av arbeidsplassene som avviker for mye fra den monosentriske modellen, og arbeidsplassene i sentrum vil ofte bestå av offentlig administrasjon og andre aktiviteter som har mindre betydning for verdiskapningen.

Spesielt kan det bli misvisende resultater om en ikke skiller mellom bransjespesifikke og byomfattende agglomerasjonseffekter (urban effects og industrial effects). De bransjespesifikke agglomerasjonseffektene regnes for å være langt de viktigste, og det ville da kreve en beregning bransje for bransje, hver med sitt "sentrum". I så fall blir det en oppgave å finne ut hvordan innflyttingen påvirker antall arbeidsplasser i hver bransje. Dessuten burde en ta med produktivitetstapet på de stedene de nye byboerne flytter fra.

## 3.5 Nettverkseffekter

Det vi her kaller mernytte, kalles noen ganger nettverkseffekter. Nettverk og nettverksteori er begreper som har en presis betydning på noen områder innen matematikk og økonomi, men her dreier det seg om en vagere forestilling om at de ulike deler av transportsektoren henger sammen, slik at en positiv utvikling i en del av systemet kan sette i gang positive prosesser i andre deler. Spesielt tenker man seg at hvis transportnettverket *ikke* henger sammen, kan byggingen av den manglende lenka utløse et skred av nye forbindelser mellom økonomiske aktører. Denne tanken ligger bak EUs satsing på det såkalte TEN-T nettverket. Med nettverkseffekter tenker man også på at transport og markedene som bruker transport henger sammen, slik at en transportforbedring vil stimulere vekst og fornyelse i næringslivet.

Laird et al (2005) går gjennom den akademiske delen av denne litteraturen. De presiserer nettverkseffekter slik at det blir identisk med det vi har kalt mernytte, og deler den i effekter internt i transportsektoren og effekter fra transport til andre sektorer. Når det gjelder de transportsektorinterne effektene, så viser de at disse effektene i de aller fleste tilfellene fanges opp i en god transportmodell med etterfølgende nyttekostnadsanalyse i transportsektoren. Problemet er altså å gjøre modellene bedre, ikke å foreta skjønnsmessige korreksjoner i nytteberegningen på grunn av at det finns priser som ikke tilsvarer samfunnsøkonomisk kostnad internt i transportsektoren.

Kidokoro (2004) forsterker og utvider denne konklusjonen, for han viser at med en høvelig måte å beregne trafikantnytte og operatørselskapets nytte på, vil resultatet bli riktig sjøl i en situasjon der prisene i transport er feil. Det er altså ikke behov for å snakke om transportsektorintern mernytte i det hele tatt. (Derimot er det naturligvis en samfunnsøkonomisk gevinst i å rette opp prisene, hvis mulig.) Feilaktige priser i transport kan samvirke med feilaktige priser utenfor transport når det gjelder å skape mernytte utenfor transport, men de krever ikke noen triks eller korreksjoner i den nytten vi har beregnet i transport. Det eneste vi behøver å bry oss om, er at transportmodellene fanger opp tilpasningene i transportsektoren så godt som mulig, og at vi tar hensyn til de viktigste formene for mernytte i økonomien *utenfor* transportsektoren.

Kidokoros resultat forutsetter etterspørselsfunksjoner som ikke har inntekt som argument. Under denne forutsetningen er resultatet så vidt jeg skjønner allerede bevist i Williams (1977). Våre transportmodeller er slik at inntekt grovt sett ikke spiller noen rolle i etterspørsels-funksjonene. Vårt opplegg for nyttekostnadsanalyser i samferdsel i Norge er helt konsistent med Kidokoros (og Williams') resultat. Så transportintern mernytte burde heller ikke være noe diskusjonstema i Norge. Sjøl med etterspørsel som er funksjon av inntekt vil vi kunne beregne nytten på en riktig måte ved hjelp av teori som er utviklet for rundt 10 år siden.

## 3.6 Samspill mellom arbeidsmarkedseffekter og lokaliseringseffekter

I teori av samme type som i Venables (2007) blir byen mer produktiv jo flere arbeidere som jobber i sentrum. Produktiviteten slår ut i høyere lønn. I eldre amerikanske studier brukte man gjennomsnittstall for utvalgte yrkesgrupper til å teste hypoteser om lønnsforskjeller som skyldes om arbeidsplassen lå i sentrum eller utenfor. Timothy og Wheaton (2001) mener at det kan lede galt av sted om det også finns andre forskjeller mellom de som jobber i sentrum og de som jobber i utkanten – hvilket det åpenbart gjør. Seinere studier har derfor brukt mikrodata om den enkelte arbeidstaker og hennes egenskaper. Det blir da mulig å korrigere for erfaring, utdanning, fag og bransje. Det blir også mulig å teste hypoteser om hva slags effekter som faktisk frambringer lønnsforskjellene, og hvor store de er i ulike byer. Jeg antar at undersøkelser som bruker de mest moderne metodene fra stokastisk nytteteori, slik som i den norske verdsettingsundersøkelsen, vil være de som egner seg best.

Modellen i Timothy og Wheaton leder fram til følgende konklusjon om sammenhengen mellom arbeidsreiser, arbeidsstedets beliggenhet og lønn: Når bedriftene ligger spredt utover byen, ikke bare sentrum, vil transportkostnadsgevinster ikke bare ende opp som endringer i boligpriser/tomtepriser, men også som endringer i lønn. Mer nøyaktig: Ulikhet i kostnaden ved å pendle mellom arbeidstakerne som jobber på samme sted, vil bli kapitalisert i boligprisen, mens ulikhet i gjennomsnittlig kostnad ved å pendle mellom arbeidstakere som jobber på forskjellig sted, vil bli kapitalisert i lønningene.

Imidlertid vil lønnsforskjellene mellom sonene (sammen med den gjennomsnittlige pendler-kostnad i hver av dem) i prinsippet bli utjamnet med tida ved at bedriftene flytter for å redusere sine lønnskostnader. Når vi opplever lønnsforskjeller mellom soner, kan det derfor være fordi det tar lang tid å etablere likevekt, eller fordi det finnes agglomerasjonsfordeler. Det er derfor ikke utelukket at det finnes agglomerasjonsfordeler. Det eneste som er utelukket er at vi kan skille dem fra andre effekter ved å bruke aggregerte data og uten å spesifisere disse fordelene på en måte som kan etterprøves empirisk.

## 3.7 Frigjorte arealer

To typer av prosjekter har til virkning at de frigjør arealer som tidligere var brukt til transport. Den første typen er tunneler, og den andre typen er nedlegging av jernbanelinjer, eller i sjeldne tilfeller nedlegging av utdaterte veger, havner eller flyplasser. De frigjorte arealene kan få en økonomisk verdi, men som regel må betydelige opprydningskostnader trekkes fra.

I disse tilfellene kan også tilliggende arealer få endret verdi – større hvis det er skapt sammen-hengende arealer av tidligere oppsplittede og verdiløse eiendommer, og mindre hvis virksom-heter som lå langs vegen eller ved stasjonene, nå søker seg bort.

Når tomter blir *mindre* verdt på grunn av nedlegging av en jernbanelinje, en veg, en havn eller en flyplass, er det i de fleste tilfellene trolig hovedsakelig en rein pekuniær eksternalitet som ikke skal tas med i nyttekostnadsregnestykket. Frigjøring av arealer som tidligere var brukt til transport er derimot en virkelig samfunnsøkonomisk effekt, og det samme er økt verdi av arealene på grunn av fjerning av den barrieren som infrastrukturen tidligere utgjorde.

Vi antar at det bare er aktuelt å ta med slike effekter dersom det har meldt seg seriøse interessenter til de frigjorte arealene. Det vil da som regel også være mulig å anslå hvor mye tomteverdien vil endre seg.[65]

## 3.8 Ringvirkninger av byggeprosjektet

I keynesianismens glansdager (eller litt seinere) blei det bygd opp et par regionaliserte kryssløpsmodeller i Norge. Den ene het PANDA og den andre REGARD. PANDA drives og vedlikeholdes framleis og blir blant annet brukt til å vurdere ringvirkninger av større byggeprosjekter, herunder også transportprosjekter. Den drives av SINTEFs avdeling for anvendt økonomi. REGARD drives av SSB, men den siste publikasjonen på SSBs hjemmeside der REGARD er brukt, er fra 2004.

Mens dagens diskusjon dreier seg om hvordan *trafikantnytten* genererer mernytte, dreier disse modellene seg om positive virkninger av *kostnadene* (keynesianske multiplikatorvirkninger av etterspørselen etter arbeidskraft og andre ressurser som anlegget krever). Det sier seg sjøl at for at slike virkninger skal ha noe å si, må det eksistere ledige ressurser, spesielt ledig arbeidskraft. Et annet vilkår er at man ikke bare regner med de positive effektene for den regionen man er interessert i, men også eventuelle negative effekter i andre regioner.

I våre dager vil geografisk oppdelte generelle likevektsmodeller kunne brukes til å vurdere denne typen effekter, forutsatt at de har en realistisk gjengivelse av markedsimperfeksjoner og et realistisk transportnettverk. Dessuten må man klare å fange opp hvor ressursene til byggingen kommer fra, hvilket ikke er så lett i en verden med globalisert produksjon og anbud med internasjonal deltakelse.

# 4 Konklusjon

Mernytte skyldes dels at vi har modeller som ikke har med alle sammenhenger, og dels at diverse imperfeksjoner i økonomien utenfor transportsektoren gir opphav til pekuniære eksterne virkninger med virkelige økonomiske effekter. Stordriftsfordeler og agglomerasjonsfordeler som gir økt produktivitet er kanskje de viktigste.

Vår tilnærming til problemet med mernytte består av flere ting. Vi ønsker større innsats i utviklingen av modeller som fanger opp flere sammenhenger i transportsektoren. Vi ønsker modellutvikling som modeller sammenhengene mellom transport og andre sektorer, som arbeidsmarkedet, eller sammenhengen mellom transport og arealbruk, eller sammenhengen mellom transport, lagerhold og terminaler. Og vi ønsker ikke minst modellutvikling på området SCGE-modeller.

[65] Det finns et par svenske forslag til metode for verdsetting i slike tilfeller – se kapittel 17 og særlig kapittel 18 i SIKA (2008). De bør ikke brukes ukritisk, verken i Norge eller Sverige. Faktisk etterspørsel etter arealene kan gi verdsettingen økt realisme.

Om det skulle bli åpnet for å plusse på mernytte i våre analyser, krever det et lengre metode-utviklingsarbeid. For det første må vi mye klarere enn nå definere og avgrense de mekanismene som skaper mernytte i konkrete tilfeller. Mernytte finns bare i tilfeller der vi har markeds-imperfeksjoner. For det andre bør en, før en får lov til å legge til mernytte, kunne gi et fornuftig svar på hvorfor en ikke kan fjerne eller ta hensyn til markedsimperfeksjonen gjennom regu-lering eller økonomiske virkemidler. For det tredje må det foretas en analyse av det enkelte tilfellet. Om det vil kunne utvikles redskaper til å gjøre det, vil framtida vise. Om en kan bruke anslag og verdier overført fra andre land eller tilfeller, er mer enn tvilsomt, siden mernyttens størrelse vil være avhengig av blant annet skatteforhold, transportforhold og arealbruk og lokalisering.

# Litteratur


Alonso, W (1964) Location and land use: Towards a general theory of land rent. Harvard University Press.

Borukhov, E. and O. Hochman (1977) Optimum and market equilibrium in a model of a city without a predetermined center. *Environment and Planning A* **9**(8), 849-856.

Fujita, M (1988) A monopolistic competition model of spatial agglomeration: a differentiated product approach. *Regional Science and Urban Economics* **18**, 87-124.

Fujita, M (1990) Urban Economic Theory. Land Use and City Size. Cambridge University Press, Cambridge.

Harvey, D. (2001) Spaces of Capital. Towards a critical geography. Edinburgh University Press, Edinburgh, Chapter 14

Ivanova. O. (2003) The Role of Transport Infrastructure in Regional Economic Development. PhD Dissertation, Department of Economics, University of Oslo.

Jansson, J.-O. And R.E. Wall (2002) A new model for identifying and measuring re-organization benefits of improvements in transport infrastructure. Prepared for presentation at the 6th Workshop of the Nordic Research Network on Modelling Transport, Land-Use and the Environment, Haugesund, Norway, September 27-29 2002.

Jara-Diaz, S. (1986) On the Relation between User Benefits and the Economic Effects of Transportation Activities. *Journal of Regional Science* **26**(2), 379-391.

Kidokoro, Y. (2004) Cost-Benefit Analysis for Transport Networks: Theory and Applications. *Journal of Transport Economics and Policy* **38**(2), 275-307.

Laird, J.J., J. Nellthorp, and P.J. Mackie (2005) Network effects and total economic impact in transport appraisal. *Transport Policy* **12**(6), 537-544.

Lakshmanan, T.R. (2010) The broader economic consequences of transport infrastructure investments. *Journal of Transport Geography* **19**, 1-12.

Minken, H. (2014) Industrial reorganisation benefits revisited. *Journal of Transport Economics and Policy* 48(1), pp. 53-63. (Eller TØI arbeidsdokument ØL/2332/2011.)

Minken, H. (2009) Rammeverk for nyttekostnadsanalyse og finansieringsanalyse. Arbeidsdokument ØL/2156/2009, TØI. (Dette er artikkelen i avsnitt 2.1 i vår rapport her. Også tilgjengelig som vedlegg 4 i Minken m.fl. (2009).

Minken, H., O.I. Larsen, J.H. Braute, S. Berntsen og T. Sunde (2009) Konseptvalgsutredninger og samfunnsøkonomisk analyse. TØI-rapport 1011/2009.

Mohring, H (1961) Land values and the measurement of highway benefits. *Journal of Political Economy* **69**(3), 236-249.

Mohring, H. (1993) Maximizing, measuring and not double counting transportation-improving benefits: A primer on closed- and open-economy cost-benefit analysis. *Transportation Research B* **27**(6), 413-424.

Mohring, H. and H.F. Williamson Jr. (1969) Scale and "industrial reorganisation" economies of transport improvements. *Journal of transport Economics and Policy* **3**(3), 251- 271.

Mori, T (2006) Monocentric versus Polycentric Models in Urban Economics. Discussion Paper No. 611, Kyoto Institute of Economic Research.

SIKA (2008) Samhällsekonomiska principer och kalkylvärden för transportsektorn: ASEK 4. SIKA PM 2008:3.

Small, KA (1997) Project evaluation. Working paper UCI-IST-97-06, University of California, Irvine.

Starrett, D (1978) Market Allocations of Location Choice in a Model with Free Mobility. *Journal of Economic Theory* **17**, 21-37.

Timothy, D and WC Wheaton (2001) Intra-Urban Wage Variation, Employment Location and Commuting Times. *Journal of Urban Economics* **50**(2), 338-366.

Venables, A.J (2007) Evaluating Urban Transport Improvements: Cost Benefit Analysis in the Presence of Agglomeration and Income Taxation. *Journal of Transport Economics and Policy* **41**(2), 173-188.

Von Thünen, J-H (1826) Der isolierte Staat in Beziehung auf Landwirtschaft und Nationaloekonomie. http://www.archive.org/details/derisoliertestaa00thuoft

Wall, R. (2002) The importance of transport costs for spatial structures and competition in goods and service industries. PhD Dissertation, University of Linköping.

Williams, H.C.W.L. (1977) On the formation of travel demand models and economic evaluation measures of user benefits. *Environment and Planning A*, **9**, 285-344.

Wheaton, WC (1977) Residential decentralization, Land Rents, and the Benefits of Urban Transportation Investment. *American Economic Review* **67**(2), 138-143.

## 3.5 Samfunnsøkonomisk optimal køpris i det lange løp[66]

[66] Dette er et utdrag fra et vedlegg i Dovre Group og TØIs kvalitetssikringsrapport (KS1) av Oslopakka trinn 2 i 2017. Vedlegget var skrevet av Harald Minken.

# Samfunnsøkonomisk optimal køpris i det lange løp

I dette avsnittet viser vi hvordan samfunnsøkonomisk riktige bompengesatser bør endres ved endringer i de ytre omstendighetene, som befolkningsøkning, forbedringer i vegsystemet og forbedringer i kollektivsystemet og for gående og syklende. Vi viser også at nullvekstmålet nødvendigvis kommer i konflikt med samfunnsøkonomiske mål (riktige køpriser) om det ligger fast over lengre tidsrom.[67]

Vi ser på en enkel modell av et bytransportsystem der optimal køprising er innført. Det er en multinomisk logitmodell med tre reisemåter – bil, kollektivtransport og gåing og sykling. Vi antar at kollektivsystemet ikke på noen måte deler infrastruktur med bilene. Heller ikke de gående og syklende gjør det. Vårt problem er hvordan optimale køpriser skal justeres når de ytre omstendighetene endrer seg.

De ytre omstendighetene vi ser på er:

- Samlet befolkning eller samlet antall reiser i byområdet, $N$
- Kvaliteten av vegsystemet, som vi vil måle ved $h = e^a$, der $a$ er den alternativspesifikke konstanten i den betingede nyttefunksjonen for bilreiser
- Den systematiske nytten av reiser med de to andre reisemåtene, kollektivtransport og gåing og sykling. Summen av de to nyttene skriver vi $K$ for enkelhets skyld.

Siden vi nå har redusert informasjonen vi trenger om de to andre reisemåtene til et eneste tall $K$, kan vi klare oss uten egen fotskrift på reisemåten bil. La $x$ være antall bilreiser, $p$ være køavgifta, $v$ tidsverdien, $t(x)$ den typiske reisetida i området (en tiltakende og konveks funksjon av antall bilreiser), og la $g$ være generaliserte reisekostnader. Alle disse variablene gjelder bilreiser.

Likning (1) nedenfor er definisjonen av generaliserte kostnader. Likning (2) er den optimale køavgifta. Enhver reise i et købelastet område påfører alle andre reisende en marginal reisetidsøkning på $t'(x)$. Dette tidstapet, som varierer med trafikknivået, vil en som skal fatte en beslutning om å reise, ikke ta hensyn til medmindre det gjenspeiles i vegavgifta $p$. For å maksimere samfunnsnytten av transportsystemet bør derfor $p$ settes lik summen av alle disse marginale tidstapene, verdsatt med tidsverdien $v$ til den gjennomsnittlige vegbrukeren. Men siden trafikknivået (og kanskje også sammensetningen av trafikken) varierer over døgnet, bør avgifta $p$ også variere. I den enkle modellen her har vi implisitt forutsatt en bestemt tid på dagen. Konklusjonene våre vil imidlertid være gyldige for et hvilket som helst tidspunkt på dagen.

Likning (3) er en etterspørselsfunksjon av typen «multinomisk logit» for reiser over en vegstrekning eller i et byområde på et gitt tidspunkt på dagen. Konstanten $K$ er naturligvis et forenklet uttrykk for summen av den indirekte nytten som er forbundet med en kollektivreise og den indirekte nytten som knytter seg til en gang- eller sykkelreise. Sagt på en annen måte: $K = h_{pt}e^{-\lambda g_{pt}} + h_{wc}e^{-\lambda g_{wc}}$. Vi trenger ikke her spesifisere nøyaktig hvilke variabler i uttrykket for $K$ som har endret seg når vi nå skal studere en eksogen endring i variabelen $K$.

[67] Framstillingen krever matematikkunnskaper.

Likning (1)-(3) nedenfor utgjør hele modellen.

$$(1) \qquad g = p + vt(x)$$

$$(2) \qquad p = xvt'(x)$$

$$(3) \qquad x = N\frac{he^{-\lambda g}}{he^{-\lambda g} + K}$$

Nå innfører vi de tre typene av eksogene endringer som vi nevnte ovenfor, altså *dN*, *dh* and *dK*. Vi totaldifferensierer hele modellsystemet, og får:

$$(4) \qquad dg = dp + vt'(x)dx$$

$$(5) \qquad dp = v\left(t'(x) + xt''(x)\right)dx$$

$$(6) \qquad \left(K + he^{-\lambda g}\right)dx + \lambda(N - x)he^{-\lambda g}dg = he^{-\lambda g}dN + (N - x)e^{-\lambda g}dh - xdK$$

Vi eliminerer nå *dg* ved å sette inn (4) i (6) og ordne, i det vi bruker forkortelsen $\lambda(N - x)he^{-\lambda g} = B$:

$$(7) \qquad \left[K + he^{-\lambda g} + Bvt'(x)\right]dx + Bdp = he^{-\lambda g}dN + (N - x)e^{-\lambda g}dh - xdK$$

Nå har vi to muligheter: vi kan bruke (5) for å eliminere *dp* fra (7), eller vi kan bruke samme likning for å eliminere *dx* i stedet. Den første muligheten vil gi oss et uttrykk for hvordan endringer i *N*, *h* og *K* påvirker biltrafikken under optimal vegprising, mens den andre uttrykker hvordan vegprisen sjøl skal endre seg når omstendighetene endrer seg. Vi forfølger begge muligheter i likning (8).

$$(8) \qquad \begin{aligned} dx &= \frac{1}{K + he^{-\lambda g} + Bv\left[2t'(x) + xt''(x)\right]} \cdot \left[he^{-\lambda g}dN + (N - x)e^{-\lambda g}dh - xdK\right] \\ dp &= \frac{v\left[t'(x) + xt''(x)\right]}{K + he^{-\lambda g} + Bvt'(x)} \cdot \left[he^{-\lambda g}dN + (N - x)e^{-\lambda g}dh - xdK\right] \end{aligned}$$

Vi ser tydelig av første linje i (8) at det optimale trafikknivået skal gå opp når befolkningen øker og dersom kvaliteten av vegsystemet øker. Derimot skal den gå ned om det skjer forbedringer i andre transportmåter. Det er vårt første funn her.

En optimal køavgift skal øke forholdsvis mer enn trafikkvolumet når det skjer endringer i *N* og *h*. Den skal også reduseres mer enn trafikknivået på veg når de konkurrerende reisemåtene forbedres. Vi skal altså reagere forholdsvis kraftig på endringer i de ytre omstendighetene. Det framgår ved å sammenlikne første og andre linje i (8). Det er vårt andre funn.

Siden *dx* åpenbart øker med *dN* og *dh* og avtar med *dK*, er det klart at en politikk for å holde privatbilbruken på et konstant nivå («nullvekstmålet») ikke lar seg forene med det økonomiske prinsippet om marginalkostnadsprising. Det er vårt tredje funn.

## 3.6 Kollektivselskapets kostnader, optimalt kollektivtilbud og verdien av forbedringer[68]

[68] Dette er et upublisert arbeidsdokument fra 2009, ØL/2157/2009. Det finnes også et seinere arbeids-dokument, nummer 50272 fra 2012, med noen forenklinger i forhold til dette, men også med eksterne kostnader, som ikke er med her.

# Kollektivselskapets kostnader, optimalt kollektivtilbud og verdien av forbedringer

## Innhold

# 1 Innledning

Dette arbeidsdokumentet har flere hensikter. For det første vil vi lansere en enkel analytisk modell som kan danne utgangspunkt for å studere eller beregne kostnadene til et kollektivselskap. Den fanger opp de viktigste aspektene som påvirker kostnadene på en kollektivlinje eller en samling av likeartede linjer. Men ikke alle. Blant de utelatte aspektene er gjennomsnittsavstanden mellom holdeplassene og tida det tar å stoppe ved en holdeplass. Driftsforstyrrelser, vedlikehold og disse faktorenes konsekvenser for behovet for reservemateriell behandles heller ikke.

For det andre behandler vi kollektivselskapets beste tilpasning av pris, frekvens og flatedekning, både når bedriftsøkonomisk overskudd og samfunnsøkonomi er målet. Det som skiller de to tilfellene er prissettingen. Med samme prissetting vil selskapet by fram samme tilbud, uansett om bedriftsøkonomi eller samfunnsøkonomi er målet.

For det tredje gjennomfører vi nyttekostnadsanalyser av følgende endringer: Tiltak som endrer maksimalt tillatt frekvens, tiltak som muliggjør større kapasitet pr. avgang, framkommelighetstiltak og eksogene skift i etterspørselen (for eksempel forårsaket av innføring av køprising på vegen). Hensikten er å få fram hvilke direkte og indirekte virkninger som gjør seg gjeldende i slike tilfeller. Konkrete analyser vil måtte ta hensyn til fordelingen av etterspørselen og tilbudet i tid og rom på en ganske annen måte enn det vi gjør i vår modell, som behandler geografien på en abstrakt og gjennomsnittlig måte. Våre nyttekostnadsanalyser vil likevel gi en pekepinn om virkningene på et konseptuelt nivå, og de har den fordelen framfor tester med store modeller at de inkluderer kollektivselskapets endringer av tilbudet når vilkårene endrer seg.

Kostnadene ved å gi et ulikt tilbud i og utenom rush og tilpasningen i toperiodetilfellet tas opp i siste kapittel.

## 1.1 Effekter av at kollektivselskapet endrer tilpasning

Når vi tester et tiltak med en transportmodell, må vi først kode et bestemt kollektivtilbud i modellen. Dvs. at vi må legge inn kollektivlinjene og driftsopplegget på hver av linjene. Dette er altså data som hentes inn utenfra og ikke endrer seg, uansett hva som skjer med etterspørselen og framkommelighetsforholdene som følge av tiltaket.

Anta at tiltaket består i en økning av avgiftene på vegsida, med mindre kø på vegene og økt etterspørsel etter kollektivreiser til følge. I virkeligheten vil det ofte kunne gi grunnlag for at kollektivselskapet endrer tilbudet. De kan endre antall avganger, øke kapasiteten pr. avgang (lengre tog, større busser), tilpasse ruteplanene til den økte framkommeligheten, eller opprette nye linjer. Vi må anta at i den grad dette er lønnsomt for dem, vil de faktisk også gjøre det. Dermed oppstår følgende effekter, som bare delvis eller slett ikke er med i nytteberegningen basert på transportmodellen:

1. Kollektivtrafikantenes generaliserte kostnader endrer seg
2. Kollektivselskapenes inntekter og kostnader endrer seg
3. Behovet for tilskudd fra det offentlige endrer seg

*Ang. punkt 1:* I den store transportmodellen vil kollektivtrafikantenes *generaliserte kostnader* ikke endre seg, og følgelig blir nytten for kollektivtrafikantene lik null. I virkeligheten vil kollektivselskapet endre sitt tilbud når flere velger kollektivt på grunn av tiltaket, og dermed vil generaliserte kostnader også endre seg. Vi kan kalle det en utelatt andreordenseffekt av prisøkningen på vegsida.

*Ang. punkt 2:* Kollektivselskapets *inntekter* vil også øke, fordi flere vil velge kollektivt. Denne førsteordenseffekten er *med* i nytteberegningen. Det som *ikke* er med, er andreordenseffekten når kollektivselskapet bruker de økte inntektene til forbedringer i tilbudet. Så lenge tilbudet er fastlagt på forhånd i modellen, vil kollektivselskapets *kostnader* også være uendret i nytteberegningen.

*Ang. punkt 3:* Førsteordenseffekten av tiltaket på kollektivselskapets inntekter vil naturligvis gi endringer i tilskuddsbehovet. Dette er *med* i nytteberegningene med transportmodellen. Det som *ikke* er med når billettinntektene øker på grunn av tiltaket, er andreordensvirkingene på tilskuddsbehovet. Når det gjelder kollektivselskapets kostnader, er verken første- eller andreordensvirkningene av kostnadsendringer på tilskuddsbehovet med i nytteberegningen.

Small (2004) viser at andreordensvirkningene av vegprising kan være betydelige. Det vil si at vegprising setter i gang en positiv, sjølforsterkende prosess på kollektivsida. Hovedelementet i denne prosessen er at når biltrafikken reduseres, vil framkommeligheten for bussene bli bedre, hvilket gir kortere reisetid for trafikantene og lavere kostnader for kollektivselskapet. Et annet element er den økte etterspørselen etter kollektivreiser, som skaper grunnlag for hyppigere avganger og bedre flatedekning (nye busslinjer). Det gir lavere reisekostnader for alle kollektivtrafikanter. Kollektivselskapet får økte inntekter, men også økte kostnader.

Den familien av tilbudsmodeller som vi utvikler i dette arbeidsdokumentet, er beslektet med Smalls. Vi viser hvordan det gir grunnlag for å beregne annenordenseffekter som ikke er med i vanlige transportmodeller, og antyder hvordan tilbudsmodellen vår kan integreres i et vanlig transportmodellsystem.

Åpenbart vil den samfunnsøkonomiske lønnsomheten av noen typer av kollektivtiltak kunne vise seg å være større enn hittil beregnet dersom annenordenseffektene inkluderes, sjøl om det i praksis må gjøres på en grov og forenklet måte. Arbeidsdokumentet antyder forøvrig også andre kilder til at kollektivprosjekter ofte kommer så dårlig ut i slike beregninger: Det nytter ikke å eliminere en kapasitetsskranke hvis kapasiteten ikke er den begrensende faktoren, eller hvis beste tilpasning begrenses av andre skranker. Resultatet kan også bli tilfeldig hvis verken nåværende driftsopplegg og prissetting eller det nye opplegget som foreslås, er optimalt. Da vil det finnes billigere (eller helt kostnadsfrie) muligheter til forbedring.

# 2 En modell av kollektivselskapets kostnader

Vi tar utgangspunkt i en samling av noenlunde likeartede kollektivlinjer. Det kan for eksempel være alle t-banelinjer eller alle de viktigste busslinjene i Oslo. De viktigste variablene og parametrene er definert i tabell 1.

*Tabell 1 Variable og parametere*

| | |
|---|---|
| $a$ | lengda av en rundtur på en gjennomsnittlig linje (km) |
| $c$ | kapasitet pr. avgang (passasjerplasser) |
| $f$ | frekvens (antall avganger pr. time) på en gjennomsnittlig linje |
| $g$ | kilometeravhengig kostnad (kr/km) |
| $G$ | generalisert reisekostnad (kr) |
| $h$ | antall driftstimer pr. år på en gjennomsnittlig linje |
| $k$ | behovet for kjøretøyer (busser, togsett) på en gjennomsnittlig linje |
| $\ell$ | bemanning pr. kjøretøy |
| $m$ | gjennomsnittlig reiselengde (km) |
| $N$ | antall linjer |
| $p$ | billettpris (kr) |
| $r$ | kapitalkostnad pr. kjøretøy inklusive påslag for reservemateriell (kr/år) |
| $s$ | kjøretøyets gjennomsnittsfart inklusive reguleringstid (km/time) |
| $t$ | gjennomsnittlig rundturtid (timer) |
| $u$ | reguleringstid (timer) |
| $v$ | marsjfart (eksklusive reguleringstid) (km/time) |
| $w$ | lønnskostnad pr. driftstime inklusive påslag for uproduktiv tid (kr/time) |
| $x$ | etterspørsel pr. time, alle linjer |
| $\varepsilon$ | skiftparameter i etterspørselen |
| $\omega$ | tidsverdi (kr/time) |
| $\varphi$ | gjennomsnittsbelegg (en parameter mellom 0 og 1 som uttrykker hvordan etterspørselen er fordelt langsmed linjene) |

## 2.1 Rundturtid, marsjfart og tid ombord

Rundturtida på en linje er i gjennomsnitt $t = av^{-1} + u$, der *a* er antall kilometer på en gjennomsnittlig linje, *v* er marsjfart og *u* er reguleringstid. [69] Gjennomsnittsfarta inklusive reguleringstida kaller vi *s*:

$$s = \frac{a}{t} = \frac{a}{\frac{a}{v} + u} = \frac{av}{a + uv} \tag{1}$$

Gjennomsnittlig tid ombord for en kollektivreise vil være $mv^{-1}$.

## 2.2 Tidsavhengige og kilometeravhengige kostnader

Vi ser bort fra at det finnes en høytrafikk- og en lavtrafikkperiode, og holder oss til en gjennomsnittlig time. [70]

Kollektivtilbudet er oppdelt i et antall kollektivlinjer. Vi antar at hver linje bruker rullende materiell og mannskap som er dedikert til denne linja. Driften på linja består av et antall avganger *f* pr. time. Avgangene er rundturer, slik at etter utløpet av rundturtida er kjøretøyet tilbake ved utgangspunktet og kan ta neste tur. Med et kjøretøy mener vi en buss eller en trikk eller et togsett, slik at det rullende materiellet som brukes til en avgang er ett og bare ett kjøretøy. Det er da enkelt å vise at om antall kjøretøyer som trengs til drifta er *k*, er $k \geq tf$. Likhet er mulig om *k* ikke behøver å være et helt tall, eller om reguleringstida kan tilpasses for å oppnå likhet sjøl om *k* er heltallig. Vi antar likhet, altså $k = tf$.

Kapasiteten pr. avgang (eller pr. kjøretøy), målt i maksimalt antall plasser for passasjerene, kaller vi *c*. Den årlige kapitalkostnaden *r* pr. kjøretøy i perioden vi ser på, er en funksjon av *c*. Kilometerkostnaden *g* er også en funksjon av *c*, mens derimot bemanningen pr. kjøretøy, λ, her antas å være uavhengig av kjøretøykapasiteten. (For de fleste driftsarter er det en sjåfør eller togfører, og ingen andre.) Med ganske godt belegg i empiri mener vi å kunne anta lineære sammenhenger:

$$\begin{aligned} r &= r_0 + r_1 c \\ \ell &= \ell_0 \\ g &= g_0 + g_1 c \end{aligned} \tag{2}$$

---

[69] Marsjfarta *v* kan i noen sammenhenger, der bussene går i vegen for hverandre eller togene må vente på hverandre, være en funksjon av frekvensen *f*: $v = v(f; v_0, \alpha)$, der $v_0$ er farta når bussene ikke går i vegen for hverandre og $\alpha$ er en skiftparameter som for eksempel kan representere forbedringer av signalsystemet for skinnegående trafikk. I dette notatet antar vi imidlertid at *v* er en gitt parameter.

Marsjfarta, slik vi har definert den, er også en funksjon av antall stoppesteder pr. kilometer. Antall stoppesteder kunne vært en handlingsvariabel i modellen, men det har vi ikke tatt høyde for her.

[70] I den grad kostnader kan allokeres entydig til enten høy- eller lavtrafikkperioden, kan en bygge separate modeller for høy- og lavtrafikken. Se denne artikkelens kapittel 7.

La lønnskostnaden pr. driftstime for et medlem av bemanningen være *w*, og anta at det er *h* driftstimer i året. Den årlige tidskostnaden på linja er da $(r+hw\ell)k$, og den årlige kilometer-kostnaden er *hksg* = *hafg*.

## 2.3 Årlig kostnad for et sett av likeartede linjer

Anta det finns *N* likeartede linjer. Kollektivselskapets årlige kostnader $C_{år}$ er summen av kostnadene på linjene. Vi bruker (1), (2) og *k* = *tf* og får:

$$
\begin{aligned}
C_{år} &= N\left\{(r+hw\ell)k+hafg\right\} \\
&= N\left\{(r_0+hw\ell_0)t+r_1ct+hag_0+hag_1c\right\}f \\
&= N\left\{\frac{(r_0+hw\ell_0)t+hag_0}{a}+\frac{r_1t+hag_1}{a}c\right\}af \qquad (3)\\
&= N\left\{\left[(r_0+hw\ell_0)s^{-1}+hg_0\right]+\left[r_1s^{-1}+hg_1\right]c\right\}af \\
&= N\left\{\left[(r_0+hw\ell_0)+hg_0s\right]+\left[r_1+hg_1s\right]c\right\}\frac{af}{s}
\end{aligned}
$$

Det er lett å sjekke at *ks* = *af*, så $C_{år}$ er proporsjonal med antall kjøretøyer i bruk, *Naf/s*.

*Kollektivselskapets handlingsrom*

I det kollektivsystemet av likeartede linjer som vi ser på, er mulige handlingsvariable *c*, *f*, *N* og billettprisen *p*. Handlingsvariablene kan ligge i hendene på myndighetene eller kollektiv-selskapet. Nedenfor antar vi først at de ligger i hendene på kollektivselskapet, som maksimerer profitt. Dernest antar vi at de ligger i hendene på myndighetene, og at myndighetene ønsker å maksimere samfunnsøkonomisk lønnsomhet.

Men uansett hvem som bestemmer virkemiddelbruken, finns det visse skranker som en blir nødt til å ta hensyn til. Det kan eksistere en øvre og en nedre grense for kjøretøyenes kapasitet, $c_{\min}\le c\le c_{\max}$. Siden vi her er interessert i kollektivtrafikken i store byer, kan vi gå ut fra at det aldri er aktuelt å bruke de aller minste kjøretøyene, så bare den øvre grensa kan være et problem. (For eksempel vil mer enn 6 vogner på t-banen kreve utvidelse av alle perrongene.) Videre kan det finnes en øvre grense for hvor stor frekvens en kan velge, $f\le\overline{f}$. (For eksempel er det ikke mulig med flere enn 28 avganger pr. time gjennom t-banetunnelen dersom en skal bygge på kvartersruter, se Minken og Dahl (2007)).[71]

Vi må også ta hensyn til at passasjerer ikke skal måtte bli stående igjen på perrongen eller holdeplassen, sjøl ikke på den strekningen hvor kjøretøyet er som fullest. Hvor stor kapasitet dette krever, avhenger av hvordan etterspørselen er fordelt langsmed linjene. La $\varphi$ være en parameter mellom 0 og 1 som uttrykker dette. Når $\varphi$ = 1, er etterspørselen helt jamt fordelt, og jo mer ujamnt den fordeler seg, jo mindre er $\varphi$. Videre vil den nødvendige kapasiteten bli større jo lenge reisene er, målt i forhold til rundturdistansen *a*. Vi har antatt at gjennomsnittlig

[71] Det er altså i høyden én linje som kan få doble avganger, med mindre en bygger ny tunnel, bygger Lørensvingen eller bruker Ensjøforbindelsen. Ved 12-minuttersruter på alle linjer kan en imidlertid oppnå 30 avganger pr. time.

reiselengde er $m$. Antar vi at etterspørselen i systemet som helhet er $x$, vil vilkåret for at alle kommer med på den avgangen de ønsker, være

$$\varphi cf \geq \frac{m}{a}\frac{x}{N}$$

Det kan vises at ulikhetstegnet i dette vilkåret ikke kan gi samfunnsøkonomisk beste løsning med mindre det finnes en bindende minste kjøretøykapasitet. Det samme gjelder hvis selskapet maksimerer profitt. Vi kan derfor anta likhetstegn:

(4) $$c = \varphi^{-1}\frac{m}{a}\frac{x}{Nf}$$

For mange formål er det greiere å operere med kostnader og inntekter pr. år enn pr. time. Det er blant annet lettere å bruke regnskapsdata. På den andre sida er det hensiktsmessig å bygge på kostnader pr. time dersom vi trenger å skille mellom driftstilbudet i ulike perioder. Så lenge vi veit hvor mange driftstimer det er i et år, dvs. når vi kjenner $h$, spiller det ingen rolle for noen av regnestykkene våre om vi gjør det ene eller det andre, bare vi er konsistente. Fra nå av regner vi kostnader og inntekter pr. time. Vi definerer kostnaden pr. time, $C$, som $C_{år}/h$. Setter vi (4) inn i (3), bruker $C = C_{år}/h$ og ordner, har vi:

(5) $$C = C_1 + C_2 = \left[\left(r_0/h + w\ell_0\right) + g_0 s\right] N\frac{af}{s} + \left[r_1/h + g_1 s\right]\varphi^{-1}\frac{m}{s}x$$

Bortsett fra terminologi er formuleringen av kollektivselskapets kostnader i likning (5) identisk med formuleringen hos Small med to unntak. Det første er beleggsprosenten $\varphi$, som mangler hos Small. Når den er mindre enn 1, gir den større vekt til den etterspørselsavhengige delen av kostnadene, $C_2$. Det andre unntaket gjelder innholdet i de to klammeparentesene. I stedet for klammeparentesene har Small to konstanter, mens vi har spesifisert kapitalkostnader, lønnskostnader og kilometeravhengige kostnader. Det som faktisk gjør en stor forskjell, er at spesifikasjonene er funksjoner av $s$.

Vi kaller den første klammeparentesen $A(s)$ og den andre $B(s)$. Videre tar vi hensyn til at etterspørselen er en avtakende funksjon av generaliserte kostnader $G$:

(6) $$x = x(G) = \varepsilon D(G), \quad \varepsilon > 0,\ D' \leq 0,\ \lim_{G\to\infty} D = 0$$

Her er $\varepsilon$ en skiftparameter som i utgangspunktet er 1, men kan få et tillegg eller fratrekk som følge av vegprising, befolkningsendring e.l. Nå kan vi kan skrive

(7) $$C = C_1 + C_2 = A(s)N\frac{af}{s} + B(s)\varphi^{-1}\frac{m}{s}x(G)$$

Vi antar videre at generaliserte reisekostnader er en lineær funksjon av billettprisen og tidskostnaden, og at tidskostnaden består av tid om bord, gangtid og skjult ventetid:

$$G = p + E(s) + Vf^{-1} + LN^{-1}$$

der

(8) $$E(s) = \omega \frac{m}{a}\left(\frac{a}{s} - u\right)$$

$$V = \frac{1}{2}\omega r_v$$

$$L = \omega r_g d$$

Parentesen i uttrykket for $E(s)$ er tida kjøretøyet bruker på en rundtur, reguleringstida ikke medregnet. Når den multipliseres med den gjennomsnittlige reiselengda som andel av rundturlengda, får vi tida om bord for en gjennomsnittsreise. Vi antar at frekvensen er såpass høy at folk ankommer tilfeldig til holdeplassen. Ventetida på holdeplass er da halvparten av tida mellom avgangene. $r_v$ og $r_g$ er henholdsvis vektene for ventetid og gangtid (se SVV 2006), og $d$ er bestemt ved at $d/N$ er den gjennomsnittlige gangtida til holdeplass i utgangspunktet.

Forutsetningen om at alle har samme tidsverdi slik at billettkostnaden og tidskostnadene kan adderes til en generalisert kostnad, er urealistisk men vanlig.

## 2.4 Nærmere om flatedekning

God flatedekning betyr at gangavstanden fra der hvor reisa starter til holdeplass eller stoppested, og fra holdeplass til der hvor reisa ender, er kort. Vi antar at den er omvendt proporsjonal med antall linjer. Det er denne antakelsen som ligger bak ønsket om å bruke antall linjer $N$ som en handlingsvariabel. Men det behøver ikke alltid være slik. For det første vil det kunne finnes kollektivlinjer av et annet slag i kollektivsystemet, slik at gangavstanden til slike andre linjer er vel så relevant som gangavstanden til en av de $N$ linjene i vår modell. For det andre kan nye linjer legges slik geografisk at de påvirker gangavstanden lite for de fleste eksisterende kollektivpassasjerer. I begge disse tilfellene er det best å ikke anta at $N$ er en handlingsvariabel, men la modellen gjelde for en enkelt linje, altså $N = 1$.

For det tredje er det ikke bare antall linjer, men også avstanden mellom stoppestedene som bestemmer gangavstanden. Men avstanden mellom stoppestedene påvirker også marsjfarta negativt, slik vi har definert den. Dette kunne også ha vært en handlingsvariabel i modellen, men vi har valgt å utelate det her. Den delen av gangavstanden som skyldes at folk bor *mellom* stasjonene, er da er fast kostnad som vi kan se bort fra. Bare den korteste avstanden hjemmefra til *linja* teller med.

Noen stiliserte eksempler på hvordan $N$ påvirker gangavstanden:

1. En rektangulær by der linjene ligger parallelt og vannrett med like stor avstand og etterspørselen oppstår jamt langs loddrette linjer gjennom stasjonene. Ingen gangtid på bestemmelsesstedet. Bredden av byen er $b$ og lengda ½$a$: Gjennomsnittlig gangavstand er da ¼ $b/N$, dvs. $d$ = ¼ $b$. Siden vi ignorerer avstanden langs linja til stasjonen, gjelder dette også om folk bor jamt spredt over hele byen.

2. En sirkelformet by der linjene går radialt til sentrum og etterspørselen er jamt fordelt over hele byen. Byen har radius $b$. Avstanden mellom stoppestedene er $d_s$. De som bor i avstand $r$ fra sentrum og skal ta bussen, går langs sirkelbuen i avstand $r$ til de kommer til en kollektivlinje, deretter oppover eller nedover linja til en stasjon. Denne siste biten

av gangavstanden ser vi bort fra. Gjennomsnittlig gangavstand er da $\frac{1}{4}\pi\, b/N$, dvs. $d = \frac{1}{4}\pi b$.

3. Vi tenker oss at rundt hver stasjon er det et influensområde av form som en regulær sekskant. All etterspørsel kommer fra disse influensområdene. Når *N* er liten, er det store områder i byen som har så lange gangavstander at kollektivtransport anses som uaktuelt. I dette tilfellet påvirkes ikke gangavstanden av at det blir flere linjer. Når *N* er stor nok, dekker de sekskantede områdene hele byen. I sentrum av hver sekskant er det en stasjon. Vi må nå anta at når *N* øker, reduseres samtidig avstanden mellom stasjonene (ellers vil våre sekskanter ikke lenger bli regulære). Vi antar også at byen er kvadratisk med sidekant *b* og at det finns like mange ”horisontale” som vertikale” linjer, altså ½ *N* linjer av hvert slag. Avstanden mellom linjene (som også er avstanden mellom stasjonene) kan vises å være $r\sqrt{3}$, der *r* er radien i sirkelen som omslutter sekskanten. *r* er også gangavstanden til de som har lengst veg til stasjonen.

   Hvis sidekanten på byen, *b*, gir plass til ½ *N* linjer, hver med bredde $r\sqrt{3}$, får vi at $r = \frac{2}{\sqrt{3}} \cdot \frac{b}{N}$. Alle i sekskanten går raskeste veg direkte til stasjonen. Det kan da vises at gjennomsnittlig gangavstand er $\frac{1}{4}\left(1 + \frac{1}{2}\sqrt{3}\right) r = \frac{2+\sqrt{3}}{4\sqrt{3}} \cdot \frac{b}{N}$. I dette tilfellet er altså $d = \frac{2+\sqrt{3}}{4\sqrt{3}} \cdot b$.

De tre eksemplene viser at i en by med dårlig kollektivtilbud vil gangavstanden ikke påvirkes nevneverdig av nye linjer, som i alle tilfeller blir liggende for langt fra hverandre til å utgjøre noe alternativ for de som sogner til en bestemt stasjon. I slike tilfeller gjør vi best i å anta *N* = 1. De viser også at når kollektivtilbudet er så godt at det er gangavstand mellom linjene og stasjonene, er det slett ikke noen urealistisk antakelse at gangavstanden er omvendt proporsjonal med antall linjer. De som mener å gjenkjenne sin egen by i et av eksemplene, kan kanskje gjøre antakelser om *d* i samsvar med det, men for øvrig gjør man nok best i å estimere *d* på grunnlag av observerte gangtider i nåsituasjonen.

# 3 En familie av tilbudsmodeller

En tilbudsmodell må bygge på en forutsetning om hvordan kollektivselskapet tilpasser seg. To typer forutsetninger er de vanligste, profittmaksimering og maksimering av samfunnsøkonomisk overskudd (velferd).

## 3.1 Profittmaksimering

Profitten er $\Pi = px - C$, så profittmaksimeringsproblemet er

(9)
$$\max_{G,p,f,N} \Pi = \left( p - B(s)\varphi^{-1}\frac{m}{s} \right) x(G) - A(s)\frac{a}{s}Nf$$
gitt
$$p + E(s) + Vf^{-1} + LN^{-1} = G \qquad (\mu)$$

Førsteordensbetingelsene for maksimum er:

(10)
$$\frac{\partial L}{\partial G} = \left( p - B(s)\varphi^{-1}\frac{m}{s} \right)\frac{\partial x}{\partial G} + \mu = 0$$

(11)
$$\frac{\partial L}{\partial p} = x - \mu = 0$$

(12)
$$\frac{\partial L}{\partial f} = -A(s)\frac{a}{s}N + \mu V f^{-2} = 0$$

(13)
$$\frac{\partial L}{\partial N} = -A(s)\frac{a}{s}f + \mu L N^{-2} = 0$$

Av (11) får vi $\mu = x$. Med denne verdien innsatt kan (10) skrives

(14)
$$\frac{p - B(s)\varphi^{-1}\frac{m}{s}}{p} = -\frac{1}{El_p x}$$

Likning (14) kan sies å være regelen for hvordan et profittmaksimerende kollektivselskap skal sette prisen. Denne regelen er identisk med den klassiske formelen for monopolprising, som sier at forholdet mellom grenseprofitten og prisen skal være omvendt proporsjonalt med tallverdien av priselastisiteten. Det framgår av formelen at i tilpasningspunktet er priselastisiteten mindre enn $-1$. Hvis den ikke er det, driver selskapet ikke profittmaksimering. Det vil nemlig alltid ha en viss markedsmakt.

Fra (12) og (13) kan vi beregne optimal frekvens og flatedekning som funksjoner av $x$:

(15)
$$f = \left( A(s)\frac{a}{s} \right)^{-\frac{1}{3}} V^{\frac{2}{3}} L^{-\frac{1}{3}} x^{\frac{1}{3}}$$
$$N = \left( A(s)\frac{a}{s} \right)^{-\frac{1}{3}} V^{-\frac{1}{3}} L^{\frac{2}{3}} x^{\frac{1}{3}}$$

Likningene (14) og (15) er ikke en eksplisitt løsning, siden $x$ og elastisiteten til $x$ er funksjoner av $G$ og dermed av $N$ og $f$.

Ved innsetting av (15) i (4) har vi også:

(16) $$c = \varphi^{-1}\frac{m}{a}\left(A(s)\frac{a}{s}\right)^{\frac{2}{3}}(VL)^{-\frac{1}{3}}x^{\frac{1}{3}}$$

## 3.2 Samfunnsøkonomi

La oss i stedet anta at kollektivselskapet maksimerer samfunnsøkonomisk velferd. For enkelhets skyld ser vi bort fra ulykkeskostnader, miljøkostnader og skatt på billetter og innsatsfaktorer. Det samfunnsøkonomiske regnestykket består da av trafikantenes brukernytte pluss kollektivselskapets profitt. Vi antar at eventuelle underskudd må dekkes av det offentlige, slik at profitten må multipliseres med faktoren 1 + $\lambda$, der $\lambda$ er skyggeprisen på offentlige midler. Problemet å maksimere samfunnets velferd $W$ kan da skrives:

$$\max_{G,p,f,N} W = \int_G^\infty x(y)dy + (1+\lambda)\left[\left(p - B(s)\varphi^{-1}\frac{m}{s}\right)x(G) - A(s)\frac{a}{s}Nf\right]$$

(17) gitt

$$p + E(s) + Vf^{-1} + LN^{-1} = G \qquad (\mu)$$

Førsteordensbetingelsene for samfunnsøkonomisk maksimum er:

(18) $$\frac{\partial L}{\partial G} = -x + (1+\lambda)\left(p - B(s)\varphi^{-1}\frac{m}{s}\right)\frac{\partial x}{\partial G} + \mu = 0$$

(19) $$\frac{\partial L}{\partial p} = (1+\lambda)x - \mu = 0$$

(20) $$\frac{\partial L}{\partial f} = -(1+\lambda)A(s)\frac{a}{s}N + \mu Vf^{-2} = 0$$

(21) $$\frac{\partial L}{\partial N} = -(1+\lambda)A(s)\frac{a}{s}f + \mu LN^{-2} = 0$$

Av (19) får vi (1 + $\lambda$)$x$ =$\mu$. Med denne verdien innsatt kan (18) skrives

(22) $$\frac{p - B(s)\varphi^{-1}\frac{m}{s}}{p} = -\frac{\lambda}{1+\lambda}\frac{1}{El_p x}$$

Siden (1 + $\lambda$)$\mu$ = $x$ blir vilkårene (20) og (21) identiske med (12) og (13). Formlene (15) og (16) for optimal frekvens, flatedekning og kjøretøykapasitet gjelder altså også dersom vi maksimerer samfunnsnytte.

## 3.3 Sammenlikning av profittmaksimering og maksimering av samfunnsnytte

Ved sammenlikning av likning (22) med likning (14) ser vi at profittmarginen i samfunnsøkonomisk optimum er bare 1/6 av profittmarginen ved profittmaksimering (gitt at $\lambda$ = 0,2). Feilen med profittmaksimering er at det tas for høye priser. Dette gir mindre etterspørsel enn det som er optimalt, og dermed *indirekte* også dårligere kvalitet (frekvens og flatedekning). Feilen vil kunne rettes gjennom regulering i form av et pristak.

Modellen vår inneholder ikke alle former for god eller dårlig kvalitet som trafikantene bryr seg om. For eksempel tar den ikke opp trengsel ombord og pålitelighet. Men for de kvalitetsaspektene den tar opp – frekvens og flatedekning – er det ingen grunn til at et profittmaksimerende selskap systematisk skulle levere for dårlig kvalitet dersom ikke etterspørselen hadde blitt redusert gjennom for høye priser. Grunnen til det er at vi har antatt at etterspørselen er en funksjon av generaliserte kostnader, dvs. at alle trafikanter opplever og verdsetter kvalitetsaspektene på samme måte. Dersom dette ikke hadde vært tilfelle, ville monopolselskapet i modellen kunne hatt en tendens til å ignorere flesteparten av kundene når det fastla kvaliteten, og bare bry seg om den marginale betalingsvilligheten for kvalitet hos den marginale kunden (Tirole 1988 avsnitt 2.2.1).

Verken monopolprisen i (14) eller den samfunnsøkonomisk riktige prisen i (22) behøver å innebære overskudd. Det som i alle fall vil dekkes av begge formene for prising er den etterspørselsavhengige delen av kostnadene, $C_2$, mens $C_1$ dekkes i varierende grad. Hvis $\lambda$ = 0 vil optimal pris være like marginalkostnaden, og samfunnsøkonomisk riktig prising vil bare dekke $C_2$. I det tilfellet vil brukernytten være høy og kostnadsdekningen lav i optimumspunktet. Dersom underskudd er mer kostbart (større $\lambda$), vil det være optimalt at brukernytten er mindre og at en større eller mindre del av $C_1$ blir dekket av billettinntektene. Monopolprising tilsvarer at $\lambda$ er uendelig stor. I det tilfellet er det bare profitten som teller. Det er likevel et åpent spørsmål om positiv profitt er mulig. Brukernytten er lavest i dette tilfellet. Om den da er så liten at den er mindre enn den udekkede delen av kostnadene, vil hele kollektivsystemet være samfunnsøkonomisk ulønnsomt.

## 3.4 Optimering med et fast tilskuddsbeløp

I noen tilfeller er det gitt et fast tilskudd, og oppgava er å få mest mulig samfunnsøkonomi ut av de tilgjengelige midlene. Hvis budsjettbetingelsen er bindende, vil den ha en skyggepris, og det er lett å vise at skyggeprisen vil spille akkurat samme rolle som skyggeprisen på offentlige midler, $\lambda$, i problemet (17). Den eneste forskjellen er at mens $\lambda$ i (17) har en gitt verdi (nemlig 0,2 i norske nyttekostnadsanalyser, i følge Finansdepartementets retningslinjer), vil skyggeprisen på budsjettbetingelsen være større jo mindre tilskuddsbeløpet er.

Problemstillingen (17) og løsningen (formlene (15), (16) og (22)) er derfor generell i den forstand at den dekker profittmaksimering ($\lambda = \infty$), samfunnsøkonomisk optimering under en budsjettbetingelse ($0.2 \le \lambda < \infty$), samfunnsøkonomisk optimering med en skyggepris på offentlige midler ($\lambda$ = 0.2) og samfunnsøkonomisk optimering uten skyggepris på offentlige midler ($\lambda$ = 0). Det er bemerkelsesverdig at forskjellen mellom tilfellene utelukkende stammer fra prissettingen i (22) og at frekvens og flatedekning vil bli bestemt på samme måte i alle tilfellene.

En praktisk metode for å løse problemet med en budsjettbetingelse kan være å gjette på en $\lambda$, beregne (15), (16) og (22) med denne $\lambda$, sette inn i $\Pi$ og kontrollere om $-\Pi$ er større eller mindre enn det gitte tilskuddsbeløpet. Hvis –Π er for stor, økes λ, og hvis den er for liten, reduseres $\lambda$. Prosessen gjentas inntil –Π er tilstrekkelig nær tilskuddsbeløpet, eller til det blir klart at sjøl ikke den største $\lambda$ kan løse problemet. I det sistnevnte tilfellet må man enten legge ned eller få økt tilskuddene.

## 3.5 Stordriftsfordeler

”Familien” av tilbudsmodeller som genereres av problemet (17) vil alle sammen innebære at det eksisterer stordriftsfordeler i produksjonen av kollektivtransport. Det vil si at jo større etterspørselsgrunnlaget er, jo billigere vil det være å produsere en gjennomsnittsreise. Det gjelder både for kollektivselskapet og den enkelte trafikanten.

Kollektivselskapets gjennomsnittskostnader pr. reise framkommer ved å dele $C$ i likning (7) på $x$ og bruke optimale verdier fra (15) for $f$ og $N$. Vi ser da at $C_2/x$ er konstant og $C_1/x$ er proporsjonal med $x^{-1/3}$, altså avtakende i $x$. De to siste leddene i trafikantens generaliserte kostnad (likning 8) er også åpenbart proporsjonale med $x^{-1/3}$, så vi har et tilsvarende resultat for trafikantens kostnader.

## 3.6 *f* eller *N* er gitt

Hittil i dette kapitlet har vi utviklet en familie av enkle tilbudsmodeller som løsninger av problem (17). Det gjenstår å komplettere dette med tilfellene der det eksisterer begrensninger på det frie valget av variablene. I dette avsnittet ser vi på tilfellet der $f$ eller $N$ er gitt. I neste kapittel behandler vi tilfeller der det finnes mer kompliserte skranker for det frie valget av variable.

Frekvensen kan for eksempel være gitt i løsningspunktet hvis fritt valgt optimal $f$ er større enn største mulige $f$. Vi vil da ha $f = \overline{f}$. Som nevnt kan en slik øvre grense være aktuelt for T-banen, og generelt for skinnegående transport, der minste tidsavstand mellom avgangene er bestemt av signalsystemet. $N$ kan være gitt hvis det er forbundet med store investeringskostnader å legge om linjene. $c$ kan være gitt hvis det bare finns én aktuell kjøretøytype, eller hvis teknologi eller infrastruktur setter grenser for størrelsen på kjøretøyene.

Hvis $c$ er fri, men enten $f$ eller $N$ er gitt, er det bare å se bort fra enten likning (20) eller (21).

Hvis $N$ er gitt har vi av (20):

(23) $$f(N) = \sqrt{\frac{sV}{aAN}} \cdot \sqrt{x}$$

Hvis $f$ er gitt har vi av (21):

(24) $$N(f) = \sqrt{\frac{sL}{aAf}} \cdot \sqrt{x}$$

Det framgår av (23) at elastisiteten av $f$ med hensyn på $x$ er ½. Det samme er elastisiteten av $N$ med hensyn på $x$ ifølge (24). Elastisiteten av $c$ med hensyn på $x$ er i begge tilfeller ½. Det eksisterer stordriftsfordeler i begge tilfeller.

# 4 Tilbudsmodeller med kapasitetsbeskrankninger

Vi innfører nå to nye bibetingelser i maksimeringsproblemet (17), nemlig en øvre grense for *Nf* og en øvre grense for *c*. En øvre grense for *Nf* kan være gitt hvis flere linjer deler samme begrensende flaskehals (spor, terminal). Anta at det maksimale antall avganger pr. time gjennom en felles flaskehals er *K*. Vi har altså $Nf \leq K$. Samtidig har vi en øvre grense for kjøretøykapasiteten, altså $c \leq c_{\max}$. Bruker vi (4) og ordner, kan vi samlet uttrykke de nye bibetingelsene slik:

$$\varphi^{-1}\frac{m}{a}x \leq c_{\max}Nf \leq c_{\max}K$$

Problemstillingen blir da:

$$\max_{G,p,f,N} W = \int_G^{\infty} x(y)\,dy + (1+\lambda)\left[\left(p - B(s)\varphi^{-1}\frac{m}{s}\right)x(G) - A(s)\frac{a}{s}Nf\right]$$

gitt

(25) B1 $p + E(s) + Vf^{-1} + LN^{-1} = G \qquad (\mu_1)$

B2 $Nf \leq K \qquad (\mu_2)$

B3 $\varphi^{-1}\frac{m}{a}x - c_{\max}Nf \leq 0 \qquad (\mu_3)$

Kuhn-Tuckerbetingelsene for maksimum er:

(26) $$\frac{\partial L}{\partial G} = -x + (1+\lambda)\left(p - B(s)\varphi^{-1}\frac{m}{s}\right)\frac{\partial x}{\partial G} + \mu_1 - \mu_3\varphi^{-1}\frac{m}{a}\frac{\partial x}{\partial G} = 0$$

(27) $$\frac{\partial L}{\partial p} = (1+\lambda)x - \mu_1 = 0$$

(28) $$\frac{\partial L}{\partial f} = -(1+\lambda)A(s)\frac{a}{s}N + \mu_1 Vf^{-2} - (\mu_2 - c_{\max}\mu_3)N = 0$$

(29) $$\frac{\partial L}{\partial N} = -(1+\lambda)A(s)\frac{a}{s}f + \mu_1 LN^{-2} - (\mu_2 - c_{\max}\mu_3)f = 0$$

(30) $$\mu_2 \geq 0,\ \mu_3 \geq 0$$

Bibetingelsene B2 og B3 kan være bindende eller ikke. Hvis de ikke er bindende, er den tilsvarende Lagrangeparameteren lik 0. Det finns fire mulige tilfeller: Ingen av dem er bindende, bare B2 er bindende, bare B3 er bindende og begge er bindende. Hvert av tilfellene kan gi opphav til en løsningskandidat, og vi går nå videre og behandler dem etter tur.

**Verken B2 eller B3 er bindende**

Dette gir oss den løsningskandidaten som vi allerede har behandlet som løsningen av problem (17).

**Bare B2 er bindende**

Siden $\mu_3 = 0$, gir (26) opphav den til samme prislikningen som i problem (17). Altså gjelder (22). Av B2 med likhetstegn og likning (28) og (29) kan vi utlede:

$$(31)\quad f=\sqrt{\frac{KV}{L}},\quad N=\sqrt{\frac{KL}{V}},\quad c=\varphi^{-1}\frac{m}{a}\frac{x}{K},\quad \mu_2=(1+\lambda)\frac{1}{K}\left[\left(\frac{VL}{K}\right)^{1/2}x-C_1\right]$$

Merk at $C_1$ er en konstant i dette tilfellet siden $Nf = K$. Det kan virke rart at optimal $f$ og $N$ bare avhenger av $K$ og parametere som inngår i generaliserte kostnader, ikke av parametere i kollektivselskapets kostnader. Men løsningskandidaten er ikke så rar som det kan virke: Siden kollektivselskapets kostnader i dette tilfellet består av en konstant pluss et ledd som er proporsjonalt med $x$, innebærer løsningen at $f$ og $N$ settes slik at generaliserte kostnader minimeres.

Et vilkår for at dette skal være en løsningskandidat, er imidlertid at $\mu_2$ er positiv. Parametrene i problemet bestemmer om det er mulig.

**Bare B3 er bindende**

Prislikningen blir

$$(32)\qquad \frac{p-\varphi^{-1}\frac{m}{a}\left(B(s)\frac{a}{s}+\frac{\mu_3}{1+\lambda}\right)}{p}=-\frac{\lambda}{1+\lambda}\frac{1}{El_p x}$$

Prisen skal altså skrus opp for å holde etterspørselen innenfor maksimumskapasiteten pr. avgang.

Av B3 med likhetstegn og likning (28) og (29) kan vi utlede:

$$(33)\qquad \begin{aligned} &f=\left(\varphi c_{\max}\frac{a}{m}\right)^{-1/2}\left(\frac{V}{L}\right)^{1/2}x^{1/2},\quad N=\left(\varphi c_{\max}\frac{a}{m}\right)^{-1/2}\left(\frac{L}{V}\right)^{1/2}x^{1/2},\quad c=c_{\max}\\ &\mu_3=\frac{1+\lambda}{c_{\max}}\left[A(s)\frac{a}{s}-\left(\varphi c_{\max}\frac{a}{m}\right)^{3/2}(VL)^{1/2}x^{-1/2}\right]=\frac{1+\lambda}{c_{\max}Nf}\left[C_1-\left(\frac{VL}{Nf}\right)^{1/2}x\right]\end{aligned}$$

Et vilkår for at dette skal være en løsningskandidat, er at $\mu_3$ er positiv. Parametrene i problemet bestemmer om det er mulig.

**Både B2 og B3 bindende**

I dette tilfellet finner vi:

$$(34)\quad f=\sqrt{\frac{KV}{L}},\quad N=\sqrt{\frac{KL}{V}},\quad c=c_{\max},\quad \mu_2=(1+\lambda)\frac{1}{K}\left[\left(\frac{VL}{K}\right)^{1/2}x-C_1\right]+c_{\max}\mu_3$$

Med disse verdiene for $f$ og $N$ innsatt finner vi $\mu_3$ og $p$ av (32) og B3. (Noen eksplisitt løsning er ikke mulig uten å spesifisere etterspørselsfunksjonen.) Til slutt kan vi sette inn den verdien vi finner for $\mu_3$ i utrykket for $\mu_2$ i (34).

## 4.1 Vilkår som må gjelde for at den enkelte løsningskandidaten skal eksistere

Vi har rapportert løsningene for Lagrangeparametrene og understreket at $\mu_2$ og $\mu_3$ må være positive. Bruker vi de optimale verdiene for $N$ og $f$ i hvert tilfelle og setter dem inn i ulikhetene som gjelder på grunn av kravet til Lagrangeparametrene eller fordi noen av bibetingelsene ikke er bindende, får vi følgende vilkår for hver av løsningskandidatene:

*Verken B2 eller B3 bindende:*

$$x < \min\left( A\frac{a}{s}K\left(\frac{K}{VL}\right)^{1/2}, \left(\varphi c_{\max}\frac{a}{m}\right)^3\left(A\frac{a}{s}\right)^{-2}VL \right)$$

Det er altså bare under denne betingelsen at bibetingelsene er uten betydning.

*Bare B2 bindende:*

$$A\frac{a}{s}K\left(\frac{K}{VL}\right)^{1/2} \leq x < \varphi c_{\max}\frac{a}{m}K$$

*Bare B3 bindende:*

$$\left(\varphi c_{\max}\frac{a}{m}\right)^3\left(A\frac{a}{s}\right)^{-2}VL \leq x < \varphi c_{\max}\frac{a}{m}K$$

*Både B2 og B3 bindende:*

$$x = \varphi c_{\max}\frac{a}{m}K$$

$$p\left(1+\frac{\lambda}{1+\lambda}\frac{1}{El_p x}\right) - \varphi^{-1}B\frac{m}{s} \geq 0$$

$$\left[\left(\frac{VL}{K}\right)^{1/2}\varphi c_{\max}\frac{a}{m} - A\frac{a}{s}\right] + \varphi c_{\max}\frac{a}{m}\left[p\left(1+\frac{\lambda}{1+\lambda}\frac{1}{El_p x}\right) - \varphi^{-1}B\frac{m}{s}\right] \geq 0$$

Ser vi på de fire tilfellene samlet, kan vi først merke oss at de til sammen dekker hele det mulige området for *x*, nemlig $0 < x \leq \varphi c_{\max} a m^{-1} K$ . Tenker vi oss at *x* først er liten, men vokser med en eksogent gitt faktor, har vi først tilfellet uten bindende restriksjoner som eneste løsningskandidat. Når *x* vokser, inntreffer det før eller siden at en av de to restriksjonene, B2 og B3, blir bindende. Hvilken av dem det er, avgjøres av parametrene i problemet. Hvis altså vilkåret for at bare B2 skal være bindende er oppfylt for en verdi av *x*, vil vilkåret for at bare B3 skal være bindende, aldri inntreffe, og omvendt. I et konkret problem vil enten bare tilfelle en, to eller fire, eller bare tilfelle en, tre og fire, være aktuelle. Parametrene avgjør om det er det andre eller det tredje tilfellet som kan utelukkes a priori. B2 og B3 gjelder samtidig bare i ett eneste punkt, $x = \varphi c_{\max} a m^{-1} K$ . Etterspørsel over dette nivået er ikke forenlig med en statisk modell.

Hvis man for eksempel har funnet ut hvilket av de to uttrykkene som er minst i vilkåret for at verken B2 eller B3 skal være bindende i løsningen, kan man straks se om det er begrensningen i kjøretøystørrelsen eller kapasiteten gjennom flaskehalsen som vil bli problemet når etterspørselen stiger. Slik kan man unngå å sette inn tiltak for å oppheve en begrensing som ikke kommer til å bli bindende før systemet er fullt utnyttet, dvs. før *begge* begrensninger er bindende.

## 4.2 *N* = 1

Ofte kan det, som nevnt, være urealistisk å anta at det er mulig å endre flatedekningen uten større investeringer. Spesielt gjelder det skinnegående systemer. I det tilfellet kan vi like gjerne betrakte linjene enkeltvis, dvs. sette *N* = 1. Siden en slik enkeltlinje vil være konstruert for et begrenset antall avganger og et begrenset utvalg av kjøretøytyper, er det av interesse å behandle problemstillingen for dette tilfellet for seg.

Problemet er:

$$\max_{G,p,f} W = \int_G^\infty x(y)dy + (1+\lambda)\left[\left(p - B(s)\varphi^{-1}\frac{m}{s}\right)x(G) - A(s)\frac{a}{s}f\right]$$

gitt

(35) B1 $p + E(s) + Vf^{-1} + LN^{-1} = G$ $(\mu_1)$

B2 $f \le \bar{f}$ $(\mu_2)$

B3 $\varphi^{-1}\frac{m}{a}x - c_{\max} f \le 0$ $(\mu_3)$

Kuhn-Tuckerbetingelsene for maksimum er:

(36) $$\frac{\partial L}{\partial G} = -x + (1+\lambda)\left(p - B(s)\varphi^{-1}\frac{m}{s}\right)\frac{\partial x}{\partial G} + \mu_1 - \mu_3\varphi^{-1}\frac{m}{a}\frac{\partial x}{\partial G} = 0$$

(37) $$\frac{\partial L}{\partial p} = (1+\lambda)x - \mu_1 = 0$$

(38) $$\frac{\partial L}{\partial f} = -(1+\lambda)A(s)\frac{a}{s} + \mu_1 V f^{-2} - (\mu_2 - c_{\max}\mu_3) = 0$$

Som før har vi fire tilfeller:

*Verken B2 eller B3 bindende:*

Løsningskandidaten i dette tilfellet er (implisitt) gitt ved (22) og

(39) $$f = \sqrt{\frac{sV}{aA}} \cdot \sqrt{x}$$

Av $f < \bar{f}$ og $c < c_{\max}$ kan vi bl.a. utlede at i løsningspunktet må ventetidskostnadene være *mindre* enn $C_1$, Vi kan også utlede at hvis dette tilfellet skal gi en løsningskandidat, må følgene ulikhet gjelde i løsningspunktet:

$$x < \min\left(A\frac{a}{s}\frac{\bar{f}^2}{V}, \left(\varphi c_{\max}\frac{a}{m}\right)^2\left(A\frac{a}{s}\right)^{-1}V\right)$$

*Bare B2 bindende:*

Løsningskandidaten i dette tilfellet er (implisitt) gitt ved (22) og $f = \bar{f}$. For den Lagrange-multiplikatoren som ikke nødvendigvis er 0, har vi dessuten

$$\mu_2 = (1+\lambda)\left[Vx\bar{f}^{-2} - A\frac{a}{s}\right] \ge 0$$

Av dette følger at i dette tilfellet er ventetidskostnadene *større* enn $C_1$ i løsningspunktet. Sammen med $c < c_{\min}$ gir det ulikhetene

$$A\frac{a}{s}\frac{\bar{f}^2}{V} \le x < \left(\varphi c_{\max}\frac{a}{m}\right)\bar{f}$$

som må gjelde hvis dette tilfellet skal gi en løsningskandidat.

*Bare B3 bindende:*

I dette tilfelle gjelder

(40) $$x = \left( \varphi c_{\max} \frac{a}{m} \right) f < \left( \varphi c_{\max} \frac{a}{m} \right) \bar{f}$$

Av (38) finner vi

$$\mu_3 = \frac{1+\lambda}{c_{\max}} \left( A \frac{a}{s} - Vxf^{-2} \right) \geq 0$$

I dette tilfellet er altså ventetidskostnadene igjen *mindre* enn $C_1$. Setter vi inn for $\mu_3$ i (36) og bruker (40), får vi

(41) $$\frac{\left( p + Vf^{-1} \right) - \varphi^{-1} \frac{m}{s} \left( B + \frac{A}{c_{\max}} \right)}{p} = -\frac{\lambda}{1+\lambda} \frac{1}{El_p x}$$

(40) og (41) fastlegger nå $p$ og $f$. I henhold til (41) settes prisen så høyt at etterspørselen ikke overstiger kapasiteten pr. avgang ved optimalt valg av $f$. Det er interessant at vi skal ta hensyn til alle kostnadselementer, både for selskapet og for brukerne, når prisen fastsettes.

Av $\mu_3 \geq 0$ og ulikheten i (40) kan vi utlede følgende ulikheter:

$$\left( \varphi c_{\max} \frac{a}{m} \right)^2 \left( A \frac{a}{s} \right)^{-1} V \leq x < \left( \varphi c_{\max} \frac{a}{m} \right) \bar{f}$$

*Både B2 og B3 bindende:*

I dette tilfellet gjelder $f = \bar{f}$ og

(42) $$x = \left( \varphi c_{\max} \frac{a}{m} \right) \bar{f}$$

(42) bestemmer $p$. Dermed vil (36) bestemme $\mu_3$, og tilslutt vil vi da finne $\mu_2$ av (38). Vi finner:

$$\mu_3 = (1+\lambda) \varphi \frac{a}{m} \left[ p \left( 1 + \frac{\lambda}{1+\lambda} \frac{1}{El_p x} \right) - \varphi^{-1} B \frac{m}{s} \right] \geq 0$$

$$\mu_2 = (1+\lambda)\, \bar{f}^{-1} \left[ Vx\bar{f}^{-1} + px \left( 1 + \frac{\lambda}{1+\lambda} \frac{1}{El_p x} \right) - B\varphi^{-1} \frac{m}{s} x - A \frac{a}{s} \bar{f} \right] \geq 0$$

De to ulikhetene som framkommer av positivitetskravet på Lagrangemultiplikatorene, har ingen likhet med vilkårene som vi avledet i de andre tilfellene. Det som derimot kan sammenliknes med de andre vilkårene, er (42). Vi ser at vilkårene for B2 og B3 ikke omfatter det øvre endepunktet for mulighetsområdet for $x$. Det er først når frekvens og kapasitet pr. avgang er maksimalt utnyttet, at $x$ antar denne øvre verdien.

Men de to kravene til Lagrangemultiplikatorene er ikke uten interesse. Uttrykket for $\mu_2$ inneholder hele kollektivselskapets profitt, pluss ventetidskostnadene, minus et ledd som er proporsjonalt med billettinntektene og omvendt proporsjonalt med priselastisiteten. Positivitetskravet på $\mu_2$ sier altså at i dette tilfellet må kollektivselskapet ikke ha for stort underskudd dersom vi krever bedrifts- eller samfunnsøkonomisk optimalitet. Hvis $\lambda = 0$ kan underskuddet ikke være større enn trafikantenes ventetidskostnader, og hvis $\lambda > 0$ skal underskuddet være enda mindre. Så vidt jeg veit er dette et resultat som ikke har vært rapportert i litteraturen tidligere. I mangel av muligheter for å utvide drifta, skal altså kollektivselskapet øke prisen for at ikke avgangene skal bli overfylt. Eller omvendt: Nytten av investeringer som gjør det mulig å utvide drifta, omfatter også nytten av lavere optimale priser.

Ser vi på positivitetskravet på $\mu_3$, oppdager vi dessuten at når $\mu_3$ er relativt liten (dvs. når tendensen til for stor etterspørsel er moderat), er det avgjørende for om $\mu_2$ er positiv eller ikke om ventetidskostnadene er større enn $C_1$ eller ikke. Det er naturlig: Om ventetidskostnadene er store ved maksimal frekvens, er det en stor gevinst ved å kunne øke øvre grense for frekvensen, og dermed en stor Lagrangemultiplikator knyttet til denne skranken.

## 4.3 Nærmere om løsningen

Tenker vi oss at $x$ først er liten, men vokser med en eksogent gitt faktor, har vi først tilfellet uten bindende restriksjoner som eneste løsningskandidat. Når $x$ vokser, inntreffer det før eller siden at en av de to restriksjonene, B2 og B3, blir bindende. Ulikhetene for de to tilfellene, B2 bindende og B3 bindende, viser at om ventetidskostnadene er større enn $C_1$ ved høyeste frekvens og kapasitet pr. avgang, er det restriksjonen B2 som først blir bindende, og omvendt. Kravene til at Lagrangemultiplikatoren skal være positiv i hvert av de to tilfellene, viser at om for eksempel B2 er den restriksjonen som først blir oppfylt når $x$ vokser, kan vi ikke seinere få et skift til at bare B3 er oppfylt – forholdet mellom ventetidskostnadene og $C_1$ kan jo ikke endre seg. Når $x$ blir tilstrekkelig stor, vil imidlertid begge restriksjonene bli bindende.

Dermed har vi vist at om ventetidskostnadene er større enn $C_1$ ved høyeste frekvens og kapasitet pr. avgang, er det første, andre og fjerde tilfelle som kan oppstå, og om det er omvendt, er det første, tredje og fjerde tilfelle som kan oppstå. For hver av de to mulighetene er det størrelsen på $x$ som avgjør hvilket av de tre tilfellene som er eneste mulige løsning.

Nå er ikke $x$ eksogent gitt, men formlene viser at når $x$ vokser med en eksogen faktor, vil prisen gå ned og frekvensen opp, slik at også generaliserte kostnader bidra til veksten. Derfor er det definert entydige punkter der ett tilfelle går over i et annet. I disse punktene vil ingen av våre variable gjøre et hopp. Så lenge parametrene som inngår i ventetidskostnadene og $C_1$ ikke endrer seg, eksisterer det da en kontinuerlig og tiltakende kostnadsfunksjon $C^*(x)$ som er definert over hele det mulige området for $x$, fra 0 til $\varphi c_{\max} a m^{-1} \bar{f}$ .

# 5 Nytten av investeringer som kan øke frekvens og kapasitet pr. avgang

Vi ønsker å finne et enkelt uttrykk for samfunnsnytten av en liten økning i den maksimalt tillatte frekvensen eller den maksimale kapasiteten pr. avgang.

De matematiske forutsetningene er tilstede for å bruke en setning fra Sydsæther (1990, kapittel 4.15) som sier at den deriverte av verdifunksjonen (den optimerte målfunksjonen) med hensyn på en skranke (her: $\bar{f}$ eller $c_{\max}$) er lik Lagrangeparameteren tilknyttet skranken. Men det finnes et lite problem: $c_{\max}$ står ikke aleine på høyresida av ulikheten B3, slik den skal for at vi kan bruke setningen. Det viser seg at om vi flytter $c_{\max}$ slik at den står aleine på høyresida, blir den relevante Lagrangemultiplikatoren i det generelle problemet (25) – la oss kalle den $\gamma$ – lik $\gamma = Nf\mu_3$. I problemet (35) blir den $\gamma = f\mu_3$.

La oss kalle de små endringene $d\bar{f}$ og $dc_{\max}$, og samfunnsnytten $dW^*$. Formlene nedenfor forutsetter at tilpasningen i utgangspunktet er nær optimal. De er dessuten førsteordens tilnærminger til den virkelige nytten, og er bare brukbare for små endringer. Større endringer, eller endringer i en situasjon der tilpasningen ikke er optimal, krever at problemet løses både for en nullsituasjon og situasjonen etter tiltaket, og differansen beregnes.

## 5.1 Både *N* og *f* kan velges

*Ingen bindende restriksjoner:*

I dette tilfellet er $dW^* = 0$, idet optimal tilpasning er mulig uten nye investeringer. (I praksis kan det hende at det likevel er en nytte i form av økt pålitelighet når togene ikke behøver å pakkes maksimalt tett, eller i form av bedre muligheter for sitteplass sjøl i de mest belastede periodene. Men dette og andre eventuelle fordeler ved investeringen ligger utafor vår modell.)

*Bare B2 bindende:*

For investeringer som muliggjør økt kapasitet har vi også her at $dW^* = 0$. For investeringer som muliggjør økt frekvens har vi:

(43) $$dW^* = \frac{\partial W^*}{\partial \bar{f}} d\bar{f} = \mu_2^* d\bar{f} = (1+\lambda)\frac{1}{K}\left[\left(\frac{VL}{K}\right)^{1/2} x^* - A(s)\frac{a}{s}K\right] d\bar{f}$$

*Bare B3 bindende:*

For investeringer som muliggjør økt frekvens er $dW^* = 0$. For investeringer som muliggjør økt kapasitet har vi:

(44) $$\begin{aligned} dW^* &= \frac{\partial W^*}{\partial c_{\max}} dc_{\max} = N^* f^* \mu_3^* dc_{\max} \\ &= \frac{1+\lambda}{c_{\max}}\left[A(s)\frac{a}{s}\left(\varphi c_{\max}\frac{a}{m}\right)^{-1} x^* - (VL)^{1/2}\left(\varphi c_{\max}\frac{a}{m}\right)^{1/2} x^{*1/2}\right] dc_{\max} \end{aligned}$$

*Begge restriksjoner bindende:*

$$dW^* = \frac{\partial W^*}{\partial \bar{f}} d\bar{f} + \frac{\partial W^*}{\partial c_{\max}} dc_{\max} = \mu_2^* d\bar{f} + N^* f^* \mu_3^* dc_{\max} =$$

$$(45)\quad (1+\lambda)\left[\left(\frac{VL}{K}\right)^{1/2}\left(\varphi c_{\max}\frac{a}{m}\right) - A(s)\frac{a}{s} + \left(\varphi c_{\max}\frac{a}{m}\right)K\left(p\left(1+\frac{1}{1+\lambda}\frac{1}{El_p x}\right) - B(s)\varphi^{-1}\frac{m}{s}\right)\right]d\bar{f}$$

$$+(1+\lambda)\varphi\frac{a}{m}K\left(p\left(1+\frac{1}{1+\lambda}\frac{1}{El_p x}\right) - B(s)\varphi^{-1}\frac{m}{s}\right)dc_{\max}$$

## 5.2 *N* = 1

*Ingen bindende restriksjoner:*

I dette tilfellet er $dW^* = 0$, idet optimal tilpasning er mulig uten nye investeringer.

*Bare B2 bindende:*

For investeringer som muliggjør økt kapasitet har vi også her at $dW^* = 0$. For investeringer som muliggjør økt frekvens har vi:

$$(46)\qquad dW^* = \frac{\partial W^*}{\partial \bar{f}} d\bar{f} = \mu_2^* d\bar{f} = (1+\lambda)\left[Vx^* \bar{f}^{-2} - A\frac{a}{s}\right]d\bar{f}$$

*Bare B3 bindende:*

For investeringer som muliggjør økt frekvens er $dW^* = 0$. For investeringer som muliggjør økt kapasitet har vi:

$$(47)\qquad \begin{aligned} dW^* &= \frac{\partial W^*}{\partial c_{\max}} dc_{\max} = f^* \mu_3^* dc_{\max} = \frac{1+\lambda}{c_{\max}}\left(A\frac{a}{s}f - Vx^* f^{-1}\right)dc_{\max} \\ &= \frac{1+\lambda}{c_{\max}}\left(A\frac{a}{s}\left(\varphi c_{\max}\frac{a}{m}\right)^{-1} x^* - V\left(\varphi c_{\max}\frac{a}{m}\right)\right)dc_{\max} \end{aligned}$$

*Begge restriksjoner bindende:*

$$dW^* = \frac{\partial W^*}{\partial \bar{f}} d\bar{f} + \frac{\partial W^*}{\partial c_{\max}} dc_{\max} = \mu_2^* d\bar{f} + N^* f^* \mu_3^* dc_{\max} =$$

$$(48)\quad (1+\lambda)\left[V\varphi c_{\max}\frac{a}{m}\bar{f}^{-1} - A(s)\frac{a}{s} + \varphi c_{\max}\frac{a}{m}\bar{f}\left(p\left(1+\frac{1}{1+\lambda}\frac{1}{El_p x}\right) - B(s)\varphi^{-1}\frac{m}{s}\right)\right]d\bar{f}$$

$$+(1+\lambda)\varphi\frac{a}{m}\bar{f}\left(p\left(1+\frac{1}{1+\lambda}\frac{1}{El_p x}\right) - B(s)\varphi^{-1}\frac{m}{s}\right)dc_{\max}$$

# 6 Nytten av etterspørselsskift og framkommelighetstiltak

## 6.1 Framgangsmåte

Vi vil nå finne virkningen av eksogene etterspørselsskift ($d\varepsilon$) og endringer i framkommelighet ($ds$) på kollektivselskapets tilpasning av tilbudet, dvs. den optimale verdien av $f$, $N$ og $c$. Deretter vil vi beregne konsekvensene som en slik mulighet for tilbudsforbedring har på selskapets profitt, eller med andre ord på tilskuddsbehovet, og endelig velferdsvirkningene, dvs. virkningen på $W$. Det sier seg sjøl at virkningene er ulike alt etter som både $f$ og $N$ eller bare f kan varieres, og alt ettersom hvilke bindende skranker som finnes på $f$ og $c$. Vi vil derfor gå gjennom alle variantene av dette.

En kan spørre seg hva som skjer hvis endringene i $f$, $N$ og $c$ medfører at en restriksjon som ikke var bindende i utgangspunktet, blir bindende undervegs til det ønskelige nye tilbudet. Svaret på det er formodentlig at en først må beregne virkningen fram til det punktet hvor en ny restriksjon gjør seg gjeldende, og deretter, under nye forutsetninger, virkningen fra dette punktet til et nytt optimalt tilbud. Ulikhetene som vi har utledet som vilkår for hvilket løsningsalternativ som utgjør den optimale løsningen, vil definere disse overgangspunktene nøyaktig.

Vi forutsetter at selskapet allerede i utgangspunktet gir et optimalt tilbud. Videre antar vi at prisene holdes fast. Det betyr jo at prisene etter etterspørselsskiftet og framkommelighetstiltakene ikke lenger vil være de optimale, men forhåpentligvis har det liten betydning. Om noe, burde det undervurdere profitten og velferdsøkningen.

Den typen av endringer som vi ser på i dette kapitlet, er nettopp emnet for artikkelen til Small som vi nevnte innledningsvis (Small 2004). I sitt hovedtilfelle antar han at økte inntekter til selskapet blir brukt til prissenkninger, som utløser en ny runde av etterspørselsøkning. Vår implisitte antakelse er i stedet at selskapets økte inntekter brukes til å redusere det offentlige tilskuddet. Ingen av delene vil være helt optimalt. Smalls antakelse vil vel gi større indirekte virkninger på etterspørselen, men vår antakelse vil muligens være vel så realistisk.

Vi skal begynne med å utlede noen generelle sammenhenger, før vi anvender dem på hvert av de mange ulike tilfellene.

## 6.2 Generelle sammenhenger

Et skift i etterspørselen og framkommeligheten vil umiddelbart gi nye optimale verdier for $f$, $N$ og $c$. Det er ikke dermed sagt at selskapet straks vil endre sitt tilbud – det kan være mange praktiske hindringer som må ryddes av vegen først. Men et klokt kollektivselskap vil skjønne at disse skiftene har gitt nye muligheter. Et skift i framkommeligheten vil også nokså umiddelbart gi raskere reisetid og dermed lavere generaliserte reisekostnader for trafikantene. Disse umiddelbare og nokså umiddelbare konsekvensene er det vi vil kalle *direkte* virkninger.

I neste omgang vil den økte frekvensen og flatedekningen og den raskere reisetida gi en ny runde av etterspørselsøkning, som så gi selskapet ytterligere grunnlag for tilbudsforbedring. Dette er det vi vil kalle annenordensvirkningene eller de indirekte virkningene. Et veldig forutseende kollektivselskap vil kunne foregripe også de indirekte virkningene når det fastlegger det nye optimale tilbudet.

De optimale $f$, $N$ og $c$ er funksjoner av $s$ og $x$. $c$ er dessuten også en funksjon av $f$ og $N$. Ved relativt små endringer i $s$ og $x$ har vi derfor:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial s}ds,\ \ dN = \frac{\partial N}{\partial x}dx + \frac{\partial N}{\partial s}ds,$$
$$dc = \frac{\partial c}{\partial x}dx + \frac{\partial c}{\partial s}ds + \frac{\partial c}{\partial f}df + \frac{\partial c}{\partial N}dN$$

Videre er *x* en funksjon av *g* og *ε*, så

$$dx = \left[\frac{\partial x}{\partial \varepsilon} + \frac{\partial x}{\partial G}\left(\frac{\partial G}{\partial f}\frac{\partial f}{\partial \varepsilon} + \frac{\partial G}{\partial N}\frac{\partial N}{\partial \varepsilon}\right)\right]d\varepsilon + \frac{\partial x}{\partial G}\left(\frac{\partial G}{\partial s} + \frac{\partial G}{\partial f}\frac{\partial f}{\partial s} + \frac{\partial G}{\partial N}\frac{\partial N}{\partial s}\right)ds$$

Vi setter nå uttrykkene for *df* og *dN* inn i utrykket for *dc*, og setter så utrykket for *dx* inn i uttrykkene for *df*, *dN* og *dc*. Vi omgjør uttrykkene for *df*, *dN* og *dc* til elastisiteter og merker oss at $El_\varepsilon x$ = 1 og $El_s c$ = 0, $El_x c$ = 1, $El_f c = El_N c$ = - 1. Endelig bruker vi følgende definisjoner:

(49)
$$I_\varepsilon \equiv El_G x\left(El_f G \cdot El_\varepsilon f + El_N G \cdot El_\varepsilon N\right)$$
$$I_s \equiv El_G x\left(El_f G \cdot El_s f + El_N G \cdot El_s N\right)$$

Det gir:

(50) $$\frac{df}{f} = El_x f\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \left(El_s f + El_x f \cdot El_G x \cdot El_s G + El_x f \cdot I_s\right)\frac{ds}{s}$$

(51) $$\frac{dN}{N} = El_x N\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \left(El_s N + El_x N \cdot El_G x \cdot El_s G + El_x N \cdot I_s\right)\frac{ds}{s}$$

(52)
$$\frac{dc}{c} = \left(1 - El_x f - El_x N\right)\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon}$$
$$+\left[\left(1 - El_x f - El_x N\right)\left(El_G x \cdot El_S G + I_s\right) - El_s f - El_s N\right]\frac{ds}{s}$$

De leddene i disse formlene som inneholder $I_\varepsilon$ eller $I_s$, utgjør annenordensvirkningene på tilbudsvariablene.

Før vi går til de konkrete tilfellene, er det enda noen elastisiteter som vi kan beregne på generelt grunnlag, nemlig

(53) $$El_s G = -\frac{\omega m}{sG},\ \ El_f G = -\frac{Vf^{-1}}{G},\ \ El_N G = -\frac{LN^{-1}}{G}$$

Vi setter inn verdiene fra (53) i (49) til (52):

(54)
$$I_\varepsilon \equiv \left|El_G x\right|\left(\frac{Vf^{-1}}{G}El_\varepsilon f + \frac{LN^{-1}}{G}El_\varepsilon N\right)$$
$$I_s \equiv \left|El_G x\right|\left(\frac{Vf^{-1}}{G} \cdot El_s f + \frac{LN^{-1}}{G}El_s N\right)$$

(55) $$\frac{df}{f} = El_x f\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \left(El_s f + \left|El_G x\right|\frac{\omega m}{sG}El_x f + El_x f \cdot I_s\right)\frac{ds}{s}$$

(56) $$\frac{dN}{N} = El_x N\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \left(El_s N + \left|El_G x\right|\frac{\omega m}{sG}El_x N + El_x N \cdot I_s\right)\frac{ds}{s}$$

(57)
$$\frac{dc}{c} = \left(1 - El_x f - El_x N\right)\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \left[\left(1 - El_x f - El_x N\right)\left(\left|El_G x\right|\frac{\omega m}{sG} + I_s\right) - El_s f - El_s N\right]\frac{ds}{s}$$

Likning (54)-(57) er vårt utgangspunkt i det følgende for å evaluere virkningen på tilbudsvariablene av eksogene endringer i etterspørselen og i framkommeligheten. Vi ser at de elastisitetene som gjenstår i disse likningene, med ett unntak vil kunne spesifiseres når vi kjenner de optimale *f*, *N* og *c* i hvert konkret tilfelle. Unntaket er elastisiteten av etterspørselen med hensyn på generaliserte kostnader.

## 6.3 Tilbudsendringene i ulike tilfeller

### 6.3.1 *f* og *N* er valgvariable

**Verken B2 eller B3 er bindende**

I dette tilfellet er de optimale *f*, *N* og *c* gitt ved (15) og (16). Vi får:

(58)
$$El_\varepsilon f = El_\varepsilon N = El_\varepsilon c = \tfrac{1}{3}$$
$$El_s f = El_s N = \tfrac{1}{3} \cdot \frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s}$$
$$El_s c = -\tfrac{2}{3} \cdot \frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s}$$

Det gir:

(59)
$$I_\varepsilon \equiv \tfrac{1}{3}\left|El_G x\right|\frac{Vf^{-1} + LN^{-1}}{G}$$
$$I_s \equiv \tfrac{1}{3}\left|El_G x\right|\frac{Vf^{-1} + LN^{-1}}{G} \cdot \frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s}$$

(60)
$$\frac{df}{f} = \frac{dN}{N} = \tfrac{1}{3}\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \tfrac{1}{3}\left(\frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s} + \left|El_G x\right|\frac{\omega m}{sG} + I_s\right)\frac{ds}{s}$$

(61)
$$\frac{dc}{c} = \tfrac{1}{3}\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \tfrac{1}{3}\left(-2\frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s} + \left|El_G x\right|\frac{\omega m}{sG} + I_s\right)\frac{ds}{s}$$

Vi ser at ved eksogen etterspørselsøkning skal *f*, *N* og *c* økes i samme takt, og denne takta er noe, men ikke mye, større enn en tredjedel av etterspørselsøkninga. Ved økt framkommelighet skal *f* og *N* økes i samme takt, og denne takta ser ut til å kunne være noe, men ikke mye, mindre enn takta i framkommelighetsforbedringa. Derimot ser det ut til at kapasiteten pr. avgang i høyden skal holdes konstant.

**Bare B2 er bindende**

Formel (31) viser at optimal *f* og *N* er konstanter, mens *c* er en funksjon av *x* aleine. Dermed har vi av (54)-(57):

(62)
$$\frac{df}{f} = \frac{dN}{N} = 0, \quad \frac{dc}{c} = \frac{d\varepsilon}{\varepsilon} + \left|El_G x\right|\frac{\omega m}{sG} \cdot \frac{ds}{s}$$

Enhver endring i etterspørselen, enten den er eksogen eller framkommer fordi økt fram-kommelighet reduserer reisetida, må i dette tilfelle fullt ut tas opp av endring i kapasiteten pr. avgang.

**Bare B3 bindende**
Fra formel (33) får vi:

(63) $$El_x f = El_x N = \tfrac{1}{2}, \quad El_s f = El_s N = 0$$

Dermed har vi av (54)-(57):

(64) $$I_\varepsilon \equiv \tfrac{1}{2}\left|El_G x\right| \frac{Vf^{-1} + LN^{-1}}{G}$$
$$I_s \equiv 0$$

(65) $$\frac{df}{f} = \tfrac{1}{2}\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \tfrac{1}{2}\left|El_G x\right| \frac{\omega m}{sG} \cdot \frac{ds}{s}$$

(66) $$\frac{dN}{N} = \tfrac{1}{2}\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \tfrac{1}{2}\left|El_G x\right| \frac{\omega m}{sG} \cdot \frac{ds}{s}$$

(67) $$\frac{dc}{c} = 0$$

**Både B2 og B3 bindende**
Alle tilbudsvariablene er nå konstanter, og ingen tilpasning til økt etterspørsel er mulig innafor modellen og våre forutsetninger. Riktignok kan vi endre forutsetningen om konstant pris, og bruke økt pris til å holde etterspørselen innafor kapasiteten. Vi kan også la avgangene være fullere enn vi tillater i modellen.

### 6.3.2 Bare *f* er valgvariabel, *N* = 1

**Verken B2 eller B3 er bindende**
Optimal frekvens er gitt i likning (39). Av den finner vi:

(68) $$El_\varepsilon f = \tfrac{1}{2}$$
$$El_s f = \tfrac{1}{2} \cdot \frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s}$$

Elastisitetene av *N* er naturligvis 0. Dermed gir likning (54)-(57):

(69) $$I_\varepsilon \equiv \tfrac{1}{2}\left|El_G x\right| \frac{Vf^{-1}}{G}$$
$$I_s \equiv \tfrac{1}{2}\left|El_G x\right| \frac{Vf^{-1}}{G} \cdot \frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s}$$

(70) $$\frac{df}{f} = \tfrac{1}{2}\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \tfrac{1}{2}\left(\frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s} + \left|El_G x\right| \frac{\omega m}{sG} + I_s\right)\frac{ds}{s}$$

(71) $$\frac{dc}{c} = \tfrac{1}{2}\left(1 + I_\varepsilon\right)\frac{d\varepsilon}{\varepsilon} + \tfrac{1}{2}\left(-\frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s} + \left|El_G x\right| \frac{\omega m}{sG} + I_s\right)\frac{ds}{s}$$

Som når N er valgvariabel, ser vi at ved eksogen etterspørselsøkning skal *f* og *c* økes i samme takt, og denne takta er noe, men ikke mye, større enn halvparten av takten i etterspørselsøkninga. Ved økt framkommelighet skal også *f* økes, men det ser ut til at kapasiteten pr. avgang skal reduseres.

**Bare B2 er bindende**
I dette tilfellet er det ikke lenger mulig å tilpasse frekvensen, og alle endringer i etterspørselen må takles ved endring i kapasiteten pr. avgang (jfr. likning (4)). Vi får:

(72)
$$El_\varepsilon f = 0, \quad El_s f = 0$$

Elastisitetene av *N* er naturligvis 0. Dermed gir likning (54)-(57):

(73)
$$\begin{aligned} & I_\varepsilon = I_s = 0 \\ & \frac{df}{f} = 0 \\ & \frac{dc}{c} = \frac{d\varepsilon}{\varepsilon} + \left| El_G x \right| \frac{\omega m}{sG} \cdot \frac{ds}{s} \end{aligned}$$

**Bare B3 bindende**
I dette tilfellet er det ikke lenger mulig å tilpasse kapasiteten pr. avgang, og alle endringer i etterspørselen må takles ved endring i frekvensen (jfr. likning (4)). Vi får:

(74)
$$El_\varepsilon f = 1, \quad El_s f = 0$$

Elastisitetene av *N* er naturligvis 0. Dermed gir likning (54)-(57):

(75)
$$\begin{aligned} & I_\varepsilon = \left| El_G x \right| \frac{Vf^{-1}}{G}, \quad I_s = 0 \\ & \frac{df}{f} = \left(1 + I_\varepsilon\right) \frac{d\varepsilon}{\varepsilon} + \left| El_G x \right| \frac{\omega m}{sG} \cdot \frac{ds}{s} \\ & \frac{dc}{c} = 0 \end{aligned}$$

**Både B2 og B3 bindende**
Alle tilbudsvariablene er nå konstanter, og ingen tilpasning til økt etterspørsel er mulig innafor modellen og våre forutsetninger. Riktignok kan vi endre forutsetningen om konstant pris, og bruke økt pris til å holde etterspørselen innafor kapasiteten. Vi kan også la avgangene være fullere enn vi tillater i modellen.

### 6.3.3 Kommentar

Tilbudet kan tilpasses i tre dimensjoner, men på grunn av likning (4) er det bare to frihetsgrader. Dette er et fundamentalt trekk ved produksjonsmulighetene til et kollektivselskap, med mindre etterspørselen er så lav at sjøl minste kjøretøystørrelse gir ledig kapasitet pr. avgang, hvilket vi har sett bort fra. Som vi så i kapittel 4, gir tilpasning i tre dimensjoner opphav til en tredjerotslov for både *f*, *N* og *c* som funksjoner av etterspørselen. Tilpasning i to dimensjoner, *f* og *c*, gir opphav til en kvadratrotlov.

Når en restriksjon blir bindende, reduseres mulighetsrommet. En bindende felles flaskehals gjør faktisk både *f* og *N* til konstanter i optimum, mens *c* blir en lineær funksjon av *x*. Hvis det er kapasiteten pr. avgang som blir den bindende restriksjonen, får vi en kvadratrotlov for de optimale *f* og *N*. I tilfellet der *N* er gitt, blir det som var en tredjerotslov, en kvadratrotlov, og det som var en kvadratrotlov blir en lineær sammenheng mellom etterspørsel og det tilgjengelige virkemiddelet. Hvis to restriksjoner gjør seg gjeldende, har vi ikke lenger noen mulighet til å tilpasse tilbudet til etterspørselen.

Disse sammenhengene gjør at virkningen på *f*, *N* og *c* av eksogene etterspørselsskift og framkommelighetstiltak, blir ulik i de ulike situasjonene.

Vi ser likevel at eksogene etterspørselsskift alltid skal medføre en like stor økning i alle de dimensjonene av tilbudet som ikke er låst fast av restriksjonene. Når det gjelder framkommelighetstiltak er det annerledes. Siden frekvens og flatedekning inngår i generaliserte kostnader, skal økning i disse dimensjonene prioriteres på bekostning av kapasitet pr. avgang. Dette resultatet er vel kjent fra litteraturen, men det framheves ofte at et kollektivselskap som maksimerer profitt, ikke vil følge en slik oppskrift, men prioritere kapasitet pr. avgang. I vår modell er dette ikke riktig – kollektivselskapet vil ha samme interesse som trafikantene av å øke frekvens og flatedekning. Dette skyldes forutsetningen om at alle trafikanter opplever alle elementer av generaliserte kostnader likt.

På grunnlag av våre funn om optimal bruk av *f* og *N* kan vi anta at dersom vår modell hadde inkludert kostnadene ved trengsel om bord i generaliserte kostnader, ville vi fått at *c* skulle øke i samme takt som *f* og *N* også når tiltaket dreier seg om framkommelighet. Det gjenstår å vise.

Hvis en vil bruke noe av dette i praksis, bør en først finne ut om tilbudet man gir i utgangspunktet, er optimalt. Er det ikke det, finnes det jo gratis muligheter til å forbedre samfunnsøkonomien eller profitten. Én årsak til at kollektivtiltak kan komme dårlig ut i en nyttekostnadsanalyse, er nettopp at tilbudet før eller etter tiltaket ikke er optimalt bestemt, slik at det er uklart om endringen fører oss i riktig retning eller ikke.

Først når en veit hva den optimale løsningen vil være i utgangspunktet, kan en bedømme virkningen av etterspørselsøkning og økt framkommelighet.

## 6.4 Virkninger på tilskuddsbehovet

Tilskuddsbehovet er kostnaden minus billettinntektene. Kostnaden er $C = C_1 + C_2$ (jfr. likning (5)). Kaller vi billettinntektene *R*, så er $R = px$. En endring i tilskuddsbehovet er altså $dC - dR$. Bruker vi samme framgangsmåte som i avsnitt 6.2, får vi

$$dR = pdx = px \cdot \frac{dx}{x} = px \cdot \left\{ \left(1 + I_\varepsilon\right) \frac{d\varepsilon}{\varepsilon} + \left( \left|El_G x\right| \frac{\omega m}{sG} + I_s \right) \frac{ds}{s} \right\} \tag{76}$$

$$dC_1 = C_1 \left\{ \left(1 + I_\varepsilon\right) \frac{d\varepsilon}{\varepsilon} + \left( \left|El_G x\right| \frac{\omega m}{sG} + I_s - \frac{r_1/h}{r_1/h + g_1 s} \right) \frac{ds}{s} \right\} \tag{77}$$

$$\begin{aligned} dC_2 &= \\ C_2 &\left\{ \left(El_s f + El_s N\right)\left(1 + I_\varepsilon\right) \frac{d\varepsilon}{\varepsilon} + \left[ \left(El_\varepsilon f + El_\varepsilon N\right) \left( \left|El_G x\right| \frac{\omega m}{sG} + I_s \right) - \frac{r_0/h + w\ell_0}{r_0/h + w\ell_0 + g_0 s} \right] \frac{ds}{s} \right\} \end{aligned} \tag{78}$$

Av dette finner vi

(79)

$$d\left(C-R\right)=\left(C-R\right)\left\{\left(1+I_{\varepsilon}\right)\frac{d\varepsilon}{\varepsilon}+\left(\left|El_{G}x\right|\frac{\omega m}{sG}+I_{s}\right)\frac{ds}{s}\right\}-C_{1}\frac{r_{1}/h}{r_{1}/h+g_{1}s}\cdot\frac{ds}{s}$$

$$-C_{2}\left\{\left(1-El_{\varepsilon}f-El_{\varepsilon}N\right)\left(1+I_{\varepsilon}\right)\frac{d\varepsilon}{\varepsilon}+\left[\left(1-El_{s}f-El_{s}N\right)\left(\left|El_{G}x\right|\frac{\omega m}{sG}+I_{s}\right)-\frac{r_{0}/h+w\ell_{0}}{r_{0}/h+w\ell_{0}+g_{0}s}\right]\frac{ds}{s}\right\}$$

Likning (79) gir endringen i tilskuddsbehovet i kroner. Hvis vi ønsker endringen relativt til *C* eller *C – R*, kan vi dele med dette på begge sider. Vi ser da at første linje i (79) sier at i utgangspunktet øker tilskuddsbehovet i samme takt som etterspørselen (se likning (76)). Linje 2 og 3 sier imidlertid at både økningen i $C_1$ og $C_2$ er mindre enn dette, slik at både kostnaden og underskuddet pr. reise reduseres når trafikken øker.

Den konkrete virkningen vil avhenge av hvilket av tilfellene som vi behandlet i avsnitt 6.3 som foreligger. Vi går ikke gjennom disse tilfellene på nytt: det er bare å sette inn elastisitetene som gjelder i hvert tilfelle, i likning (79).

### 6.4.1 Kommentar

I noen tilfeller, slik som i den nylig gjennomførte konseptvalgsutredningen av Oslopakke 3, brukes avsetninger til drift av kollektivtrafikken som er virkemiddel, uten at det er nærmere spesifisert hva pengene skal brukes til. Det er da vanskelig å bedømme om disse midlene faktisk trengs til det driftsopplegget som er valgt. Det kan tenkes at de brukes til fornuftige tiltak som ikke omfattes av vår modell. I det tilfellet kan man betrakte den resulterende etterspørsels-økningen som eksogent gitt i vår forstand, og bruke likning (79) til å undersøke om de midlene som *ikke* blir anvendt på denne måten, er tilstrekkelig til å finansiere den optimale omlegginga av drifta som bør bli følgen av den økte etterspørselen. Hvis midlene brukes til framkommelig-hetstiltak, kan man på samme måte bruke formelen til å undersøke om den optimale tilpasningen til framkommelighetstiltakene lar seg finansiere.

Som vi pekte på i forrige avsnitt, bør en først finne ut om tilbudet man gir i utgangspunktet, er optimalt. Er det ikke det, finnes det jo gratis muligheter til å forbedre samfunnsøkonomien eller profitten.

## 6.5 Velferdsvirkningen

I forrige avsnitt har vi allerede beregnet en stor del av virkningen på *W*, nemlig det som står i klammeparentesen i (25) eller (35). Det er bare å snu fortegnet i likning (79) og multiplisere med 1 + $\lambda$. Det som gjenstår er trafikantnytten. Kall trafikantnytten *UB*. Med henvisning til likning (49) og likning (79) har vi:

$$dW = dUB - (1+\lambda)d(C-R)$$

$$= \left(\int_G^\infty x(y)dy\right)\frac{d\varepsilon}{\varepsilon} - xdG - (1+\lambda)d(C-R)$$

(80) $$= \left(\int_G^\infty x(y)dy\right)\frac{d\varepsilon}{\varepsilon} - Gx\frac{1}{G}\left[\left(\frac{\partial G}{\partial f}\frac{\partial f}{\partial \varepsilon} + \frac{\partial G}{\partial N}\frac{\partial N}{\partial \varepsilon}\right)d\varepsilon + \left(\frac{\partial G}{\partial s} + \frac{\partial G}{\partial f}\frac{\partial f}{\partial s} + \frac{\partial G}{\partial N}\frac{\partial N}{\partial s}\right)ds\right]$$

$$-(1+\lambda)d(C-R)$$

$$= \left(\int_G^\infty x(y)dy\right)\frac{d\varepsilon}{\varepsilon} - \frac{Gx}{El_G x}\left(I_\varepsilon\frac{d\varepsilon}{\varepsilon} + I_s\frac{ds}{s}\right) - (1+\lambda)d(C-R)$$

Under visse vilkår kan denne formelen anvendes som en justering av det samfunnsøkonomiske regnestykket som gjøres på grunnlag av en transportmodellkjøring.

Anta for eksempel at en bruker transportmodellen til å teste vegprising. Vegprisingen leder til en viss nedgang i biltrafikken, og en del av denne finner vegen til kollektivtransport. I transportmodellen gir det ingen endringer i generaliserte kostnader for kollektivtrafikantene, og dermed ingen nytte for dem. Det gir heller ingen endring i driftsopplegget for kollektivtrafikken. Om nå kollektivselskapets kostnader i transportmodellen er beregnet på grunnlag av det kodede driftsopplegget, skjer det heller ingenting med dem. I det tilfellet kan resultatet fra en beregning etter formel (80) i sin helhet legges til resultatet fra nytteberegning med transportmodellen, som et grovt anslag på den virkelige nytten som oppstår på kollektivsida dersom selskapet utnytter mulighetene som de nye kundene gir.

Det kan tenkes at kostnadsberegningene i transportmodellen ikke er basert på kodede detaljerte data i modellen, men på en empirisk basert aggregert kostnadsfunksjon for hele området. Denne funksjonen vil tilta med etterspørselen. I så fall har en allerede tatt hensyn til kollektivselskapets inntekter og kostnader, så langt som førsteordensvirkningene rekker. Det eneste man ikke har fått med seg, er nytteendringen for brukerne og annenordensvirkningene for selskapet. En kan da bruke de to første leddene i (80) som et tillegg, og kanskje også legge til de leddene i (79) som avhenger av $I_\varepsilon$ og $I_s$.

Endelig er det mulig at tiltaket som testes med modellen, ikke bare omfatter vegprising, men også det nye driftsopplegget på kollektivsida som man antar vil trenges i den forbindelsen. Tilsvarende kan det tenkes at man endrer driftsopplegget i transportmodellen i forbindelse med testing av framkommelighetstiltak. I de tilfellene fanges alle førsteordensvirkningene opp av beregningene i ttransportmodellen. Det gjelder også alle annenordensvirkningene, hvis man har vært forutseende. Da blir det dobbelttelling å ta med noe fra likning (80). Man kan imidlertid bruke likning (79) for å prøve om finansieringsforutsetningene holder.

En må altså tenke seg om før en bruker likning (80) som et tillegg til andre nytteberegninger. Vi behøver vel ikke minne om at en annen forutseting for å bruke likningene er at tilbudet i utgangspunkt er noenlunde optimalt.

# 7 To perioder

Anta nå at antall driftstimer pr. år, $h$, fordeler seg på $h_H$ høybelastningstimer og $h_L$ lavbelastningstimer, og at kollektivtilbudet er forskjellig i de to periodene. Vi definerer andelen av høybelastningstimer $\eta$ og andelen lavbelastningstimer $1 - \eta$ :

(81) $$\eta = \frac{h_H}{h}, \quad 1 - \eta = \frac{h_L}{h}$$

Uansett hvordan vi bruker tilgjengelige virkemidler i de to periodene, vil vi alltid anta at etterspørselen er større i høybelastningsperioden enn i lavbelastningsperioden.

Det er mange ulike måter som kollektivtilbudet kan være forskjellig på i de to periodene. Et vanlig brukt opplegg er å ha et grunntilbud som er det samme hele døgnet og kjøre ekstra-avganger i rush. Ekstraavgangene kan kjøres med større busser enn de som ellers brukes. Et opplegg som bare kan brukes med tog, er å kople på flere vogner eller kople sammen flere togsett i rush. Disse to metodene gjør at kollektivselskapets optimale tilpasning i hver av periodene kan bestemmes helt eller nesten helt uavhengig av tilpasningen i den andre perioden, sjøl om de to periodene må dele på en felles ressurs, nemlig det rullerende materiellet i grunntilbudet/lavbelastningsperioden. Hvis derimot kapasiteten pr. avgang nødvendigvis må være den samme i begge perioder, er tilpasningen i periodene knyttet nærmere sammen, og dette gir et vanskeligere problem.

Prinsipielt kan vi tenke oss at billettprisen er ulik i de to periodene, og dette er også ønskelig ut fra et samfunnsøkonomisk synspunkt. Her vil vi legge til grunn at billettprisen faktisk er ulik i de to periodene, og at kapasiteten pr. avgang kan tilpasses enten ved egen kapasitet på ekstra-avgangene i rush eller når det gjelder antall vogner/antall sammenkoplede togsett i rush. For enkelhets skyld forutsetter vi å anta at kapasiteten pr. avgang er en kontinuerlig variabel.

Nestbestetilfellene med felles kapasitet pr. avgang og felles billettpris vil likevel bli kort kommentert til slutt.

Som tidligere vil målfunksjonen være samfunnsøkonomisk velferd pr. driftstime, men den driftstimen vi nå ser på, er et veid gjennomsnitt av høy- og lavbelastningsperioden med andelen timer av hvert slag som vekter. For enkelhets skyld tar vi her bare for oss tilfellet $N = 1$.

## 7.1 Grunntilbud med ekstraavganger i rush

Vi bruker fotskrift $L$ på lavtrafikkvariable, $H$ på høytrafikkvariable, $G$ på grunntilbudsvariable og $E$ på ekstraavgangsvariable. Etterspørselen i de to periodene er altså $x_H$ og $x_L$, med $x_H > x_L$. Åpenbart har vi $f_G = f_L$ og $f_G + f_E = f_H$. Frekvensen er derfor minst like høy i høybelastnings-perioden som i lavbelastningsperioden, med likhet bare dersom $f_E = 0$. Den teoretiske trans-portkapasiteten pr. time er $c_G f_G = c_L f_L$ i lavtrafikkperioden og $c_G f_G + c_E f_E$ i høybelastnings-perioden. Vi kan definere $c_H$, gjennomsnittskapasiteten pr. rushtidsavgang, ved $c_G f_G + c_E f_E = c_H f_H$.

Kostnaden pr. gjennomsnittlig driftstime, $C$, er kostnaden pr. år delt på antall driftstimer. Vi har:

$$
\begin{aligned}
C &= \left\{\left[{}^{r_{0E}}\!/_{h} + \eta\left(w\ell_0 + g_{0E}s\right)\right] + \left[{}^{r_{1E}}\!/_{h} + \eta g_{1E}s\right]c_E\right\}\frac{af_E}{s} \\
&+ \left\{\left[{}^{r_{0G}}\!/_{h} + \left(w\ell_0 + g_{0G}s\right)\right] + \left[{}^{r_{1G}}\!/_{h} + g_{1G}s\right]c_G\right\}\frac{af_G}{s} \\
&= \eta\left\{\left[{}^{r_{0E}}\!/_{h_H} + \left(w\ell_0 + g_{0E}s\right)\right] + \left[{}^{r_{1E}}\!/_{h_H} + g_{1E}s\right]c_E\right\}\frac{af_E}{s} \\
&+ \left\{\left[{}^{r_{0G}}\!/_{h} + \left(w\ell_0 + g_{0G}s\right)\right] + \left[{}^{r_{1G}}\!/_{h} + g_{1G}s\right]c_G\right\}\frac{af_G}{s} \\
&= \eta\left(A_E + B_E c_E\right)\frac{a}{s} f_E + \left(A_G + B_G c_G\right)\frac{a}{s} f_G \\
&= \frac{a}{s}\left\{\eta A_E\left(f_H - f_L\right) + \eta B_E\left(c_H f_H - c_L f_L\right) + A_G f_L + B_G c_L f_L\right\} \\
&= \frac{a}{s}\left\{\eta A_E f_H + \eta B_E c_H f_H + \left(A_G - \eta A_E\right) f_L + \left(B_G - \eta B_E\right) c_L f_L\right\}
\end{aligned}
\tag{82}
$$

Første linje er kostnaden pr. gjennomsnittstime av ekstratilbudet. Legg merke til at de tids- og kilometeravhengige driftskostnadene er vektet med andelen av året der et ekstratilbud blir gitt. Andre linje er kostnaden pr. gjennomsnittstime av grunntilbudet. Tredje og fjerde linje er lik første og andre, bortsett fra at vi har satt $\eta$ utenfor parentesen i tredje linje. Femte linje definerer $A_E$, $B_E$, $A_G$ og $B_G$ på samme måte som vi definerte $A$ og $B$ i tilfellet med én periode, nemlig som innholdet i de ulike klammeparentesene. Sjette linje bruker sammenhengene $f_G = f_L$ og $f_G + f_E = f_H$, $c_G f_G = c_L f_L$ og $c_G f_G + c_E f_E = c_H f_H$. Syvende linje er en omgruppering av sjette linje.

Etterspørselen i høy- og lavbelastningsperioden er henholdsvis $x_H(G_H)$ og $x_L(G_L)$. Siden det aldri lønner seg å kjøre med ubenyttet ekstrakapasitet over det mest belastede snittet, og siden ekstraavgangene har valgbar kapasitet pr. avgang slik at kapasiteten kan tilpasses i både høy- og lavbelastningsperioden, gjelder disse sammenhengene:

$$
\begin{aligned}
c_L f_L &= \varphi^{-1}\frac{m}{a}x_L \\
c_H f_H &= c_G f_G + c_E f_E = \varphi^{-1}\frac{m}{a}x_H
\end{aligned}
\tag{83}
$$

Setter vi dette inn i $C$ og bruker forkortelsene $A_G - \eta A_E = A_{EG}$ og $B_G - \eta B_E = B_{EG}$, får vi

$$
C = \left\{\eta A_E f_H + A_{EG} f_L\right\}\frac{a}{s} + \left\{\eta B_E \varphi^{-1}\frac{m}{s}x_H + B_{EG}\varphi^{-1}\frac{m}{s}x_L\right\}
\tag{84}
$$

Vi bemerker at $A_{EG}$ og $B_{EG}$ har usikre fortegn, men med mindre andelen av driftsåret med høybelastningstilbud er godt over 50 %, er de høyst sannsynlig positive.

### 7.1.1 Maksimeringsproblemet

Vi skal maksimere velferden i en gjennomsnittlig driftstime. Som i enperiodetilfellet vil profittmaksimering være et spesialtilfelle av dette, nemlig tilfellet der $\lambda = \infty$. Bibetingelsene er at frekvensen $f_H$ i høybelastningsperioden ikke er større enn maksimalt mulig eller tillatt frekvens, og at for begge typer avganger er kapasiteten pr. avgang ($c_G$ og $c_E$) ikke større enn den maksimalt mulige, $c_{max}$. Nå vil vi ikke bruke $c_G$ og $c_E$ som valgvariable, derfor må vi skrive om disse bibetingelsene. $c_G = c_L$, og med likning (83) finner vi lett at vilkåret $c_L \le c_{\max}$ er ekvivalent med

$$\frac{m}{a}x_L - \varphi c_{\max} f_L \leq 0$$

Med definisjonen av $c_H$ og likning (83) finner vi likeledes at siden $f_H \geq f_L$ er $c_E \leq c_{\max}$ ekvivalent med

$$\frac{m}{a}\left(x_H - x_L\right) - \varphi c_{\max}\left(f_H - f_L\right) \leq 0$$

Maksimeringsproblemet blir:

(85)
$$\underset{p_H, p_L, G_H, G_L, f_H, f_L}{Max} W = \eta \int_{G_H}^{\infty} x_H(u)\,du + (1-\eta)\int_{G_L}^{\infty} x_L(u)\,du$$
$$+(1+\lambda)\left\{\eta p_H x_H + (1-\eta) p_L x_L - \left[\eta A_E f_H + A_{EG} f_L\right]\frac{a}{s} - \left[\eta B_E \varphi^{-1}\frac{m}{s}x_H + B_{EG}\varphi^{-1}\frac{m}{s}x_L\right]\right\}$$

gitt

$$p_H + E + Vf_H^{-1} + LN^{-1} = G_H \quad (\mu_1)$$
$$p_L + E + Vf_L^{-1} + LN^{-1} = G_L \quad (\mu_2)$$
$$f_H \leq \bar{f} \quad (\mu_3)$$
$$\frac{m}{a}\left(x_H - x_L\right) - \varphi c_{\max}\left(f_H - f_L\right) \leq 0 \quad (\mu_4)$$
$$\frac{m}{a}x_L - \varphi c_{\max} f_L \leq 0 \quad (\mu_5)$$

Alle valgvariablene er større enn null. Kuhn-Tuckerbetingelsene for maksimum blir da:

(KT 1) $\frac{\partial L}{\partial p_H} = (1+\lambda)\eta x_H - \mu_1 = 0$

(KT 2) $\frac{\partial L}{\partial p_L} = (1+\lambda)(1-\eta)x_L - \mu_2 = 0$

(KT 3) $\frac{\partial L}{\partial G_H} = -\eta x_H + (1+\lambda)\left[\eta p_H - \eta B_E \varphi^{-1}\frac{m}{s} - \mu_4\frac{m}{a}\right]\frac{\partial x_H}{\partial G_H} + \mu_1 = 0$

(86) (KT 4) $\frac{\partial L}{\partial G_L} = -(1-\eta)x_L + (1+\lambda)\left[(1-\eta)p_L - B_{EG}\varphi^{-1}\frac{m}{s} + (\mu_4 - \mu_5)\frac{m}{a}\right]\frac{\partial x_L}{\partial G_L} + \mu_2 = 0$

(KT 5) $\frac{\partial L}{\partial f_H} = -(1+\lambda)\eta\frac{a}{s}A_E + \mu_1 V f_H^{-2} - \mu_3 + \mu_4\varphi c_{\max} = 0$

(KT 6) $\frac{\partial L}{\partial f_L} = -(1+\lambda)\frac{a}{s}A_{EG} + \mu_2 V f_L^{-2} - \mu_4\varphi c_{\max} + \mu_5\varphi c_{\max} = 0$

(KT 7) $\mu_3, \mu_4$ og $\mu_5 \geq 0$ ( = 0 når vedkommende bibetingelse ikke er bindende)

### 7.1.2 Prissetting

Av KT1 og KT3 får vi optimal prissetting i høybelastningsperioden:

(87)
$$\frac{p_H - B_E\varphi^{-1}\frac{m}{s} - \frac{\mu_4}{\eta(1+\lambda)}\frac{m}{a}}{p_H} = -\frac{\lambda}{1+\lambda}\cdot\frac{1}{El_{G_H} x_H}$$

Av KT2 og KT4 får vi optimal prissetting i lavbelastningsperioden:

$$(88) \qquad \frac{p_L - \frac{1}{1-\eta} B_{EG} \varphi^{-1} \frac{m}{s} + \frac{\mu_4 - \mu_5}{(1-\eta)(1+\lambda)} \frac{m}{a}}{p_L} = -\frac{\lambda}{1+\lambda} \cdot \frac{1}{El_{G_L} x_L}$$

Lagrangemultiplikatoren $\mu_4$ kan ses som skyggeprisen på ressursen kapasitet pr. ekstraavgang. Den bidrar til å heve $p_H$ når optimal kapasitet på ekstravgangene er $c_{max}$, den høyest mulige. Skyggeprisen på kapasitet på grunntilbudet er $\mu_5$, og den bidrar på tilsvarende måte til å øke prisen i lavbelastningsperioden. Samtidig bidrar eksistensen av en kapasitetsskranke på ekstra-avgangene til å *redusere* prisen i lavtrafikkperioden. Grunnen er at det overfører trafikk til grunntilbudet i begge perioder.

Det kan vises at differansen mellom $B_E$ og $(1-\eta)^{-1}B_{EG}$ er $(h/h_L)(B_E - B_G)$. Dersom det rullende materiellet i ekstraavgangene og grunntilbudet er av samme type, men ulik størrelse, vil denne differansen være null eller nær null (parametrene i $B_E$ og $B_G$ vil være like). I fraværet av en bindende skranke på kapasiteten pr. avgang vil det derfor bare være én mulig grunn til at prisen i de to periodene skulle være ulik, nemlig ulik etterspørselselastisitet. Det stiller seg annerledes om man må sette inn en helt annen og tyngre type materiell i rush. Det er også mulig at det finnes faktorer utenom vår modell som tilsier prisdifferensiering. For eksempel kan fraværet av vegprising tilsi lavere pris i rushtida. Det er likevel interessant at det slett ikke er umulig at det er lite å vinne i samfunnsøkonomisk effektivitet ved prisdifferensiering.

Parameteren $\lambda$ er enten skyggeprisen på offentlige midler eller skyggeprisen på et sektorbudsjett eller lokalt budsjett. Vi ser at når $\lambda$ er null eller nær null, har eventuelle ulikheter i etterspørsels-elastisiteten ingen eller svært liten betydning, men dersom sektorbudsjettet er svært stramt, blir det viktig å utnytte denne kilden til inntektsøkning. Det samme gjelder i enda større grad om profittmaksimering er målet (uendelig $\lambda$).

### 7.1.3 Frekvens

Siden det er tre bibetingelser som er ulikheter, har vi åtte mulige kandidater til løsning av maksimeringsproblemet. Ingen av dem gir opphav til motsigelser, men det er mulig at noen av dem likevel er uforenlige med de forutsetningene vi har gjort, men som ikke inngår eksplisitt i maksimeringsproblemet. Det gjelder at $x_H > x_L$ og definisjonen av $A_{EG}$ og $B_{EG}$.

**Ingen bindende bibetingelser**
$\mu_3 = \mu_4 = \mu_5 = 0$ og

$$(89) \qquad f_H^* = \sqrt{\frac{sV}{aA_E}} \cdot \sqrt{x_H} \ , \quad f_L^* = \sqrt{\frac{(1-\eta)sV}{aA_{EG}}} \cdot \sqrt{x_L}$$

**Bare bibetingelse 3 er bindende**
$\mu_4 = \mu_5 = 0$ og

$$(90) \qquad \begin{gathered} f_H^* = \bar{f}, \quad f_L^* = \sqrt{\frac{(1-\eta)sV}{aA_{EG}}} \cdot \sqrt{x_L} \\ \mu_3 = (1+\lambda)\eta\left(Vx_H \bar{f}^{-2} - \frac{a}{s} A_E\right) \end{gathered}$$

**Bare bibetingelse 4 er bindende**

$\mu_3 = \mu_5 = 0$. $f_H$ og $f_L$ kan finnes av følgende to likninger:

$$f_H = f_L + (\varphi c_{\max})^{-1}\frac{m}{a}(x_H - x_L) \tag{91}$$
$$(1-\eta)Vx_L f_L^{-2} + \eta V x_H f_H^{-2} = \frac{a}{s}A_G$$

Problemet er at dette gir opphav til en fjerdegradslikning, og sjøl om vi i prinsippet går inn for eksplisitte analytiske løsninger i dette arbeidsdokumentet, er det måte på hvor stygge formler vi vil ta med. Av samme grunn oppgir vi ikke $\mu_4$.

**Bare bibetingelse 5 er bindende**

$\mu_3 = \mu_4 = 0$ og

$$f_H = \sqrt{\frac{sV}{aA_E}}\cdot\sqrt{x_H}\ ,\quad f_L = (\varphi c_{\max})^{-1}\frac{m}{a}x_L \tag{92}$$
$$\mu_5 = (1+\lambda)(\varphi c_{\max})^{-1}\frac{a}{s}A_{EG}x_L^{-1}\left[x_L - (1-\eta)(\varphi c_{\max})^2\left(\frac{a}{m}\right)^2\frac{sV}{aA_{EG}}\right]$$

Denne løsningskandidaten er fullt mulig slik maksimeringsproblemet er formulert, men vi skal seinere se at den likevel er utelukket på grunn av $x_H > x_L$ og definisjonen av $A_{EG}$.

**Bibetingelse 3 og 4 er bindende**

$\mu_5 = 0$ og

$$f_H^* = \bar{f},\quad f_L^* = \bar{f} - (\varphi c_{\max})^{-1}\frac{m}{a}(x_H - x_L) \tag{93}$$
$$\mu_3 = (1+\lambda)\left[(1-\eta)Vx_L f_L^{*-2} + \eta V x_H \bar{f}^{-2} - \frac{a}{s}A_G\right]$$
$$\mu_4 = \frac{1+\lambda}{\varphi c_{\max}}\left[(1-\eta)Vx_L f_L^{*-2} - \frac{a}{s}A_{EG}\right]$$

**Bibetingelse 3 og 5 er bindende**

$\mu_4 = 0$ og

$$f_H^* = \bar{f},\quad f_L^* = (\varphi c_{\max})^{-1}\frac{m}{a}x_L \tag{94}$$
$$\mu_3 = (1+\lambda)\left[\eta V x_H \bar{f}^{-2} - \frac{a}{s}\eta A_E\right]$$
$$\mu_5 = \frac{1+\lambda}{\varphi c_{\max}}\left[\frac{a}{s}A_{EG} - (1-\eta)Vx_L f_L^{*-2}\right]$$

**Bibetingelse 4 og 5 er bindende**

$\mu_3 = 0$ og

$$f_H^* = (\varphi c_{\max})^{-1}\frac{m}{a}x_H,\quad f_L^* = (\varphi c_{\max})^{-1}\frac{m}{a}x_L \tag{95}$$
$$\mu_4 = (1+\lambda)\frac{a}{s}\eta A_E x_H^{-1}\left[x_H - (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2\frac{sV}{aA_E}\right]$$
$$\mu_5 = \frac{1+\lambda}{\varphi c_{\max}}\left[\frac{a}{s}A_G - (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2 V\left(\eta x_H^{-1} + (1-\eta)x_L^{-1}\right)\right]$$

**Alle bibetingelser er bindende**
$f_H$ er da dobbeltbestemt, hvilket bare er mulig hvis parametrene har helt bestemte verdier. Vi kan se bort fra denne løsningskandidaten.

### 7.1.4 Kommentar

Ved sammenlikning med avsnitt 4.2 finner vi: Dersom bibetingelse nr. 4 ikke er bindende, oppløses forbindelsen mellom tilbudet i høybelastningsperioden og lavbelastningsperioden, slik at løsningskandidatene blir formelt de samme som i énperiodetilfellet, bortsett da fra definisjonen av $A$ og $B$ i det tilfellet og $A_G$, $A_E$, $B_E$ osv. i toperiodetilfellet. For å være mer presis:

- Løsningskandidaten når ingen bibetingelser er bindende er formelt lik kandidatløsningene når ingen bibetingelser er bindende fra to separate maksimeringsproblemer, ett for hver periode.
- Løsningskandidaten med største mulige frekvens i høybelastningsperioden er formelt lik løsningskandidatene fra to separate maksimeringsproblemer, ett (høybelastning) hvor frekvensrestriksjonen er bindende og ett (lavbelastning) hvor ingen restriksjoner er bindende.
- Løsningskandidaten med største mulige kapasitet pr. avgang i lavbelastningsperioden er formelt lik løsningskandidatene fra to separate maksimeringsproblemer, ett (høybelastning) hvor ingen restriksjoner er bindende og ett (lavbelastning) hvor kapasitet pr avgang er bindende.
- Løsningskandidaten med største mulige frekvens i høybelastningsperioden og største mulige kapasitet i lavbelastningsperioden er formelt lik løsningskandidatene i to separate maksimeringsproblemer, ett (høybelastning) hvor frekvensrestriksjonen er bindende og ett (lavbelastning) hvor kapasitet pr avgang er bindende.
- Dessuten viser det seg at tilfellet der våre restriksjoner nr. 4 og 5 begge er bindende, er formelt lik løsningskandidatene fra to separate maksimeringsproblemer der restriksjonen på kapasitet pr avgang er bindende i begge. Grunnen er at tilpasningen av ekstratilbudet da kan gjøres som om etterspørselen etter grunntilbudets avganger i høybelastningsperioder er konstant.

Det er derfor to tilfeller hvor tilpasningen i høybelastningsperioden ikke kan gjøres uavhengig av tilpasningen i lavbelastningsperioden og omvendt, nemlig tilfellet der bare bibetingelse 4 er bindende og tilfellet der bibetingelse 3 og 4 er bindende. Disse tilfellene gir samtidig opphav til mer kompliserte formler for optimale frekvenser, naturligvis. Men alt i alt kan vi konkludere med at når tilbudet er organisert som et grunntilbud pluss ekstraavganger i rush, er det mulig å beregne optimal tilpasning i hver av periodene separat, unntatt dersom transportkapasiteten pr. avgang i rushtida ikke kan økes i optimumspunktet. Men sjøl da er det mulig, dersom transportkapasiteten pr. avgang i optimumspunktet heller ikke kan økes utenom rush.

### 7.1.5 Vilkår som må være oppfylt for at de enkelte løsningskandidatene skal være aktuelle

En mulig løsningskandidat må tilfredsstille visse vilkår som er gitt av ikke-negativitetskravet på Lagrangemultiplikatorene og ulikhetene som oppstår når en bibetingelse ikke er bindende.

Vilkårene er oppsummert i tabellen nedenfor. De er nødvendige, men ikke alltid tilstrekkelige, siden det kan finnes vilkår som vi ikke har regnet ut.

Tabell 1 Nødvendige vilkår for at løsningskandidatene skal være aktuelle

| **Binden-de** | **Grenser for $x_H$** | **Grenser for $x_L$** |
|---|---|---|
| Ingen | $x_H < \frac{aA_E}{sV}\bar{f}^2$ | $x_L < (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2(1-\eta)\frac{sV}{aA_{EG}}$ |
| 3 | $x_H \ge \frac{aA_E}{sV}\bar{f}^2$ | $x_L < (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2(1-\eta)\frac{sV}{aA_{EG}}$ |
| 4 | $x_i < \varphi c_{\max}\frac{a}{m}f_i\,,\ f_i \le \bar{f}\,,\ i = H,L \quad \& \quad (1-\eta)Vx_L f_L^{-2} + \eta V x_H f_H^{-2} \ge \frac{a}{s}A_G$ | |
| 3 og 4 | $x_H < \varphi c_{\max}\frac{a}{m}\bar{f}$<br>$(1-\eta)Vx_L f_L^{-2} + \eta V x_H \bar{f}^{-2} \ge \frac{a}{s}A_G$ | $c_{\max}\frac{a}{m}f_L > x_L \ge \frac{aA_{EG}}{(1-\eta)sV}f_L^2$ |
| 5 | $x_H < \min\left(\frac{aA_E}{sV}\bar{f}^2, (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2\frac{sV}{aA_E}\right)$ | $x_L \ge (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2(1-\eta)\frac{sV}{aA_{EG}}$ |
| 3 og 5 | $\varphi c_{\max}\frac{a}{m}\bar{f} > x_H \ge \frac{aA_E}{sV}\bar{f}^2$ | $x_L \ge (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2(1-\eta)\frac{sV}{aA_{EG}}$ |
| 4 og 5 | $\varphi c_{\max}\frac{a}{m}\bar{f} > x_H \ge (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2\frac{sV}{aA_E}$ | $\frac{a}{s}A_G \ge (\varphi c_{\max})^2\left(\frac{a}{m}\right)^2\left[(1-\eta)Vx_L^{-1} + \eta V x_H^{-1}\right]$ |

Vilkårene når bare bibetingelse 5 er bindende impliserer $A_{EG} > (1-\eta)A_E$. Definisjonen av $A_{EG}$ gir oss da at vi må ha $A_G > A_{EG}$. Ved å inspisere formel (82) ser vi at dette ikke kan være tilfelle i praksis: Det forutsetter at minste kapasitet pr. avgang på ekstratilbudet er betydelig billigere i drift enn minste kapasitet pr. avgang på grunntilbudet. Vi kan derfor konkludere at denne løsningskandidaten er umulig i praksis.

## 7.2 Togtilfellet: Tilpasning av kapasitet pr. avgang er mulig

I tilfellet med et grunntilbud og et ekstratilbud i rush kunne vi øke kapasiteten pr. avgang i rush ved å bruke egne kjøretøy på ekstraavgangene. Disse kjøretøyene var da ikke i bruk i lav-belastningsperioden. Vi ser nå på et tilfelle der alle kjøretøyene er i bruk hele døgnet, men med mindre kapasitet pr. avgang i lavbelastningsperioden. Dette er naturligvis bare mulig i form av et tog, der vogner kan koples fra og parkeres når de ikke trengs. Utviklinga har gått vekk fra tog og over til faste togsett, som eventuelt kan koples sammen. Dette gjør vår antakelse om kontinuerlig tilpasning av kapasitet tvilsom for både T-bane, trikk og lokaltogsett. Ideen om et tog med vogner som kan koples av og på er imidlertid så god at vi ikke kan se bort fra at toget kan komme tilbake.

### 7.2.1 Kostnadene

Trekkenheten er den samme uansett tidsperiode, og vognene er også like, uansett hvor mange de er. Vi behøver altså ikke lenger noen fotskrift på variable som $r_0$ og $r_1$, $g_0$ og $g_1$. Derimot er det mulig at mannskapsbehovet øker med antall vogner. Vi erstatter følgelig bemanningen pr. avgang i de tilfellene vi har behandlet til nå, $\ell_0$, med $\ell_0 + \ell_1 c_H$ rush og $\ell_0 + \ell_1 c_L$ utenom rush. Kapitalkostnaden for det rullende materiellet er uavhengig av om det brukes eller ikke, derfor er den proporsjonal med frekvensen i rush, $f_H$. Den årlige kostnaden blir:

$$\begin{aligned} C_{år} &= \left(r_0 + r_1 c_H\right)\frac{a}{s} f_H + \left(w\ell_0 + g_0 s\right)\left(h_H \frac{a}{s} f_H + h_L \frac{a}{s} f_L\right) + \left(w\ell_1 + g_1 s\right)\left(h_H \frac{a}{s} c_H f_H + h_L \frac{a}{s} c_L f_L\right) \\ &= h_H\left({}^{r_0}\!/_{h_H} + w\ell_0 + g_0 s\right)\frac{a}{s} f_H + h_H\left({}^{r_1}\!/_{h_H} + w\ell_1 + g_1 s\right)\frac{a}{s} c_H f_H \\ &\quad + h_L\left(w\ell_0 + g_0 s\right)\frac{a}{s} f_L + h_L\left(w\ell_1 + g_1 s\right)\frac{a}{s} c_L f_L \\ &= h_H A_H \frac{a}{s} f_H + h_H B_H \frac{a}{s} c_H f_H + h_L A_L \frac{a}{s} f_L + h_L B_L \frac{a}{s} c_L f_L \end{aligned} \tag{96}$$

Den siste likheten i denne formelen definerer $A_H$, $A_L$, $B_H$ og $B_L$. Det er aldri effektivt å kjøre rundt med mer kapasitet enn nødvendig, og siden vi fritt kan tilpasse kapasiteten i begge perioder, har vi:

$$\frac{m}{a} x_i = \varphi c_i f_i \text{ , } i = H, L \tag{97}$$

Setter vi dette inn i likning (96) og deler på $h$, får vi gjennomsnittskostnaden pr. driftstime, $C$:

$$C = \eta\left(A_H \frac{a}{s} f_H + B_H \varphi^{-1} \frac{m}{s} x_H\right) + \left(1 - \eta\right)\left(A_L \frac{a}{s} f_L + B_L \varphi^{-1} \frac{m}{s} x_L\right) \tag{98}$$

### 7.2.2 Maksimeringsproblemet

I tilfellet med et grunntilbud og ekstraavganger var $c_{max}$ den øvre grensa for kapasitet pr. avgang, både for avgangene i grunntilbudet og ekstraavgangene. Siden grunntilbudet og ekstratilbudet blei kjørt samtidig i rush, ga dette opphav til en bibetingelse (bibetingelse 4) som involverte både frekvensen i lavtrafikkperioden og høybelastningsperioden. Det var dette som gjorde at optimal frekvens i begge perioder måtte fastlegges simultant dersom bibetingelse 4 var bindende. Videre gjaldt restriksjonen på største frekvens bare for høybelastningsperioden, siden sammenhengen $f_H = f_G + f_E$ innebærer en lavere frekvens enn den maksimale i lavbelastningsperioden, med unntak av grensetilfellet der $f_E = 0$. I togtilfellet blir alle restriksjoner like uansett hvilken periode det dreier seg om, og de involverer utelukkende variable som tilhører en av periodene. Vi får derfor en tilpasning i hver periode som utelukkende bygger på data om denne perioden, eller m.a.o: Vi kan finne løsningen ved å gå tilbake til avsnitt 4.2.

For å vise det, stiller vi opp maksimeringsproblemet og Kuhn-Tusckerbetingelsene for løsning.

$$\underset{p_H, p_L, G_H, G_L, f_H, f_L}{Max} \quad W =$$

$$(99)\ \eta \int_{G_H}^{\infty} x_H(u)\,du + (1-\eta)\int_{G_L}^{\infty} x_L(u)\,du$$

$$+(1+\lambda)\left\{\eta p_H x_H + (1-\eta) p_L x_L - \eta\left(A_H \frac{a}{s} f_H + B_H \varphi^{-1}\frac{m}{s} x_H\right) - (1-\eta)\left(A_L \frac{a}{s} f_L + B_L \varphi^{-1}\frac{m}{s} x_L\right)\right\}$$

gitt

$$p_H + E + Vf_H^{-1} + LN^{-1} = G_H \quad (\mu_1)$$
$$p_L + E + Vf_L^{-1} + LN^{-1} = G_L \quad (\mu_2)$$
$$f_H \le \bar{f} \quad (\mu_{3a})$$
$$f_L \le \bar{f} \quad (\mu_{3b})$$
$$\frac{m}{a} x_H - \varphi c_{\max} f_H \le 0 \quad (\mu_4)$$
$$\frac{m}{a} x_L - \varphi c_{\max} f_L \le 0 \quad (\mu_5)$$

Alle valgvariablene er større enn null. Kuhn-Tuckerbetingelsene for maksimum blir da:

(KT 1) $\dfrac{\partial \mathsf{L}}{\partial p_H} = (1+\lambda)\eta x_H - \mu_1 = 0$

(KT 2) $\dfrac{\partial \mathsf{L}}{\partial p_L} = (1+\lambda)(1-\eta) x_L - \mu_2 = 0$

(KT 3) $\dfrac{\partial \mathsf{L}}{\partial G_H} = -\eta x_H + (1+\lambda)\left[\eta p_H - \eta B_H \varphi^{-1}\dfrac{m}{s} - \mu_4 \dfrac{m}{a}\right]\dfrac{\partial x_H}{\partial G_H} + \mu_1 = 0$

(100) (KT 4) $\dfrac{\partial \mathsf{L}}{\partial G_L} = -(1-\eta) x_L + (1+\lambda)\left[(1-\eta) p_L - (1-\eta) B_L \varphi^{-1}\dfrac{m}{s} - \mu_5 \dfrac{m}{a}\right]\dfrac{\partial x_L}{\partial G_L} + \mu_2 = 0$

(KT 5) $\dfrac{\partial \mathsf{L}}{\partial f_H} = -(1+\lambda)\eta \dfrac{a}{s} A_H + \mu_1 V f_H^{-2} - \mu_{3a} + \mu_4 \varphi c_{\max} = 0$

(KT 6) $\dfrac{\partial \mathsf{L}}{\partial f_L} = -(1+\lambda)(1-\eta)\dfrac{a}{s} A_L + \mu_2 V f_L^{-2} - \mu_{3b} + \mu_5 \varphi c_{\max} = 0$

(KT 7) $\mu_{3a}, \mu_{3b}, \mu_4$ og $\mu_5 \ge 0$ ( = 0 når vedkommende bibetingelse ikke er bindende)

Det er lett å forsikre seg om at dette gir akkurat samme løsning for hver av periodene som løsningen i énperiodetilfellet i avsnitt 4.2, og samme vilkår for at de enkelte løsningskandidatene skal gjelde. (Naturligvis med $A_H$ og $B_H$ eller $A_L$ og $B_L$ i stedet for $A$ og $B$.)

Alt i alt har vi altså vist at med høvelige definisjoner kan optimal tilpasning i toperiodetilfellet separeres fullstendig i optimal tilpasning i hver av periodene dersom kapasitet pr. avgang kan tilpasses som når et tog kopler vogner på og av. Bortsett fra når største mulige kapasitet pr. avgang er optimalt, gjelder det samme også når tilbudet tilpasses ved hjelp av et grunntilbud og ekstraavganger i rush.

## 7.3 Samme kapasitet pr. avgang i begge perioder

Vi antar nå at det er de samme typene av busser som må brukes både i og utenom rush. Det gjelder da å velge rett kapasitet for disse bussene, i tillegg til valget av pris og frekvens i hver periode. Vi kan ta utgangspunkt i togtilfellet (siste linje i likning (96)) når vi definerer kostnadene $C$. Vi setter altså $c_H = c_L = c$ i likning (96). Framleis kan vi anta at det ikke er lønnsomt å kjøre med ekstrakapasitet i rush, men vi kan ikke unngå det utenom rush. Vi har altså

(101) $$\frac{m}{a} x_H = \varphi c f_H$$

For sikkerhets skyld tar vi også med en restriksjon som sikrer at antall passasjerer pr. avgang utenom rush ikke overstiger antall passasjerer pr. avgang i rush. M.a.o.:

(102) $$x_L f_H - x_H f_L \leq 0$$

Maksimeringsproblemet blir nå

(103) $$\begin{aligned} &\underset{p_H, p_L, G_H, G_L, f_H, f_L, c}{Max} W = \eta \int_{G_H}^{\infty} x_H(u)\,du + (1-\eta) \int_{G_L}^{\infty} x_L(u)\,du \\ &+ (1+\lambda)\left\{\eta p_H x_H + (1-\eta) p_L x_L - \frac{a}{s}\left[\eta A_H f_H + (1-\eta)(A_L + B_L c) f_L\right] - \eta B_H \varphi^{-1} \frac{m}{s} x_H\right\} \end{aligned}$$

gitt

$$\begin{aligned} p_H + E + V f_H^{-1} + LN^{-1} &= G_H \quad (\mu_1) \\ p_L + E + V f_L^{-1} + LN^{-1} &= G_L \quad (\mu_2) \\ f_H &\leq \bar{f} \quad (\mu_3) \\ c &\leq c_{\max} \quad (\mu_4) \\ x_L f_H - x_H f_L &\leq 0 \quad (\mu_5) \end{aligned}$$

Alle valgvariablene er større enn null, med unntak av $p_L$, som også kan være null. Kuhn-Tucker-betingelsene for maksimum blir da:

$$\text{(KT 1)} \quad \frac{\partial \mathrm{L}}{\partial p_H} = (1+\lambda)\eta x_H - \mu_1 = 0$$

$$\text{(KT 2)} \quad \frac{\partial \mathrm{L}}{\partial p_L} = (1+\lambda)(1-\eta)x_L - \mu_2 \leq 0 \quad (= 0 \text{ hvis } p_L > 0)$$

$$\text{(KT 3)} \quad \frac{\partial \mathrm{L}}{\partial G_H} = -\eta x_H + (1+\lambda)\left[\eta p_H - \eta B_H \varphi^{-1}\frac{m}{s} + \frac{\mu_5}{(1+\lambda)} f_L\right]\frac{\partial x_H}{\partial G_H} + \mu_1 = 0$$

$$\text{(104)} \quad \text{(KT 4)} \quad \frac{\partial \mathrm{L}}{\partial G_L} = -(1-\eta)x_L + (1+\lambda)\left[(1-\eta)p_L - \frac{\mu_5}{(1+\lambda)} f_H\right]\frac{\partial x_L}{\partial G_L} + \mu_2 = 0$$

$$\text{(KT 5)} \quad \frac{\partial \mathrm{L}}{\partial f_H} = -(1+\lambda)\eta\frac{a}{s}A_H + \mu_1 V f_H^{-2} - \mu_3 - \mu_5 x_L = 0$$

$$\text{(KT 6)} \quad \frac{\partial \mathrm{L}}{\partial f_L} = -(1+\lambda)(1-\eta)\frac{a}{s}(A_L + B_L c) + \mu_2 V f_L^{-2} + \mu_5 x_H = 0$$

$$\text{(KT 7)} \quad \frac{\partial \mathrm{L}}{\partial c} = -(1+\lambda)(1-\eta)\frac{a}{s}B_L f_L - \mu_4 = 0$$

(KT 7) $\mu_3, \mu_4$ og $\mu_5 \geq 0$ ( = 0 når vedkommende bibetingelse ikke er bindende)

### 7.3.1 Prissetting

Anta først at $p_L > 0$. De optimale prisene finner vi fra KT1 pluss KT3 og fra KT2 (med likhet) pluss KT3. De er:

$$\text{(105)} \qquad \frac{p_H - B_H \varphi^{-1}\frac{m}{s} + \frac{\mu_5}{\eta(1+\lambda)} f_L}{p_H} = -\frac{\lambda}{1+\lambda}\frac{1}{El_{G_H} x_H}$$

$$\text{(106)} \qquad \frac{p_L - \frac{\mu_5}{(1-\eta)(1+\lambda)} f_H}{p_L} = -\frac{\lambda}{1+\lambda}\frac{1}{El_{G_L} x_L}$$

Vi ser at dersom antall reisende pr. avgang blir like stor utenom rush som i rush ($\mu_5$ større enn null), skal vi justere prisen i rush ned og prisen utenom rush opp. Hvis dette ikke er tilfelle ($\mu_5$ = 0), blir prissettingen i rush den samme som før. Vi skal ikke bry oss om hva som skjer utenom rush, sjøl om vi har samme kapasitet i begge perioder. Utenom rush skal vi sette prisen som om det ikke fantes kostnader. Hvis $p_L$ skal være større enn null i det hele tatt, reduserer likning (106) seg til $El_{G_L} x_L = -\frac{\lambda}{1+\lambda}$. Det er ikke sikkert at dette er oppnåelig for noen verdi av priser og frekvenser, og i så fall skal vi sette prisen lik null.

Det er altså stor forskjell på prisene i rush og utenom rush dersom kapasiteten pr. avgang er lik i begge perioder.

### 7.3.2 Frekvens og kapasitet

Anta først at bibetingelse 5 ikke er bindende, dvs. $\mu_5 = 0$. Hvis bibetingelse 4 er bindende, har vi $c = c_{max}$ og dermed umiddelbart av likning (103), som vil gjelde, at $f_H = (\varphi c_{\max})^{-1}\frac{m}{a}x_H$ .

Hvis bibetingelse 3 er bindende, har vi $f_H = \bar{f}$ og dermed $c = \varphi^{-1}\frac{m}{a}x_H \bar{f}^{-1}$.

Hvis ingen bibetingelse er bindende, har vi $f_H = \sqrt{\frac{sV}{aA_H}} \cdot \sqrt{x_H}$.

I alle disse tilfellene er $f_L = \sqrt{\frac{sV}{a\left(A_L + B_L c\right)}} \cdot \sqrt{x_H}$, der $c$ i hvert tilfelle er gitt sammen med $f_H$ som beskrevet. Vi ser at frekvensen utenom rush er lavere enn den ville vært dersom kapasiteten hadde vært valgbar. Det er altså frekvensen, ikke prisen, som skal ta inn over seg at det ikke er kostnadsfritt å drive transport utenom rush.

Hvis bibetingelse 5 er bindende, blir formlene betydelig mer kompliserte.

## 7.4 Samme pris i begge perioder

Å kreve samme billettpris i begge perioder er en nestbesteløsning, dvs. samfunnsøkonomisk nytte er mindre når denne restriksjonen pålegges. Prisen i rush vil bli for lav, og dermed etterspørselen i rush for høy. Det omvendte vil være tilfelle utenom rush. Ut fra det vi har sagt om prissetting i avsnitt 7.1 er det likevel grunn til å anta at det samfunnsøkonomiske tapet i tilfellet med et grunntilbud og ekstraavganger er lite eller ingenting. Det stiller seg litt annerledes i togtilfellet (avsnitt 7.2). Den optimale prisforskjellen vil være større, fordi $B_H$ inneholder kapitalkostnaden for vognene, noe $B_L$ ikke gjør. Dermed blir nestbesteløsningen relativt dårligere. Med utgangspunkt i maksimeringsproblemet for togtilfellet kan vi utlede følgende vilkår for prissettingen når prisen $p$ skal være den samme i begge perioder:

$$\begin{aligned}
&\eta \frac{\partial x_H}{\partial G_H}\left[\frac{\lambda}{1+\lambda}\frac{1}{El_{G_H} x_H} + \frac{p - B_H\varphi^{-1}\frac{m}{s} - \frac{\mu_4}{\eta(1+\lambda)}\frac{m}{a}}{p}\right] \\
&+(1-\eta)\frac{\partial x_L}{\partial G_L}\left[\frac{\lambda}{1+\lambda}\frac{1}{El_{G_L} x_L} + \frac{p - B_L\varphi^{-1}\frac{m}{s} - \frac{\mu_5}{(1-\eta)(1+\lambda)}\frac{m}{a}}{p}\right] = 0
\end{aligned} \tag{107}$$

Hvis uttrykkene i klammeparentesene settes til null, får vi den optimale prissettingen i hver av periodene. Men de kan vanskelig begge bli null så lenge prisen skal være den samme i begge perioder, medmindre $B_H$ er lik $B_L$, ingen restriksjoner er bindende og elastisitetene er de samme. Følgelig sier likningen at vi skal avveie en litt for høy pris i den ene perioden mot en litt for lav pris i den andre. Antall driftstimer og elastisiteten i hver av periodene teller med i denne avveiningen.

Når det gjelder tilfellet der kapasiteten nødvendigvis må være lik i begge perioder, viser avsnitt 7.3 at vi risikerer et mye større tap ved å kreve samme pris i begge perioder. Her er det jo nemlig optimalt med en svært lav pris utenom rush og en tilsvarende høy pris i rush.

Vi har sett at den optimale prisdifferensieringen mellom periodene i høy grad avhenger av hvordan kapasiteten pr. avgang kan tilpasses. Det samme gjelder behovet for slik prisdifferensiering og tapet ved å kaste vrak på den. Denne konklusjonen har muligens ikke vært trukket før i litteraturen.

# 8 Konklusjon og drøfting av anvendelser

I dette notatet har vi stilt opp et optimeringsproblem for driften av et kollektivsystem i to varianter (formel (25) og (35)) og løst problemet fullstendig, med angivelse av vilkårene for hvilken av løsningskandidatene som gjelder i et konkret tilfelle. Vi har vist at problemet gir samfunnsøkonomisk beste løsning og profittmaksimering som spesialtilfeller, og har pekt på at forskjellen mellom de to tilfellene ligger i prissettingen, ikke i prinsippene for valg av frekvens, flatedekning og kapasitet pr. avgang.

Vi har deretter brukt skiftanalyse på løsningen for å finne nytten av fire typer av tiltak: Tiltak for å øke den øvre grensa for frekvens, tiltak for å øke den øvre kapasiteten pr. avgang, tiltak for å bedre framkommeligheten, og tiltak utenom modellen som øker antall kollektivtrafikanter. Resultatene fra skiftanalysene kan enten brukes som generell vegledning til hvordan driftsopplegget bør endres når denne typen skift forekommer, eller til å gjennomføre enkle nyttekostnadsanalyser av tiltak som gir slike skift. To ting er av spesiell betydning i den forbindelsen.

For det første gjelder det å klargjøre om driftsopplegget i utgangspunktet tilsvarer den optimale løsningen av vårt optimeringsproblem, for eksempel når det gjelder forholdet mellom antall linjer og frekvensen på linjene, eller forholdet mellom frekvensen og kapasiteten pr. avgang. Om det ikke er tilfelle, finnes det en billig eller gratis måte å forbedre systemet på.

For det andre gjelder det å klargjøre hvilke restriksjoner som er bindende i den optimale løsningen. Dette påvirker i vesentlig grad hvilke endringer som bør gjøres i driftsopplegget når det forekommer eksogene skift i etterspørselen eller når det gjennomføres framkommelighetstiltak. Det påvirker naturligvis også nytten av tiltak for å avskaffe restriksjonene.

En viktig generell retningslinje for utforming og endring av kollektivtilbudet, er den som omtales i avsnitt 4.2: I mangel av muligheter for å utvide drifta, skal kollektivselskapet øke prisen for at ikke avgangene skal bli overfylt. Eller omvendt: Nytten av investeringer som gjør det mulig å utvide drifta, omfatter også nytten av lavere optimale priser. I dette avsnittet vises det hvordan kollektivselskapets underskudd ikke må være større enn ventetidskostnadene dersom det skal drives optimalt med største mulige frekvens og største mulige kapasitet pr. avgang. [72] Samtidig må grenseinntekta være større enn null.

En annen viktig generell retningslinje trekkes i avsnitt 6.3.3: Eksogene etterspørselsskift skal alltid medføre en like stor økning i alle de dimensjonene av tilbudet som ikke er låst fast av restriksjonene. Når det gjelder framkommelighetstiltak er det annerledes. Frekvens og flatedekning skal da prioriteres på bekostning av kapasitet pr. avgang. Vi har grunn til å anta at dette ville vært annerledes dersom kostnadene ved trengsel om bord hadde vært inkludert i modellen, og at resultatet da hadde vært at tilbudet skal utvides like mye i alle frie dimensjoner.

I kapittel 7 viser vi at med høvelige definisjoner kan optimal tilpasning i toperiodetilfellet separeres fullstendig i optimal tilpasning i hver av periodene dersom kapasitet pr. avgang kan tilpasses som når et tog kopler vogner på og av. Bortsett fra når største mulige kapasitet pr. avgang er optimalt, gjelder det samme også når tilbudet tilpasses ved hjelp av et grunntilbud og ekstraavganger i rush. Dette er nyttige resultater når det gjelder å bruke modellen sammen med en transportmodell som skiller mellom etterspørselen i rush og utenom rush. Den optimale

[72] Dette gjelder ved profittmaksimering. Det samfunnsøkonomiske vilkåret er litt slakkere.

prisdifferensieringen mellom periodene i høy grad vil avhenge av hvordan kapasiteten pr. avgang kan tilpasses. Det samme gjelder behovet for slik prisdifferensiering og tapet ved å kaste vrak på den. Behovet er størst når en er nødt til å bruke samme kapasitet pr. avgang både i og utenom rush. Denne konklusjonen har muligens ikke vært trukket før i litteraturen.

Den viktigste forutsetningen for vår modell er at alle trafikanter opplever reisekostnadene, inkludert kvalitetsaspektene ved reisa, på samme måte. Så urealistisk som det enn er, er det likevel en normal forutsetning i transportøkonomien, og denne forutsetningen hindrer ikke sammenlikning med andre beregninger som bygger på det samme, eller integrasjon med transportmodeller som også gjør det. Andre viktige forutsetninger har med graden av aggregering å gjøre. Innfører vi en uakseptabel forenkling om vi antar at alle reiser har gjennomsnittlig lengde, eller at alle kollektivlinjer har samme linjelengde, rundturtid og frekvens? Den sistnevnte forutsetningen kan testes innafor modellsystemet i notatet vårt, i og med at vi også har utarbeidet løsningene for enkeltlinjer.

Vår modell kan som nevnt anvendes som en tilbudsmodell, som sammen med en etterspørselsmodell (en transportmodell) kan utgjøre et mer fullstendig system for å evaluere transporttiltak. Siden de to sidene har ulikt aggregeringsnivå, er det imidlertid en del utfordringer. Hovedproblemet er hvordan resultater fra tilbudsmodellen skal føres tilbake til transportmodellen på den enkleste og raskeste måten. Dersom for eksempel tilbudsmodellen tilsier endring i antall linjer, vil det medføre omkoding i nettverksdelen av transportmodellen, slik at en ”kjøring” av modellsystemet som helhet vil bestå av mye manuelt arbeid. Det beste en kan oppnå når det gjelder antall linjer, er derfor trolig at tilbudsmodellen i utgangspunktet gir visse ideer til kodingen av nettverket, som siden holdes fast.

Når det gjelder frekvens, er mulighetene for automatisk oppdatering av transportmodellen trolig bedre. En full kjøring vil likevel bestå av mange gjentatte kjøringer av etterspørselsmodell og tilbudsmodell i loop, og vil ta tid.

En fordel ved å bruke vår modell som tilbudsmodell, er at den forhindrer ”overcrowding”, altså at etterspørselen pr. avgang overstiger kapasiteten pr. avgang, slik at kostnadene for kollektivsystemet ikke kan bli urealistisk lave.

Både når det gjelder generell vegledning, nyttekostnadsanalyse og integrering i et transportmodellsystem gjelder det at vi har riktige data til modellen vår. Det er det langt fra umulig å skaffe. Mye relevante data finnes i årsmeldingene fra kollektivselskapene, for eksempel Oslo sporveier. Enkelte typer data, som hvordan drivstofforbruk og kapitalkostnad henger sammen med kjøretøystørrelse (kapasitet pr. avgang), kan det trenges egne analyser for å frambringe. Arbeidet med å finne riktige data til aktuelle norske anvendelser gjenstår.

## Litteraturliste

Minken, H. og G. Dahl (2007) Trange grenser for frekvensøkning på T-banen. *Samferdsel* nr. 10/2007.

Small, K.A (2004) Road pricing and public transport. In: G. Santos (ed.) *Road Pricing – Theory and Evidence.* Elsevier, Amsterdam.

Steinsland, C. (2008) Systemdokumentasjon Stratos. Arbeidsdokument ØL/2097/2008, TØI.

Sydsæter, K. (1990) Matematisk analyse. Bind II. Universitetsforlaget.

SVV (2006) Konsekvensanalyser. Håndbok 140.

Tirole, J. (1988) The Theory of Industrial Organization. MIT Press, New York.

# 4 Noen enkeltvirkninger

## 4.1 Køkostnader på grunn av hastighetsforskjeller[73]

[73] Dette er arbeidsdokument TØ/1296/2001.

# Køkostnader på grunn av hastighetsforskjeller

## Innhold



## 1 Bakgrunn

Dette arbeidsdokumentet ble opprinnelig utarbeidet i prosjektet om den samfunnsøkonomiske gevinsten av å øke maksimal vogntoglengde, som ble gjennomført ved TØI høsten 2000. Deler av teksten er innarbeidet i arbeidsdokumentet som redegjør for detaljene i det prosjektet (arbeidsdokument TØ/1294/2000).

Det er to grunner til å gjøre det til et arbeidsdokument i TØIs formelle system. Den ene er at det kan være aktuelt med prosjektforslag fra TØI som sikter mot å innarbeide køkostnader på grunn av hastighetsforskjeller i vegvesenets opplegg for samfunnsøkonomiske beregninger. Den andre grunnen er at denne typen køkostnader er framtredende i jernbanedrift, og at modellen vi behandler kanskje egner seg enda bedre i en slik sammenheng.

## 2 Kø bak et saktegående kjøretøy

Tungtrafikken kan ofte gi opphav til tidstap for andre biler på vegen. Dette gjelder naturligvis i første rekke på smale og svingete tofeltsveger og i oppoverbakke uten krabbefelt. Lengre vogntog med større last kan øke problemene. Dette kan skje på to måter – både ved at det blir lengre mellom hvert sted hvor det er trygt å kjøre forbi, og ved at de nye vogntogstypene kan ha lavere gjennomsnittsfart på de utsatte strekningene.

Litteraturen om køer i vegsystemet har til nå dreid seg om køer av tre slag. Det er mest vanlig å bygge på en modell med en jamn strøm av like kjøretøyer på en veg som er lik overalt. Siden førerne vil ønske å holde en avstand til kjøretøyet foran som avhenger av farta, vil tettere trafikk føre til at førerne må redusere farta for å holde den ønskelige avstanden. Denne modellen gir opphav til en tidskostnad pr. kilometer som stiger med trafikkvolumet, og stiger brattere jo større trafikkvolumet er. Denne typen køkostnader er det som er lagt inn i nettverksmodeller til bruk i byområder. Samme problemstilling i en dynamisk sammenheng fører til modeller for hvordan køene tetner til og løser seg opp på en mer eller mindre rytmisk måte. Slike modeller er mest brukt i teoretisk sammenheng. Den tredje modellen er den såkalte flaskehalsmodellen (Vickrey 1969, Arnott, de Palma og Lindsey 1993), der vegen ikke er lik overalt, men har en innsnevring med lavere kapasitet.

Mens køproblemene i disse modellene skyldes stor trafikk av kjøretøyer som er like i trafikkteknisk henseende, er køproblemer som skyldes at kjøretøyene *ikke* er like, men holder ulik fart, langt mindre studert. Dette er imidlertid en problemstilling som er vel så relevant i vanlig norsk sammenheng. Våre veger er smale, svingete og med stigninger. Tungtransporten utgjør bare en del av problemet. I tillegg har vi bobiler om sommeren, og til dels store variasjoner mellom bilførerne når det gjelder hva slags hastigheter de føler seg komfortable med. Noen av disse variasjonene skyldes at bilparken er satt sammen av biler med svært forskjellig alder og motorkraft.

Christensen (1997) bygde på sannsynlighetsteori og utledet en formel for forsinkelsen til et enkelt kjøretøy som risikerer å komme bak en saktegående tungtransport. Verhoef, Rouwendal og Rietveld (1999) gjør greie for en modell som er noe mer avansert, bl.a. fordi den tar med seg på en mer eksplisitt måte muligheten for at vogntoget kan samle opp en kø av biler bak.[74] Her vil vi bygge på dette siste arbeidet.

En kø av biler bak et saktegående kjøretøy kan tenkes å gi opphav til to typer av kostnader. Den første er naturligvis køkostnaden inntil de bakenforliggende bilene slipper forbi. Den andre er ulykkeskostnaden ved forbikjøring på uegnede steder. Disse kostnadene er alternativer til hverandre. Bilførerne kan velge å finne seg i køkostnaden eller å pådra seg en ulykkesrisiko for å unngå den.

# 3 Formelen for forventet tidstap

Vi antar at det finnes to typer trafikanter. De som vil og kan kjøre raskt, har hastighet $s_1$. De som enten ikke kan eller ikke vil kjøre raskt, har hastighet $s_2$. Hastighet er målt i kilometer i timen.

Betrakt en strekning på m kilometer. Denne strekningen er definert ved at det ikke er mulig å kjøre forbi før slutten av strekningen (med der finnes det til gjengjeld en mulighet). Vi antar at alle bilførere ønsker en avstand på d* (kilometer!) til bilen foran, uansett hastighet. Dette er ikke helt realistisk, men bra nok i vår sammenheng. De raske kjører derfor i hastigheten $s_1$ inntil de kommer i en avstand av d* til et kjøretøy som kjører i $s_2$. Deretter kjører de i hastigheten $s_2$ til slutten av strekningen. På denne måten kan det samle seg opp en kø bak det saktegående kjøretøyet. Ved slutten av strekningen løser *hele* køen seg opp. Når vi kommer til neste strekning av liknende slag, ankommer derfor raske og langsomme kjøretøyer i en tilfeldig rekkefølge.

Når vi skal anvende formlene nedenfor, er forutsetningen om at hele køen oppløser seg til slutt, av vesentlig betydning for hvordan vi skal anslå strekningslengda m. Hvis vi antar at vogntoget er meget omtenksomt og slipper alle bakenforliggende biler forbi ved første rimelige anledning, får vi korte m. Hvis derimot vogntoget bare slipper forbi noen få hver gang det er mulighet til det, må vi regne enden av m til å ligge et sted mellom første og andre forbikjøringsmulighet, eller til å være enda lengre.

Hver time ankommer det $\rho_1$ raske og $\rho_2$ langsomme kjøretøyer til strekningen m. Disse ankomstratene er såpass lave at det ikke oppstår vanlig kø foran eller på hele strekningen m.

[74] Verhoef et al kjenner bare til ett tidligere arbeid med denne typen av kø, nemlig Tzedakis (1980). Christensen på sin side var ikke oppmerksom på arbeidet til Tzedakis da han utledet sin formel.

Uten hindringer ville de raske kjøretøyene brukt tida $t_0 = \frac{m}{s_1}$ på strekningen. Forventet kjøretid på strekningen for de raske bilene *ut over $t_0$*, eller med andre ord forventet *forsinkelse* for raske biler på strekningen, er:

(1) $$F_1 = \frac{1}{2} p \cdot \left( \frac{m}{s_2} - \frac{m - d^*}{s_1} \right)$$

der

(2) $$p = \frac{m(s_1 - s_2) + s_2 d^*}{(s_1 - d^* \rho_1)\left( s_2 / \rho_2 \right)}$$

OBS! For at formelen (1) + (2) skal gjelde, må parametrene være gitt slik at p er mindre enn 1. I motsatt fall gjelder en annen formel, se Verhoef et al (1999). Variabelen p kan vi tolke som sannsynligheten for at et raskt kjøretøy vil oppleve en forsinkelse på strekningen. Hvis det er så mange langsomme kjøretøyer at denne sannsynligheten er 1, må vi altså anvende en annen formel.

Formel (1) + (2) er egnet til å gi en forståelse av hvordan forsinkelsene opptrer, nemlig som sannsynligheten for at det skal forekomme forsinkelse i det hele tatt, multiplisert med den gjennomsnittlige forsinkelsen som oppstår i så fall. Men ved å regne ut (1) med (2) innsatt kan vi komme fram til en formel for forsinkelsen som er mer hensiktsmessig for våre formål.

(3) $$F_1 = \frac{1}{2} \frac{\rho_2}{s_1(s_1 - d^* \rho_1)} \cdot \left( \frac{s_1 - s_2}{s_2} m + d^* \right)^2$$

Hver enkelt av de raske bilene som benytter strekningen i løpet av en time, vil altså oppleve en forventet forsinkelse på $F_1$. Det vil være ett og bare ett saktegående kjøretøy ("gåsemor") som forårsaker forsinkelsen for en kø av biler. Siden det er $\rho_2$ saktegående og $\rho_1$ raske kjøretøyer pr. time, vil det enkelte saktegående kjøretøyet ved å trafikkere strekningen gi opphav til et gjennomsnittlig forventet samlet tidstap på $\rho_1/\rho_2$ timer for andre kjøretøyer. Vi kommer da fram til følgende formel for de forventede eksterne køkostnadene, målt i timer, $K_2$, som et saktegående kjøretøy gir opphav til ved å trafikkere en enkelt strekning uten forbikjøringsmuligheter:

(4) $$K_2 = \frac{1}{2} \frac{\rho_1}{s_1(s_1 - d^* \rho_1)} \cdot \left( \frac{s_1 - s_2}{s_2} m + d^* \right)^2$$

Denne eksterne kostnaden er uavhengig av ankomstraten $\rho_2$. De marginale og de gjennomsnittlige eksterne køkostnadene av dette slaget er derfor like. Dette er vesensforskjellig fra andre former for eksterne køkostnader på veg.

### 3.1 Kvalitativ vurdering av omfanget av køproblemer ved lengre vogntog

Formel (4) er utviklet for å kunne beregne virkningen på kødannelse når vogntogene tillates å være lengre. Som nevnt innledningsvis kan den ha en breiere anvendelse.

Vi vurderer først om modellen har svakheter, deretter ser vi nærmere på hvordan den virker når vi får endringer i kjøretøyenes egenskaper og i trafikken.

I virkelighetens verden finns det vegstrekninger som er skiltet eller merket slik at forbikjøring ikke er tillatt. Modellen passer bra for slike strekninger. Problemet er at den passer relativt dårlig der hvor forbikjøring er tillatt, men byr på problemer. Det er nemlig slik at jo større fart det saktegående kjøretøyet har, jo vanskeligere blir det å finne et sted der forbikjøring er tilrådelig. Dermed kan det tenkes at jo mindre fartsdifferansen er, jo lengre vil det gå før forbikjøring kan foretas, og jo lengre vil forsinkelsene vare. Denne tendensen motvirker den tendensen som er modellert, nemlig at større fartsdifferanse gir større forsinkelse pr. kilometer. Det er tenkelig at på slike strekninger vil det finnes en "verste" fartsdifferanse, med mindre samlede forsinkelser både om differansen går opp og om den går ned. Hvis spesielt det saktegående kjøretøyet holder *variabel* fart, slik at det når fartsgrensa der hvor forbikjøring hadde vært mulig, men holder 50-60 km/t der hvor det ikke er mulig, vil vi kunne ha en type problemer som slett ikke gjenspeiles i modellen.

Dette må vi leve med, men det kan gjøre at en reduksjon i den gjennomsnittlige fartsdifferansen i praksis ikke vil redusere forsinkelseskostnadene slik som beregnet i modellen.

Ved lengre vogntog vil flere ting endre seg i formel (3) og (4). Hastigheten til vogntogene, $s_2$, kan gå ned. Avstanden m til nærmeste forbikjøringsmulighet kan gå opp. Dette er de to forholdene som er relevante for omfanget av eksterne kostnader som det enkelte kjøretøyet forårsaker, altså formel (4).

Det tredje forholdet som vil endre seg er antall kjøretøyer og turer som er nødvendig for å transportere et gitt volum. En reduksjon i det totale antall turer som er nødvendig, kan innebære at det bli færre vogntog pr. time på strekningen. $\rho_2$ kan altså gå ned. Dette har betydning for forventet kjøretid for raske biler, altså formel (3). Men her må en passe på å ta i betraktning hvordan andre saktegående kjøretøyer, som bobiler, virker sammen med vogntogene. Parameteren $\rho_2$ må inkludere denne andre saktegående trafikken også. Den prosentvise reduksjonen i $\rho_2$ blir altså ikke like stor som den prosentvise reduksjonen i antall vogntog.

Av formel (3) og (4) kan vi beregne de følgende elastisitetene:

$$El_{\rho_2}F_1 = 1, \quad El_{\rho_2}K_2 = 0$$

$$El_m F_1 = El_m K_2 = 2\frac{\frac{s_1 - s_2}{s_2}m}{\frac{s_1 - s_2}{s_2}m + d^*} \approx 2$$

$$El_{s_2}F_1 = El_{s_2}K_2 = -2\frac{s_1 m}{s_1 m - s_2\left(m - d^*\right)} \approx -2\frac{1}{1 - {s_2}/{s_1}}$$

Første linje sier at en reduksjon i antall saktegående kjøretøyer på strekningen reduserer forventet forsinkelse for en rask bil tilsvarende, men endrer ikke køproblemene som det enkelte saktegående kjøretøyet forårsaker. Andre linje sier at en økning på 1% i lengda på strekningen hvor forbikjøring er umulig, øker køproblemene med 2%.

Fra tredje linje er det klart at når $s_2$ reduseres med 1%, øker køproblemene med svært mange prosent, avhengig av forholdet mellom $s_1$ og $s_2$. La oss si at vogntogene holder 60 km/t og privatbilene 80 km/t opp en bakke. En liten reduksjon i vogntogenes fart vil da øke køproblem-ene med 8%. Det kan derfor være viktig å stille krav til motorytelse som kan motvirke at tiltaket medfører en fartsreduksjon. Dette er desto viktigere hvis vi tenker på den situasjonen som eksisterer på strekninger hvor forbikjøring formelt sett er mulig, men vanskeliggjøres av at det saktegående kjøretøyet holder høyere fart der hvor det er mulig enn der hvor det ikke er mulig. På den andre sida vil en del av miljøgevinsten ved mer last pr. bil forsvinne hvis motorene må gjøres mye sterkere.

Hva skal til for at vi kan se bort fra at det blir en effekt av tiltaket for andre trafikanter? Med andre ord, hvilke samtidige endringer kan skje i de tre variablene på en slik måte at $F_1$ ikke endrer seg? Ved å totaldifferensiere formel (3) kan vi få svaret på det. Vi får:

$$dF_1 = \left[\rho_2^{-1}d\rho_2 + 2\frac{\frac{s_1 - s_2}{s_2}}{\frac{s_1 - s_2}{s_2}m + d^*}dm - 2s_2^{-1}\frac{\frac{s_1}{s_2}m}{\frac{s_1 - s_2}{s_2}m + d^*}ds_2\right]F_1$$

Her er $d\rho_2$ endringen i antall saktegående kjøretøyer pr. time som følge av tiltaket. Denne er negativ (en reduksjon). Videre er dm endringen i lengda på strekningen som ikke tillater forbikjøring. Den er positiv (en økning). Endelig er $ds_2$ endringen i hastigheten til de sakte-gående kjøretøyene. Den er negativ (en reduksjon). $dF_1$ er endringen i forventet forsinkelse for det enkelte raske kjøretøyet på grunn av tiltaket, og vi lurer altså på om den kan tenkes å bli null, og i tilfelle hvordan.

Vi ser at hvis $d\rho_2$ er nær null (f.eks. fordi det er et tungt innslag av andre saktegående kjøretøyer), finns det små muligheter for å unngå at tiltaket forverrer situasjonen, enten gjennom økt m eller redusert $s_2$ eller på begge måter. Derimot er mulighetene tilstede dersom en meget stor andel av den saktegående trafikken utgjøres av vogntog som blir påvirket av tiltaket. Hvis f.eks. de større vogntogene har samme fart som de nåværende ($ds_2 = 0$), så kan $dF_1$ bli null dersom $d\rho_2 \approx - 2\rho_2 dm$, dvs. dersom vogntogene er ekstremt flinke til å slippe folk forbi, eller forbikjøringsmulighetene etter strekningen er så gode at den nye vogntoglengda spiller liten rolle.

# 4 Nyttekostnadsanalyse basert på gjennomsnittsverdier

Vi utvider nå perspektivet fra en enkeltstrekning til en hel rute. En viss andel av ruta vil være uten problemer med forbikjøring, og kjøretida der blir ikke påvirket av tiltaket. Så la oss se på problemstrekningene samlet. Hensikten er å komme fram til et grovt mål på de samlede forsinkelsene som en gjennomsnittlig kilometer med et saktegående kjøretøy forårsaker. Dette er vår tilnærming fordi en detaljert gjennomgang av ruter med sikte på å registrere de n problemstrekningene og deres respektive lengder $m_i$ ($i$ = 1, 2, ...,n) åpenbart krever et omfattende kartleggingsarbeid.

Vi ser bort ifra kjøreavstanden d* i kvadratet i formel (4), men beholder den i den første faktoren. Vi skriver derfor de eksterne køkostnadene pr. vogntog på problemstrekning i, $K_{2i}$, på en forenklet måte:

$$(5) \qquad EK_{2i} = x_i y_i m_i^{\,2}, \quad \text{der } x_i = \rho_{1i} \text{ og } y_i = \frac{1}{2}\frac{1}{s_{1i}\left(s_{1i} - d^*\rho_{1i}\right)}\left(\frac{s_{1i} - s_{2i}}{s_{2i}}\right)^2$$

Legg merke til at i (5) har vi gjort trafikktettheten $\rho_{1i}$ og hastighetene $s_{1i}$ og $s_{2i}$ streknings-spesifikke. $\rho_{1i} = x_i$ er trafikktettheten av raske biler, og $y_i$ er et uttrykk som i det vesentlige avhenger av forskjellen i hastighet mellom raske og saktegående kjøretøyer. Vi vil nå komme fram til et uttrykk for den gjennomsnittlige køkostnaden for et vogntog på de n problematiske strekningene. Vi oppfatter X, Y og M som stokastiske variable med realiserte verdier $x_i$, $y_i$ og $m_i$. Vi skriver $Z = XYM^2$, og spør altså etter et uttrykk for E(Z).

Vi benytter regnereglene EXY = EXEY + cov(X,Y) og $EX^2 = (EX)^2$ + varX, og kommer fram til følgende formel:

$$(6) \qquad EZ = \left((EM)^2 + \operatorname{var} M\right)\left(EXEY + \operatorname{cov}(X,Y)\right) + \operatorname{cov}(XY, M^2)$$

I utgangspunktet ser vi ingen god grunn til at noen av kovariansene skal være forskjellig fra null. Hvorfor skulle hastighetsforskjellene samvariere med trafikktettheten? Hvorfor skulle hastigheter og trafikktetthet ha noe samband med hvor langt det er mellom hvert punkt hvor forbikjøring er mulig? Vi trur derfor ikke at kovariansene vil innvirke vesentlig på regnestykket. Hva mere er: Vi tror ikke kovariansene vil endre seg som følge av tiltaket. Vi kommer til å se bort fra dem her, hvor et grovt anslag for køkostnadene er alt vi kan satse på å etablere.

Betrakt en gjennomsnittskilometer med stamveg (eller annen veg som påvirkes av tiltaket). Problemstrekningene utgjør en andel $\alpha$ av denne kilometeren, og er følgelig $\alpha$/EM i antall. Det oppstår en forventet køkostnad EZ hver gang et vogntog av tilstrekkelig størrelse begir seg ut på vår gjennomsnittskilometer. Total forventet køkostnad når de transportene som påvirkes av tiltaket har et årlig trafikkvolum på D kjøretøykilometer, er da α*D*EZ*EM/EM = α*D*EZ. Vi bruker indeks 0 for situasjonen uten tiltaket, og indeks 1 for situasjonen med tiltaket, og har:

$$\text{(7)} \quad \begin{aligned} &\textit{Sparte køkostnader pga. tiltaket} \\ &\approx \alpha\left[D_0\left((EM_0)^2 + \operatorname{var} M_0\right)EXEY_0 - D_1\left((EM_1)^2 + \operatorname{var} M_1\right)EXEY_1\right] \end{aligned}$$

der

α = et anslag på den andelen av vegsystemet der forbikjøring bare er mulig på enkelte punkter. (Bare vegsystemet som er med i tiltaket regnes med her).

D = kjøretøykilometer pr. år for kjøretøyer som erstattes med lengre vogntog etter tiltaket.

EM = gjennomsnittlig lengde av strekninger der forbikjøring er umulig.

varM = variansen til lengda av strekninger der forbikjøring er umulig.

EX = gjennomsnittlig timetrafikk av lette biler på veglenker i det relevante vegsystemet. (Endres ikke av tiltaket).

EY = gjennomsnittlig størrelse på faktoren $y = \frac{1}{2}\frac{1}{s_1\left(s_1 - d^*\rho_1\right)}\left(\frac{s_1 - s_2}{s_2}\right)^2$ på strekninger der forbikjøring er umulig.

For å anvende formel (7) i det foreliggende prosjektet trenger vi anslag på disse størrelsene, altså:

- Hvor stor del av det vegnettet er slik at forbikjøring av et stort vogntog bare er mulig på enkelte punkter?
- I den delen av vegnettet der forbikjøring bare er mulig på enkelte punkter, hvor langt vil det gjennomsnittlig være mellom punkter hvor forbikjøring *er* mulig (før og etter tiltaket)?
- Er det mulig, f.eks. ved å granske en eller to stamvegruter mer i detalj, å gi oss et datamateriale med lengder mellom to punkter der forbikjøring er mulig?
- Hva er gjennomsnittlig timetrafikk av lette biler på veglenker i det relevante vegsystemet?
- I den delen av vegnettet der forbikjøring bare er mulig på enkelte punkter, hva er den gjennomsnittlige forskjellen mellom hastigheten til en privatbil og et typisk tungt kjøretøy som berøres av tiltaket (før og etter)?

# Litteratur


Arnott, R., A. de Palma and R. Lindsey (1993) A structural model of peak-period congestion: a traffic bottleneck with elastic demand. *American Economic Review* **83**(1), 161-179.

Christensen, P. (1997) Forventet forsinkelse for personbiler pga tungtransport. Vedlegg i Eriksen, K.S (1997) *Massetransport ved byggeprosjekter*. TØI-rapport 376/1997, TØI, Oslo.

Eidhammer, O, Minken H, Killi, M: *Samfunnsøkonomiske virkninger av å innføre vogntog med lengde 25,25 m og totalvekt 60,0 tonn.* Oslo, Transportøkonomisk institutt. Arbeidsdokument TØ/1294/2000.

Tzedakis, A. (1980) Different vehicle speeds and congestion costs. *Journal of Transport Economics and Policy* **14**, 81-103.

Verhoef, E., J. Rouwendal and P. Rietveld (1999) Congestion caused by speed differences. *Journal of Urban Economics* **45**(3), 533-556.

Vickrey, W. (1969) Congestion theory and transport investment. *American Economic Review* **59**, 251-260.

## 4.2 Hendelsesrelaterte forsinkelser på lenker[75]

[75] Dette avsnittet er arbeidsdokumentet TØ/1682/2004, datert 20. oktober 2004 og revidert 10. november samme år. Med to unntak er det identisk med vedlegg 4 i Minken, H. og Samstad, H. (2004): Virkningsberegninger av tiltak for raskere og mer pålitelig godstransport – en ny metode. TØI-rapport 825/2006. Unntaket er at arbeidsdokumentet har et kort første avsnitt som ikke er med i TØI-rapportens vedlegg 4, og at vedlegg 4 også mangler hele avsnitt 8 om Erlangfordelingen.

# En modell av hendelsesrelaterte forsinkelser på lenker

## Innhold



## 1. Innledning

Hendelser (ulykker, motorstopp m.m.) er trolig en vesentlig årsak til variabilitet i transporttida. Vi tenker da ikke i første rekke på hendelser som den enkelte bilist sjøl har forårsaket eller blitt utsatt for, men hendelser som skjer med andre, og som fører til forsinkelser for vår bilist og mange fler. Siden hundrevis og kanskje tusenvis av bilister forsinkes av en hendelse, er konsekvensene for de som bare passivt er innblandet i hendelsen, trolig større i sum enn konsekvensene for de som er aktivt innblandet, med mindre hendelsen er en alvorlig ulykke.

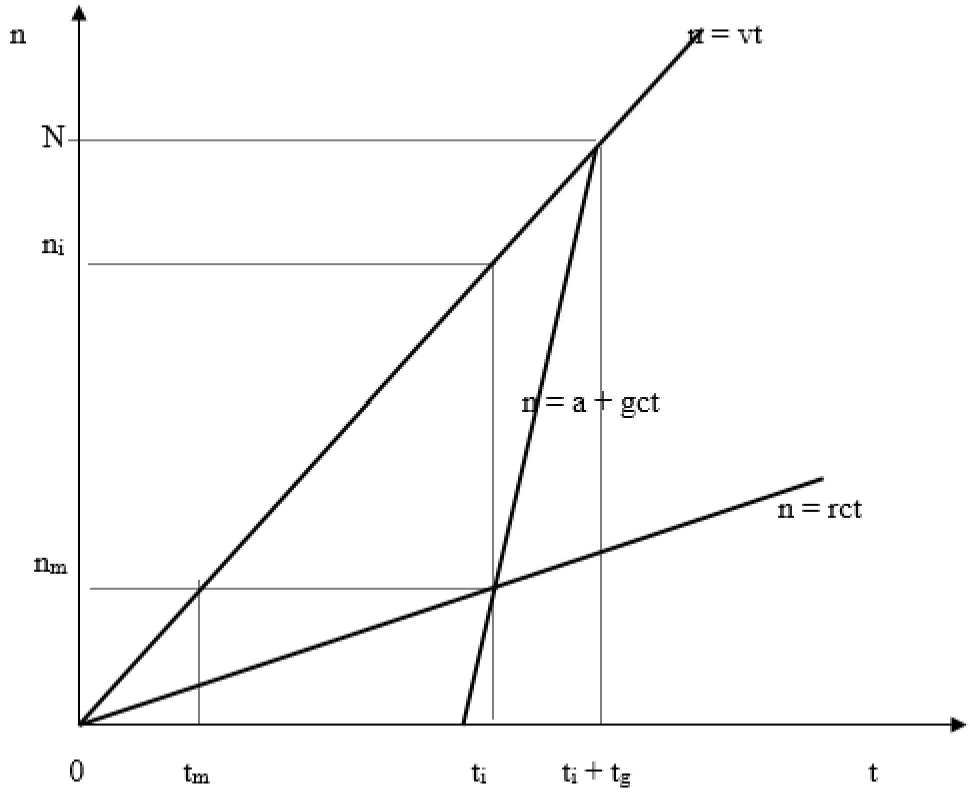
En strøm v av biler passerer over en lenke med kapasitet c. For hver kilometer som den enkelte bilen kjører, er det en fast sannsynlighet $\lambda$ for at en hendelse inntreffer. Når hendelsen inntreffer, reduseres vegkapasiteten på dette stedet fra c til rc. Tida det tar å rydde vegbanen er $T_i$. $T_i$ er en stokastisk variabel, som i tilfellet vi ser på antar verdien $t_i$. En kø oppstår så lenge vegbanen ennå ikke er ryddet, siden vi forutsetter at $v > rc$. Når vegbanen er ryddet, oppløses køen ved at bilene strømmer videre med raten gc, $c \geq gc > v > rc$. På tidspunkt $t_i + t_g$ er køen avviklet, og alt er som før.

Dette er modellen i Cohen og Southworth (1999). Utledningen av sannsynlighetsfordelingen for forsinkelsen som en tilfeldig bilist som bruker denne lenken vil oppleve, er ikke eksplisitt i artikkelen deres, og det er dette vi vil rette opp her. Vi ønsker altså å finne forventning og varians for hendelsesrelaterte forsinkelser for biler som bruker denne lenka, basert utelukkende på de forutsetninger som er gitt ovenfor.

Vi vil vi ta utgangspunkt i en grafisk framstilling av modellen. Vi benytter oss av et såkalt kumulativt plott (Newell 1982), der den horisontale aksen er tida og den vertikale aksen er det kumulative antall biler som ankommer og etter hvert passerer hendelsesstedet. Hendelsen inntreffer på tidspunkt 0. Bilene er nummerert slik at den første bilen som ankommer hendelsesstedet etter hendelsen har nummer 1, osv. La t være tida etter hendelsen, og la n være det kumulative antallet biler som registreres på hendelsesstedet. Vi antar som en tilnærming at n er en kontinuerlig variabel. Etter hva vi har sagt allerede, vil n = vt vise hvor mange biler som har ankommet ved tidspunkt t, og n = rct viser hvor mange som har passert hendelsesstedet ved tidspunkt $t < t_i$ , altså i tida før vegbanen er ryddet. For $t > t_i$ vil n = a + gct vise hvor mange biler som har passert hendelsesstedet, forutsatt at a tilpasses slik at $rct_i = a + gct_i$. Alle disse linjene er tegnet inn i det kumulative plottet.

På et gitt tidspunkt t vil den *vertikale* avstanden mellom n = vt og den nærmeste av kurvene under vise hvor mange biler som akkurat da står i kø ved hendelsesstedet. For bil nummer n vil den *horisontale* avstanden mellom n = vt og den nærmeste av kurvene til høyre for den, vise hvor lenge denne bilen må stå i kø på hendelsesstedet. Det er disse to egenskapene ved plottet som gjør det nyttig for oss. Vi har tegnet inn tidspunktet $t_m$, som er ankomsttidspunktet for den bilen som får den lengste ventetida i kø, og tidspunktet $t_i + t_g$, som er ankomsttidspunktet for den siste bilen som blir berørt av hendelsen. $t_g$ er altså tida det tar fra vegbanen er ryddet til trafikken går som normalt igjen.

*Figur 1.*




Åpenbart er arealet av trianglet mellom de tre kurvene den totale ventetida på grunn av hendelsen, og ventetida pr. bil stiger lineært fra den første bilen som blir berørt til den bilen som får lengst ventetid, og avtar deretter lineært igjen til den siste bilen som blir berørt. Vi vil nå bruke diagrammet til å regne ut den totale ventetida og fordelinga av ventetid på hver bil.

## 2. Beregning av a, tg og tm

Det framgår av figuren at parameteren a er løsninga av $rct_i = a + gct_i$, dvs.

$$a = -(g-r)ct_i$$

Dermed er likningen for «startnummeret» på den bilen som passerer ut av hendelsestedet på tidspunkt t i køavviklingstida ($t_i < t < t_i + t_g$) lik

$$n = -(g-r)ct_i + gct \,.$$

Siden $t_i + t_g$ er løsninga av likningen $a + gc\left(t_i + t_g\right) = v\left(t_i + t_g\right)$, får vi ved innsetting av verdien for a:

$$t_g = \frac{v - rc}{gc - v} t_i$$

Tidspunktet da køen er helt avviklet, $t_i + t_g$, er dermed

$$t_i + t_g = \frac{g-r}{gc-v} ct_i$$

Den bilen som ankommer på dette tidspunkt, er den siste som blir berørt av hendelsen. Kall den N. Fra $n = vt$ får vi nå at det totale antall berørte biler er

$$N = \frac{\left(g-r\right)cv}{gc-v} t_i$$

Parameteren $t_m$, som representerer ankomsttidspunktet til den siste bilen som ankommer mens køen enda øker, er løsninga av $n = rct_i = vt_m$. Det gir

$$t_m = \frac{rc}{v} t_i$$

Inspeksjon av diagrammet viser at den lengste forsinkelsen noen bil opplever, er $t_i - t_m$. Vi får:

$$t_i - t_m = \frac{v - rc}{v} t_i$$

Den bilen som opplever den lengste forsinkelsen, kan vi kalle $n_m$. Den ankommer åpenbart på tidspunkt $t_m$. Av $n = vt$ får vi da:

$$n_m = rct_i$$

Den lengste køen finns på tidspunkt $t_i$, rett før avvikling av køen begynner. Køen på dette tidspunkt er $n_i - n_m$, siden $n_i$ har ankommet og $n_m$ har unnsluppet (se diagrammet). Vi har:

$$n_i - n_m = \left(v - rc\right)t_i$$

Vi har nå uttrykt alle ukjente størrelser i det kumulative plottet ved de kjente størrelsene, og har skaffet oss det som trengs til å beregne den totale forsinkelsen og fordelingen av forsinkelsene på de enkelte bilene.

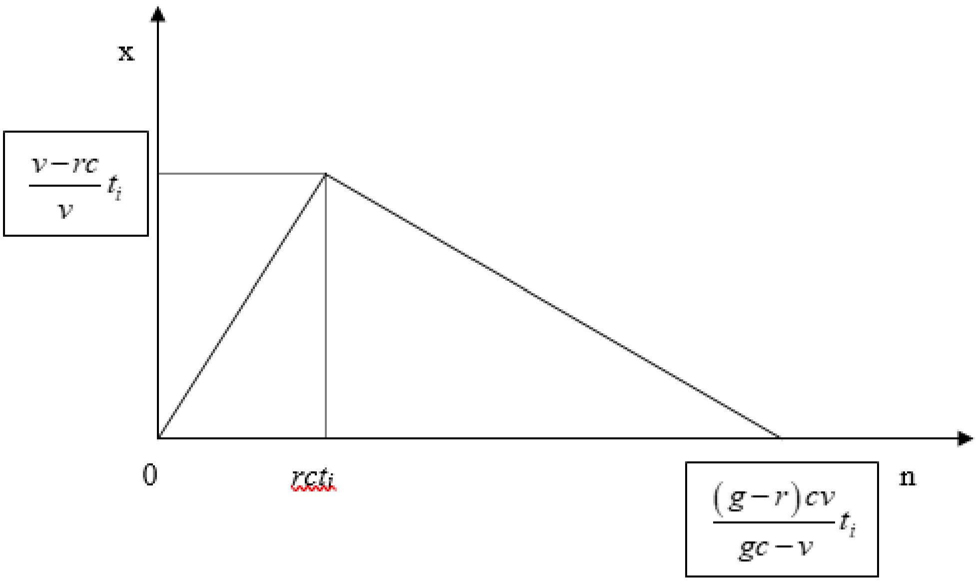
## 3. Total forsinkelse og fordeling på biler

En hendelse har inntruffet og det har tatt tida $t_i$ å rydde vegbanen. Den totale forsinkelsen D som de berørte bilene til sammen vil oppleve, er arealet som begrenses av de tre kurvene i figur 1. Vi har:

(1)
$$D=\int_0^{t_i}(v-rc)t\,dt+\int_{t_i}^{\frac{g-r}{gc-v}ct_i}\left\{vt-\left[-(g-r)ct_i+gct\right]\right\}dt$$
$$=\frac{1}{2}\frac{(g-r)(v-rc)}{gc-v}ct_i^2$$

Forsinkelsen som den enkelte bil opplever, stiger lineært fra 0 for den første bilen til et maksimumspunkt for bil nr. $rct_i$, og avtar deretter lineært til forsinkelsen 0 for den siste berørte bilen, bil nr. N. Kall forsinkelsen for den enkelte bil x. Av beregningene i forrige avsnitt har vi:

*Figur 2*




Enkel beregning av arealet til trekanten i figur 2 bekrefter at dette arealet er lik den totale forsinkelsen D.

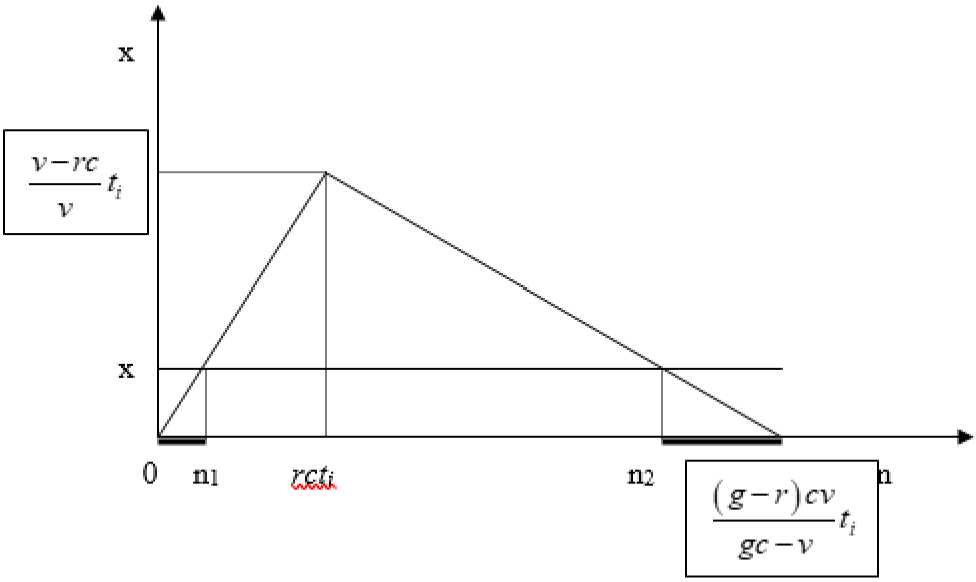
## 4. Sannsynlighetstettheten for den betingede sannsynligheten for forsinkelse, gitt at den oppstår, og gitt $t_i$

På en lenke av lengde 1 km er det en fast sannsynlighet λ for at en bil skal utløse en hendelse. Vi ser bort fra muligheten av at flere hendelser skal forsinke en og samme bil, og vi holder foreløpig $t_i$ fast. En bil kan da bli forsinket dersom det skjer en hendelse med en av de N bilene foran, der N er gitt i formelen ovenfor. Det er like sannsynlig at det skjer med den ene som den andre av disse N bilene. La Z være en stokastisk variabel som antar verdien 0 hvis en gitt bil ikke blir forsinket, og 1 hvis den blir det. La X være lengden av forsinkelsen – også det en stokastisk variabel.

Sannsynligheten for Z = 1 er $\lambda N=\lambda\frac{(g-r)cv}{gc-v}t_i$ og sannsynligheten for Z = 0 er $1-\lambda N=1-\lambda\frac{(g-r)cv}{gc-v}t_i$.

La oss nå finne den kumulative sannsynlighetsfordelingen for X, gitt at en forsinkelse inntrer, altså $\Pr(X \le x \mid Z = 1)$. Sannsynligheten for at den i så fall er mindre enn x, er sannsynligheten for at bilen som utløser en hendelse, er en bil tilstrekkelig nært foran vår bil eller tilstrekkelig nært etter bil nr. N foran vår bil. (Se figur 3, der en nivåkurve for x er tegnet inn i det samme diagrammet som figur 2, og n-verdier som gir lavere verdier enn x er tegnet med tjukkere strek).

*Figur 3*




Vi ser av figuren at $\Pr(X \le x \mid Z = 1)$ er lengden av de tjukke strekene, delt på N . Lengden av de tjukke strekene kan vi beregne om vi kjenner likningen for de to kurvene som avgrenser arealet. Ved hjelp av de kjente koordinatene for toppunktet og de to hjørnene på grunnlinjen finner vi at disse likningene er:

$$x = \frac{v - rc}{rcv} n$$

og

$$x = \frac{g - r}{g} t_i - \frac{gc - v}{gcv} n$$

og ved å løse disse for n finner vi n-verdiene for $n_1$ og $n_2$ i figuren. Den søkte sannsynligheten er nå:

$$
\begin{aligned}
&\Pr(X \le x \mid Z = 1) \\
&= \Pr\left(n \in \left[0, \frac{rcv}{v - rc} x\right]\right) + \Pr\left(n \in \left[\frac{gcv}{gc - v}\left(\frac{g - r}{g} t_i - x\right), \frac{(g - r)cv}{gc - v} t_i\right]\right) \\
&= \frac{\frac{rcv}{v - rc} x + \frac{(g - r)cv}{gc - v} t_i - \frac{gcv}{gc - v}\left(\frac{g - r}{g} t_i - x\right)}{\frac{(g - r)cv}{gc - v} t_i} \\
&= \frac{v}{v - rc} t_i^{-1} x
\end{aligned}
$$

Derivasjon av denne betingede kumulative sannsynligheten med hensyn på x gir den betingede sannsynlighetstettheten. Den er betinget av to ting – både at forsinkelse inntreffer og at $T_i = t_i$:

(2) $$f(x \mid Z = 1, T_i = t_i) = \frac{v}{v - rc} t_i^{-1}$$

## 5. Forventet hendelsesrelatert forsinkelse

Forventningen til x gitt $t_i$, $E[x|t_i]$, er ut fra det foregående:

$$
\begin{aligned}
&E[X \mid t_i] \\
&= E_Z[E_X(X \mid Z, T_i = t_i)] \\
&= (1 - \lambda N) E_X(X \mid 0, T_i = t_i) + \lambda N \cdot E_X(X \mid 1, T_i = t_i) \\
&= (1 - \lambda N) \cdot 0 + \lambda \frac{(g - r)cv}{gc - v} t_i \int_0^{\frac{v - rc}{v} t_i} \frac{v}{v - rc} t_i^{-1} x \, dx \\
&= \frac{1}{2} \lambda \frac{(g - r)(v - rc)}{gc - v} c t_i^2
\end{aligned}
$$

I utledningen har vi brukt regelen $E(X) = E_Y(E_X(X \mid Y))$. Anvender vi den en gang til, og bruker formelen for varians, har vi som et av våre hovedresultater:

(3) $$E(X) = \frac{1}{2} \lambda \frac{(g - r)(v - rc)}{gc - v} c\left[(ET_i)^2 + \operatorname{var} T_i\right]$$

Av formelen (3) framgår hva slags data som skal til for å anslå forventet hendelsesrelatert forsinkelse på en lenke av lengde 1 km. Vi ser blant annet betydningen av rask og sikker deteksjon av hendelsene.

Det framgår også av formelen at forsinkelsen er en funksjon av trafikken v på lenka. Funksjonen er tiltakende og konveks. Spørsmålet er nå om trafikantene vil regne med forventet forsinkelse i generaliserte kostnader og når de velger rute. Hvis ja, ser vi at sjøl uten kø i vanlig forstand vil brukerlikevekt avvike fra systemoptimum, og dermed vil forutsetningen i godsmodeller som SAMGODS være brutt.

Legg for øvrig merke til likheten mellom formel (1) og formel (3). Faktisk er $E(X) = \lambda E(D)$. Dette er ikke underlig. Den forventede forsinkelsen for en bil, multiplisert med antall biler N, bør jo i dette tilfellet være lik sannsynligheten for en hendelse med en av bilene, $\lambda N$, multiplisert med total forsinkelse om det skjer en hendelse. Det er mulig dette kan brukes til å anslå E(X) på grunnlag av statistikk over hyppigheten av hendelser og omfanget av tapt tid ved hendelser.

## 6. Den hendelsesrelaterte forsinkelsens varians

Vår neste oppgave er å finne var(X). Vi får da bruk for den generelle regelen $\text{var}(Y) = E_X\left[\text{var}_Y(Y \mid X)\right] + \text{var}_X\left[E_Y(Y \mid X)\right]$. Vi anvender den først med X og Z, idet vi i notasjonen på høyresida undertrykker at dette også dreier seg om sannsynligheter som er betinget av $T_i$. Vi har altså

$$\text{var}(X \mid t_i) = E_Z\left(\text{var}_X(X \mid Z)\right) + \text{var}_Z\left(E_X(X \mid Z)\right)$$

Her er

$$E_X(X \mid 0) = 0$$
$$E_X(X \mid 1) = \frac{1}{2}\frac{v - rc}{v} t_i$$
$$\text{var}_X(X \mid 0) = 0$$
$$\text{var}_X(X \mid 1) = \int_0^{\frac{v-rc}{v}t_i} \left[x - E_X(X \mid 1)\right]^2 f(x \mid 1)dx = \frac{1}{12}\left(\frac{v - rc}{v}\right)^2 t_i^2$$

Dermed er

$$E_Z[\text{var}(X \mid Z)] = \lambda N \cdot \frac{1}{12}\left(\frac{v - rc}{v}\right)^2 t_i^2$$

og

$$\text{var}_Z\left[E_X(X \mid Z)\right] = E_Z\left[\left(E_X(X \mid Z)\right)^2\right] - \left[E_Z\left[E_X(X \mid Z)\right]\right]^2$$
$$= \lambda N \frac{1}{4}\left(\frac{v - rc}{v}\right)^2 t_i^2 - \left(\lambda N\right)^2 \frac{1}{4}\left(\frac{v - rc}{v}\right)^2 t_i^2$$
$$= \left(1 - \lambda N\right)\lambda N \cdot \frac{1}{4}\left(\frac{v - rc}{v}\right)^2 t_i^2$$

Samlet har vi:

$$\text{var}(X \mid t_i) = \frac{1}{12}\lambda N\left(\frac{v - rc}{v}\right)^2 t_i^2\left(4 - 3\lambda N\right)$$
$$= \frac{1}{3}\lambda\frac{(g - r)(v - rc)^2}{gc - v}\frac{c}{v}t_i^3 - \left[\frac{1}{2}\lambda\frac{(g - r)(v - rc)}{gc - v}c\right]^2 t_i^4$$

Gjennom N er $\mathrm{var}(X \mid t_i)$ et fjerdegradsuttrykk i $T_i$.

Vi anerkjenner nå endelig at oppryddingstida er stokastisk, dvs. vi setter inn $T_i$ i stedet for $t_i$. Deretter anvender vi variansformelen enda en gang for å finne den siste størrelsen vi søker etter, nemlig var(X). Vi antar $T_i$ har en kjent fordeling med forventning $E(T_i)$ og varians $\mathrm{var}(T_i)$.

Vår regel i dette tilfellet:

$$\mathrm{var}(X) = E_{T_i}\left[\mathrm{var}_X(X \mid T_i)\right] + \mathrm{var}_{T_i}\left[E_X(X \mid T_i)\right]$$

Uttrykkene i klammeparentesen er allerede beregnet, og vi får:

$$\mathrm{var}_{T_i}\left[E_X(X \mid T_i)\right] = \left[\frac{1}{2}\lambda\frac{(g-r)(v-rc)}{gc-v}c\right]^2 \mathrm{var}\left(T_i^2\right)$$

$$= \left[\frac{1}{2}\lambda\frac{(g-r)(v-rc)}{gc-v}c\right]^2 \left(E(T_i^4) - \left(E(T_i^2)\right)^2\right)$$

og

$$E_{T_i}\left[\mathrm{var}_X(X \mid T_i)\right]$$

$$= \frac{1}{3}\lambda\frac{(g-r)(v-rc)^2}{gc-v}\frac{c}{v}E(T_i^3) - \left[\frac{1}{2}\lambda\frac{(g-r)(v-rc)}{gc-v}c\right]^2 E(T_i^4)$$

Til sammen:

$$(4) \quad \mathrm{var}(X) = \frac{1}{3}\lambda\frac{(g-r)(v-rc)^2}{gc-v}\frac{c}{v}E(T_i^3) - \left[\frac{1}{2}\lambda\frac{(g-r)(v-rc)}{gc-v}c\right]^2 \left(E(T_i^2)\right)^2$$

## 7. Oppsummering

Vi kan forenkle og sammenfatte våre resultater.

Sett

$$K = \frac{1}{2}\lambda\frac{(g-r)(v-rc)}{gc-v}c$$

Da er forventning og varians til den enkelte bilists forsinkelse på lenka:

$$(5) \qquad \mu = E(X) = K \cdot \left[\left(ET_i\right)^2 + \mathrm{var}\,T_i\right]$$

$$(6) \qquad \sigma^2 = \mathrm{var}(X) = \frac{2}{3}\frac{v-rc}{v}K \cdot E\left(T_i^3\right) - K^2 \cdot \left[\left(ET_i\right)^2 + \mathrm{var}\,T_i\right]^2$$

Fordelingen til $T_i$ avgjør om dette kan forenkles videre.

Forventning og varians for forsinkelsen på en lenke av lengde L fås ved å multiplisere med L.

Formel (5) samsvarer med Cohen og Southworth, mens formel (6) ikke gjør det. Ifølge Arup (2002), Appendix D, punkt D3.16, har Cohen innrømmet ved personlig kommunikasjon til Arup at hans formel er feil. Arups tilsvarende formel virker å være identisk med vår, men jeg har ikke studert Arups utledning og notasjon grundig.

## 8. Erlangfordelt oppryddingstid

Det er svært naturlig å anta at oppryddingstida $T_i$ er Erlangfordelt. Erlangfordelingen er Gammafordelingen i det tilfellet der en av parametrene (her kalt $R$) er heltallig. Den ble formulert første gang for hundre år siden i forbindelse med en analyse av behovet for kapasitet i Københavns telefonsentral, og skulle da representere sannsynlighetsfordelingen til varigheten av telefonsamtalene. Sannsynlighetstettheten er

$$f(x) = \frac{A^R}{\Gamma(R)} x^{R-1} e^{-Ax}, \quad x \geq 0, \; A > 0, \; R \in \{0,1,2,\ldots\}$$

der $A$ og $R$ er parametre som må tilpasses i det konkrete tilfellet, og $\Gamma(.)$ er Gammafunksjonen. I det heltallige tilfellet, som her, er

$$\Gamma(n) = (n-1)!$$

I utgangspunktet vil vi tro at $R$ rundt 10 vil passe i vårt tilfelle.

Forventning og varians har en enkel form:

$$EX = \frac{R}{A}, \quad \text{var}\, X = \frac{R}{A^2}$$

Erlangfordelingen er skeiv, og har toppunkt til venstre for forventningen. Toppunktet er

$$x^* = \frac{R-1}{A}$$

Ved for eksempel å bruke definisjonene av EX, $\mathrm{EX}^2$ og $\mathrm{EX}^3$ og anvende partiell integrasjon på dem, er det enkelt å vise at:

$$EX = \frac{R}{A}$$
$$EX^2 = \frac{R(R+1)}{A^2}$$
$$EX^3 = \frac{R(R+1)(R+2)}{A^3}$$

osv.

Vi kan nå anvende disse resultatene i formlene (5) og (6) for forventet hendelsesrelatert forsinkelse pr. kilometer, $\mu$, og for variansen til forsinkelsen, $\sigma^2$. Vi får:

(7) $$\mu = \frac{1}{2} \frac{R(R+1)}{A^2} \cdot \lambda \frac{(g-r)(v-rc)}{gc-v} c$$

(8) $$\sigma^2 = \mu \left( \frac{2(R+2)}{3A} \frac{v-rc}{v} - \mu \right)$$

## 9. Parametrene g og r

Cohen og Southworth har estimert $g$ til 1.07. Arup påpeker at hvorvidt $g$ kan bli større enn 1, avhenger av definisjonen av kapasitet, $c$. Det er ikke kjent hva Cohen og Southworth har gått ut fra når det gjelder dette. De argumenterer med at bilene som starter opp fra køen kjører tettere enn det de ellers ville gjort. Uansett mener vi det er urimelig å sette $g$ til mer enn 1, men det er heller ingen gode argumenter for å sette $g$ til mindre enn 1. Vi går derfor ut fra at vi kan sette $g = 1$.

Vi anser det som åpenbart at parameteren $r$ vil avhenge av vegkapasiteten $c$, som igjen avhenger av antall filer, filenes bredde og vegskulderens bredde. På en trang bygdeveg kan en hendelse blokkere vegen i begge kjøreretninger. På en vanlig tofeltsveg vil en hendelse ofte blokkere hele den ene kjøreretningen, og $r$ vil i så fall være mindre enn 0.5, avhengig av vegskulderens bredde, sikten i kjøretningen og volumet på den motgående trafikken. Ved flere filer i samme kjøretning vil det bare være alvorlige hendelser som blokkerer alle disse filene.

Vi kan derfor anta at $r$ som funksjon av $c$ er 0 når $c$ går mot 0, og 1 når $c$ går mot uendelig, og at den er en monotont tiltakende og konkav funksjon. Mange funksjoner kan tenkes å representere dette. Blant dem er mange kumulative sannsynlighetsfunksjoner, hyperbler med vannrett asymptote, og hyperbolsk tangens. (Et tilleggskrav kan være at $c – rc$ aldri faktisk blir 0, siden hendelsen i alle fall vil legge beslag på noe kapasitet.)

Vi betrakter først funksjonen

$$f(x) = \frac{x}{a+x}, \quad x \geq 0$$

Denne funksjonen er alltid mellom 0 og 1. Den antar verdien 0 i punktet $x = 0$, og går mot 1 når $x$ går mot uendelig. Den har positiv førstederivert og negativ andrederivert. Grensa for $x - xf(x)$ når $x$ går mot uendelig er $a$. I vårt tilfelle kan vi anvende funksjonen

$$r = f(c) = c(a+c)^{-1}$$

Parameteren $a$ i dette tilfellet har da en tolkning som den kapasiteten som en hendelse i gjennomsnitt vil legge beslag på på en brei veg.

Denne funksjonen er meget enkel. Det er naturligvis et problem for tilpasning til den virkelige kurven for $r$ at funksjonen bare har en parameter. Om den egner seg i virkeligheten, gjenstår derfor å se.

En annen funksjon som også bare tillater innarbeiding av en parameter, er den hyperbolske tangens,

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Innsetter vi her parameteren $a$ og anvender funksjonen på $ac$, kan vi anta

$$r = f(c) = \frac{e^{ac} - e^{-ac}}{e^{ac} + e^{-ac}}$$

Også denne funksjonen er alltid mellom 0 og 1. Den antar verdien 0 i punktet $c = 0$, og går mot 1 når $c$ går mot uendelig. Den har positiv førstederivert og negativ andrederivert. Grensa for $c - cf(c)$ når $c$ går mot uendelig er imidlertid 0. Parameteren $a$ har ikke her noen naturlig tolkning, og kan tilpasses fritt for å gi riktig kapasitetsrest for en bestemt vegtype. Forhåpentligvis vil den da også tilnærmet gi riktig verdi av $r$ for andre vegtyper.

En av de mest aktuelle kumulative sannsynlighetsfordelingene som kan brukes for $f(c)$ er den kumulative Weibullfordelingen:

$$r = 1 - e^{-ac}, \quad a > 0$$

Hvis ikke noe av dette skulle passe, får vi se oss om etter en kumulativ sannsynlighetsfordeling med to parametre.

Uansett valg av funksjon *f(c)* kan vi etter dette skrive:

$$\mu = \frac{1}{2}\lambda \cdot \frac{R(R+1)}{A^2} \cdot \frac{(1-f(c))(v-cf(c))}{c-v}c \tag{9}$$

$$\sigma^2 = \mu\left(\frac{2(R+2)}{3A}\frac{v-cf(c)}{v} - \mu\right) \tag{10}$$

Formel (9) og (10) med en høvelig funksjon *f(c)* for kapasitetsresten vil danne utgangspunktet for resten av arbeidet i prosjektet.

La oss kontrollere at forventet forsinkelse i timer pr. kilometer, $\mu$, har benevning som forutsatt. En enkel dimensjonsanalyse viser at $\lambda$ bør være benevnt $km^{-1}$, *R/A* og *(R+1)/A* bør være benevnt i timer, *f(c)* er dimensjonsløs og *v* og *c* er benevnt $timer^{-1}$. Totalt blir det timer pr. kilometer. Dette må multipliseres med antall kilometer. Når det gjelder variansen, vil vi anta at begivenheten ”hendelse på den ene kilometeren” er uavhengig av begivenheten ”hendelse på den forrige kilometeren”. Dermed blir variansen for flere kilometer lik summen av variansene pr. kilometer. Begge størrelser må altså multipliseres med transportens lengde.

I et nettverk vil hver lenke ha sin egen kapasitet, lengde og lenketrafikk. Forsinkelsene i dette tilfellet vil vi behandle videre i et seinere dokument.

## 10. Mange typer hendelser

Det finns mange ulike typer av hendelser, og hver kan antas å ha sin egen gjennomsnittlige verdi for $\lambda$, *r* og $T_i$. Vi antar imidlertid at det ikke finnes data som kan brukes til å estimere forventning og varians for mange typer hendelser. Hvis vi opererer med bare en type hendelser, er hendelsesfrekvensen $\lambda$ summen av frekvensene av alle hendelser der det faktisk skjer en forsinkelse, dvs. der $v \geq cf(c)$. Sannsynlighetsfordelingen til $T_i$ vil ta hensyn til alle typer hendelser. Dette skulle være uproblematisk. Problemet er at *r* i virkeligheten vil ha stor spredning. Hvor stor feil det bringer inn i formlene, bør undersøkes ved følsomhetsanalyse.

Det er viktig når vi henter inn data at vi klarer å skille ut og ikke ta med tilfellene der $v < cf(c)$

.

## 11. Det breiere bildet

Hendelsesrelatert forsinkelse av det slaget vi har behandlet her, er ikke den eneste årsak til forsinkelse. La oss kort kommentere noen andre situasjoner som kan oppstå.

1. Det er mulig at en hendelse ikke fører til kø på hendelsesstedet, men bare redusert fart over en større eller lengre strekning. Redusert fart vil også ofte forekomme ved ”planlagte” hendelser som vegarbeid, hvor farten ofte er skiltet ned, eller ved dårlig vær. Det er mulig at det kan lages en generell teori som omfatter både nedsatt fart over en strekning og nedsatt fart forbi et punkt og kø ved dette stedet.
2. Hvis tida det tar å rydde vegbanen er lang, vil bilistene lete etter andre tilpasninger. De som står i kø, vil snu og forsøke å finne andre ruter. De som ennå ikke er ankommet, vil

bli varslet av myndighetene og velge andre ruter uten å kjøre inn på den stengte lenka. De som enda ikke har startet, vil kunne revurdere avreisetidspunktet i lys av den nye situasjonen.

3. En hendelse kan ramme egen bil eller sjåfør, og vil da ofte være mer langvarig og ha mer omfattende konsekvenser for transporttida. Sjøl om dette forekommer sjeldnere, kan det bidra vesentlig til variabilitet i transporttida. En ulykke kan involvere vår bil direkte. Bilen kan gå tom for bensin, punktere eller få motorhavari. Sjåføren kan ta feil av vegen. Sjukdom eller presserende personlige behov eller interesser kan få sjåføren til å avbryte transporten, stoppe for kortere eller lengre tid eller ta en omveg.
4. Myndighetene kan gripe inn med midlertidige restriksjoner, nedskriving av fartsgrenser, vegarbeid, kontroller og eventuelt avskilting. Transportbedriften kan omdirigere bilen til viktigere eller mer presserende oppdrag. Været kan snu.
5. I noen tilfeller kan en ta inn igjen en oppstått forsinkelse ved å kjøre fortere eller ta mindre stopp enn det som normalt er forsvarlig. Muligheten for dette er mindre på korte turer og forsvinner langt på veg hvis farten er bestemt av trafikken, dvs. ved kø. Kø virker derfor på mange måter: Den inngår i formlene for hendelsesrelatert forsinkelse i form av avstanden mellom v og c. Den har implikasjoner for mange flere trafikanter når omkjøring på købelastede veger blir aktuelt. Og den bestemmer mulighetene til å ta det tapte inn igjen.

## Litteratur

ArupTransportPlanning (2002) *Journey Time Variability. Deliverable D6.1: Modelling and Appraisal of Journey Time Variability – Review of Earlier research, Research on this Contract, and Detailed Proposals for Further Research.* Downloaded from the net.

Cohen, H. and F. Southworth (1999) On the Measurement and Valuation of Travel Time Variability Due to Incidents on Freeways. *Journal of Transportation and Statistics*, December 1999.

Newell, G.F. (1982) *Applications of queueing theory ($2^{nd}$ edition).* Chapman Hall, London.

## 4.3 Statistiske stordriftsfordeler i kollektivselskaper[76]

[76] Upublisert arbeidsdokument fra 1997, TØ 1054/97.

# Statistiske stordriftsfordeler i kollektivselskaper

## Innhold

# 1. Innledning

Stordriftsfordeler (stigende skalautbytte) i kollektivselskaper er i utgangspunktet et mangetydig begrep, fordi det kollektivselskapene produserer er en eller flere transporttjenester, hver med mange egenskaper. Med den liberaliseringen i næringen som har funnet sted i mange land i de seinere åra, ser vi imidlertid ofte at selskaper fusjonerer, slik at markedet i et enkelt land domineres av for eksempel 3-5 store selskaper. Dette tyder på at det fantes en eller annen form for stordriftsfordeler ved det produksjonsnivået som de fleste kollektivselskapene hadde for kort tid sida.

Vi sparer til en annen gang å definere helt nøyaktig hva vi bør mene med kollektivselskapets tjenesteproduksjon og med et produksjonsnivå i denne forbindelsen. For vårt formål er det tilstrekkelig å peke på følgende fire viktige indikatorer på tjenesteproduksjonen til et kollektiv-selskap:

1. Passasjerer pr. tidsenhet og flateenhet
2. Gjennomsnittlig reiselengde pr. passasjer
3. Gjennomsnittlig fart
4. Flatedekning

Holder vi indikator 2-4 konstante, vil en økning i indikator 1 (økt etterspørselstetthet) gi muligheter til reduserte kostnader pr. passasjer gjennom bruk av større kjøretøyer og gjennom mer effektive ruteopplegg[77]. Med de etterspørselstettheter som er vanlige i Norge, vil det dermed sannsynligvis finnes stordriftsfordeler med hensyn til den første indikatoren.

Økt produksjon av indikator 2 for konstant nivå på indikator 1, 2 og 4 vil nok redusere gjennomsnittskostnaden pr. passasjerkilometer. Økt produksjon av indikator 3 for konstant nivå på de øvrige vil kanskje øke gjennomsnittskostnaden pr. fartsenhet på grunn av økt drivstoff-forbruk, men redusere den på grunn av redusert behov for materiell og mannskaper. Total-virkningen på kostnadene av endringer i disse indikatorene er usikker og situasjonsavhengig, og vil ikke bli vurdert videre her.

Vår interesse i dette notatet er knyttet til eventuelle stordriftsfordeler mhp. indikator 4. I forbindelse med fusjonsplaner i bussnæringen er det pekt på innsparingsmuligheter i innkjøp og administrasjon som begrunnelse for planene, og som argument for at de ikke har sitt grunnlag i ønsker om monopolmakt. Disse fusjonsplanene har involvert busselskaper som opererer i forskjellige byer og fylker. I mindre grad har de involvert to eller flere busselskaper i samme by. Problemstillingen i den enkelte by har ofte vært den motsatte, nemlig oppsplitting av produksjonen på flere selskaper. I den forbindelsen har man imidlertid ofte uten videre antatt at det *ikke* eksisterer stordriftsfordeler ved flatedekning.

Utenom stordriftsfordeler i innkjøp og administrasjon vil det også eksistere en stordriftsfordel knyttet til reparasjoner og vedlikehold, så rart det kan høres. Dels skyldes den at en stor vognpark gir bedre grunnlag for spesialisering og arbeidsdeling i vedlikeholdsverkstedet, dels at variansen til den usikre variablen "andel av vognparken som trenger reparasjon" blir mindre, ifølge de store talls lov. Dette gir enten jamnere beskjeftigelse på vedlikeholdsverkstedet eller

[77] Det kan imidlertid tenkes at gjennomsnittsfarta går ned når antall på- og avstigninger pr. stoppested går opp, og at antall stoppesteder pr. linje samt gjennomsnittlig reiselengde pr. passasjer går ned. Produksjonen av en av indikatorene er ikke uavhengig av produksjonen av de andre.

kortere gjennomsnittlige reparasjonstider. Men det gir også en annen viktig effekt, nemlig et mindre behov for å holde en reservekapasitet av kjøretøyer som sikkerhet mot at materiellet bryter sammen. Hvis altså to busselskaper i samme by slår seg sammen, vil en kunne klare seg med færre busser enn før, fordi en kan redusere den relative reservekapasiteten av busser ut over det antallet som er i bruk hver dag, uten dermed å løpe økt risiko for å måtte innstille ruter på grunn av materiellmangel.

Et tilsvarende forhold vil også gjelde for behovet for sjåfører, som vil være påvirket av den usikre variablen ”andel sjukmeldte”.

Det finnes en kostnad som vil motvirke slike innsparinger. Det er transportkostnaden mellom de to markedene som tidligere blei betjent med hver sin bussflåte. Jo mer kostbart det er å overføre en buss fra det ene til det andre markedet, jo troligere er det at det som vinnes på en mindre bussflåte, tapes i form av transportkostnader ved slike overføringer, som kan måtte gjøres relativt hyppig for å møte svingninger i antall disponible kjøretøyer i hvert marked når reservekapasiteten i begge markeder er redusert. Transportkostnaden fram og tilbake til vedlikeholdsverkstedet er også relevant. Hvis det nye verkstedet for eksempel ligger der et av de gamle lå, vil transportkostnaden til og fra verksted trolig øke for busser fra det andre markedet.

Stordriftsfordeler som skyldes at variansen til variablen ”andel av vognparken som trenger reparasjon” blir mindre, vil vi kalle ”statistiske stordriftsfordeler”. Notatet vil gi et grunnlag for å beregne dem. Dette grunnlaget er lagt i litteraturen omkring ”kvadratrotlova” og ”portefølje-effekten” i logistikk. Kvadratrotlova slik den blei foreslått av Maister (1976) er bare gyldig under spesielle vilkår. Eppen (1979) utledet en litt mer generell versjon av kvadratrotlova av avisgutt-problemet. Zinn et al (1989) definerte portefølje-effekten som en generalisering av kvadratrotlova. Tallon (1993) er en grei innføring i begge deler. Han tar også hensyn til at ledetida er usikker. Chang og Lin (1991) introduserer overføringer mellom lagre. Endelig introduserer Das og Tyagi (1997) transportkostnader i problemet. Flere referanser er gitt i Das og Tyagi.

I kapittel 2 begrunner vi at denne litteraturen lar seg anvende på problemet om stordriftsfordeler i kollektivselskaper, og gir de viktigste formlene. I kapittel 3 drøfter vi betydningen av statistiske stordriftsfordeler.

Det kan være på sin plass å understreke at de statistiske stordriftsfordelene er svært avhengig av at det ikke finns muligheter til å substituere selskapets busser med annet tilgjengelig materiale. Det kan være fordi selskapet har som politikk at alle busser skal ha selskapets farger, eller i sporvognstilfellet at det ikke eksisterer et leiemarked for det materiellet som kan brukes. Tilsvarende gjelder det at det bare er statistiske stordriftsfordeler på sjåførsida dersom vikarer ikke kan brukes uten etter en lengre opplæring.

## 2. Portefølje-effekten og dens virkning ved fusjonering av kollektivselskaper

### 2.1 Formelen for portefølje-effekten

La $\sigma_1, \ldots, \sigma_n$ være standardavvikene til n stokastiske variable $X_1, \ldots, X_n$, og betegn korrelasjonskoeffisienten mellom $X_i$ og $X_j$ med $\rho_{ij}$. Korrelasjonskoeffisienten antar verdier mellom –1 og 1. Pr. definisjon er $\rho_{ij} \equiv \mathrm{cov}(X_i,X_j)/(\sigma_i\sigma_j)$. Formelen for variansen $\sigma^2$ til summen $X = \sum X_i$ kan derfor skrives

(1) $$\sum_i \sigma_i \geq \sigma = \sqrt{\sum_i \sigma_i^{\,2} + 2\sum_{i>j}^{n}\sum_{j=1}^{n-1}\sigma_i\sigma_j\rho_{ij}}$$

*Bevis:* Begge sider i (1) er positive. Ulikheten er derfor ekvivalent med den ulikheten som framkommer ved å opphøye begge sider i annen. Siden $\rho_{ij}$ er mindre eller lik 1, følger (1).

La nå de n stokastiske variable $X_1,\ldots,X_n$ være etterspørselen pr. tidsenhet etter en og samme vare i n adskilte markeder, hvert og ett av dem betjent fra et eget lager. Bedrifter vil ofte følge en leveringsdyktighetspolitikk som innebærer at sikkerhetslageret er proporsjonalt med standardavviket til etterspørselen pr. tidsenhet. Vi ser umiddelbart av (1) at under en slik politikk vil sikkerhetslageret bli mindre om de n markedene kan betjenes fra samme lager.

Vi antar at kravet til leveringsdyktighet er det samme uansett lagerlokalisering. Hvis alle korrelasjonskoeffisientene $\rho_{ij}$ er null og alle variansene er like, gir forholdet mellom venstresida og høyresida av ulikheten i (1) da et veldig enkelt uttrykk for hvor mye sikkerhetslageret kan reduseres ved sentralisering av lageret til ett sted. Vi får nemlig at det samlede sikkerhetslageret ved desentralisert lagring på n steder, er $\sqrt{n}$ større enn det sentraliserte sikkerhetslageret. Det er dette som kalles ”kvadratrotlova”.

Den prosentvise reduksjonen i det samlede sikkerhetslageret når n adskilte lagre sentraliseres til ett, kalles portefølje-effekten. Portefølje-effekten lar seg beregne i det helt generelle tilfellet med korrelasjonskoeffisienter som ikke er null og med ulike varianser.

Med vår notasjon vil den være:

(2) $$PE = 1 - \frac{\sigma}{\sum \sigma_i} = 1 - \frac{\sqrt{\sum_i \sigma_i^{\,2} + 2\sum_{i>j}^{n}\sum_{j=1}^{n-1}\sigma_i\sigma_j\rho_{ij}}}{\sum_i \sigma_i}$$

Forutsetningene for definisjonen av porteføljeeffekten PE ifølge formel (2) er som nevnt: 1. Samme krav til leveringsdyktighet uansett lagerlokalisering. Dessuten forutsettes: 2. Sentralisering av lageret påvirker ikke usikkerheten i leveransene *til* lageret (dette vil være oppfyllt hvis ledetida inn til lagrene ikke er usikker). 3. Overføring mellom lagrene når det er flere lagre er ikke mulig eller vanlig.

For tilfellet n = 2 kan PE uttrykkes enkelt som en funksjon av $\rho_{12}$ og M, der $M = \sigma_1/\sigma_2$. Formelen blir:

(3) $$PE = 1 - \frac{\sqrt{M^2 + 1 + 2M\rho_{12}}}{M+1}$$

### 2.2 ”Sikkerhetslagre” av vogner i et kollektivselskap

Vi ser på en gitt mengde av separate kollektivselskaper, og vil finne ut om det er stordriftsfordeler knyttet til å slå sammen materiellet og bruke det under ett. Vi betrakter disse kollektivselskapene i en periode der ruteplanene er gitt, og som dessuten er så kort at det er umulig eller uaktuelt å få inn mer materiell i systemet som helhet. Hvor lang denne perioden er, vil avhenge mye av kollektivselskapenes situasjon og politikk. I noen tilfeller vil en ikke bry seg om at bussen ikke har selskapets farger og utforming for øvrig, og vil da kunne få tak i ekstramateriell raskt ved å leie fra andre selskaper[78]. I andre tilfeller kan det være snakk om år, nemlig når materiellet må produseres på bestilling.

Etterspørselen etter materiellet til vanlig drift over denne perioden er i hovedsak deterministisk og gitt av ruteplanene. I vår sammenheng kan vi se bort fra denne deterministiske komponenten av etterspørselen. Men i tillegg finns det en stokastisk etterspørselskomponent, nemlig ”etterspørsel” etter materiell til verksted. Et visst antall av bussene eller vognsettene vil bryte sammen i løpet av perioden og trenge reparasjon. For å møte denne ”etterspørselen” trenger vi et ”lager” av busser eller vognsett ut over det antallet som trenges i rute. Hvis etterspørselen ikke kan møtes, oppstår det et tap i form av avganger som må innstilles.

Anta nå at vi ut fra erfaring kan angi en normalfordeling for det antallet av busser eller vognsett i hvert selskap som står på verksted i en uke[79]. Vi antar at de stokastiske variablene $Y_{kj}$ = ”antall busser i selskap j som står på verksted i uke k” er uavhengige og identisk fordelte for alle ukene j = 1, …, J i perioden (noe som vel forutsetter reparasjonstider på en uke eller mindre og ingen sesongsvingninger i reparasjonsbehovet). Altså: $Y_{kj} \sim N(\mu_{kj}, \sigma_{kj})$ for alle k og j, eller fordi vi like gjerne kan sløyfe indeksen k for uke: $Y_j \sim N(\mu_j, \sigma_j)$.

Kollektivselskapets leveringsdyktighetspolitikk i perioden kan da defineres med hensyn på hvor stor sannsynlighet vi maksimalt tillater at det skal være for at avganger må innstilles på grunn av materiellmangel i en slik typisk uke. La oss si at denne sannsynligheten er fastlagt som p %. Vi definerer $Z_j = (Y_j - \mu_j)/\sigma_j$. $Z_j$ er fordelt N(0,1), og $Y_j = \mu_j + Z_j\sigma_j$. Kaller vi antall busser ut over det det er bruk for i rute for B, vil vi altså ha sannsynligheten for at $Y_j > B$ mindre enn p%. Vi oppnår det ved å sette

der N er mengda av naturlige tall og $\bar{z}_j$ er den øvre p-fraktilen til standard-normalfordelingen.

$$B = \min\left\{\xi \middle| \xi \in N \,\&\, \xi \geq \mu_j + \bar{z}_j \sigma_j \right\}$$

Den er tabulert i alle vanlige lærebøker i statistikk.

Vi ser at bortsett fra heltallighetsproblemer, vil ”lageret” av busser B bestå av forventet antall busser på reparasjon i enhver uke, pluss et ”sikkerhetslager” som er proporsjonalt med standard-avviket $\sigma_j$. Det er dette sikkerhetslageret som kan reduseres ved sammenslåing av kollektiv-selskapene eller deres materiell. Dersom heltallighet blir mindre av et problem i det sammen-slåtte selskapet, vil dette også bidra til stordriftsfordelene.

Det kan tenkes at på grunn av felles ytre vilkår som vær og føre vil det være en viss korrelasjon mellom de stokastiske variablene $Y_j$ for de ulike selskapene. Men grovt sett kan vi nok anta uavhengighet mellom dem. Det betyr at korrelasjonskoeffisientene mellom dem er null. Videre kan vi nok som regel anta at variansen til $Y_j$ er lik for alle j, med mindre noen av selskapene har kjøpt inn notorisk upålitelig materiell. Dette betyr at dersom alle selskapene er like store og har samme politikk mht. leveringssikkerhet, vil kvadratrotlova gjelde for reduksjonen i sikkerhetslagerets størrelse ved en fusjon. Omvendt vil kvadratrotlova også angi hva det koster å splitte opp materiellparken under en gitt leveringssikkerhetspolitikk. Når kvadratrotlova ikke antas å gjelde, vil en kunne bruke formel (2) eller (3).

---

[78] Hvis en raskt kan få inn ekstramateriell fra andre selskaper innen den mengden av selskaper vi ser på, har vi tilfellet av overføring mellom lagre, som vi så bort fra i forrige avsnitt. Vi behøver ikke nødvendigvis utelukke denne muligheten, men vi har definitivt utelukket muligheten for å få materiell utenfra inn i systemet vi ser på i den perioden som er definert.

[79] For å kunne bruke normalfordeling her, må det være svært usannsynlig at antallet er null.

Ved hjelp av Das og Tyagi (1997) vil en dessuten kunne beregne innsparingene som kan oppnås ved sentralisering også når det i utgangspunktet er mulig å overføre materiell mellom de desentraliserte lagrene til en kostnad (transportkostnad eller transaksjonskostnad). Videre vil en kunne ta hensyn til endring i transportkostnader på grunn av ny lokalisering av lagret (buss-stallen) i forhold til kundene (verkstedene). Vi går ikke nærmere inn på dette her.

## 2.3 Optimal leveringssikkerhet

Den litteraturen vi bygger på, gir ikke bare mulighet til å kalkulere innsparingspotensialet med hensyn til bussparken ved sammenslåing av kollektivselskaper når leveringssikkerhets-politikken er gitt, men også en mulighet til å kalkulere optimal leveringssikkerhetspolitikk, dvs. optimal p i problemstillingen fra forrige avsnitt. Forutsetningen er at det er mulig å sette kroner og øre på et gjennomsnittlig tilfelle av leveringssvikt. I vår definisjon oppstår leveringssvikt når avganger må innstilles i løpet av en uke. Fordi det kan dreie seg om mange avganger, må en anslå et gjennomsnitt av avlyste avganger pr. forekomst av leveringssvikt i en uke. Deretter må en avlyst avgang verdsettes. For kollektivselskapet vil den koste tapte billettinntekter og tap av goodwill i markedet. Dersom en også skal inkludere tapet for kundene, vil det hovedsakelig bestå av ventetider på neste avgang.

Gitt at en kan verdsette avlyste avganger, er problemstillingen å velge riktig leveringssikker-hetsnivå og dermed busspark lett å formulere som et såkallt ”avisgutt-problem”. Løsningen av avisgutt-problemet for hvert av selskapene er å velge det bestillingskvantum (busspark) $B_j$ som tilfredsstiller

$$(4) \qquad B_j = \min\left\{\xi \,\middle|\, \xi \in N \;\&\; \Phi_j(\xi) \geq \frac{h}{h+k}\right\}$$

der $\Phi_j(.)$ er den kumulative fordelingsfunksjonen for etterspørselen pr. uke etter busser til verkstedet for selskap j, h er den faste kostnaden pr. uke ved å eie eller disponere en buss som ikke blir brukt, og k er tapet pr. tilfelle av leveringssvikt.

Som vist i Eppen (1979) kan vår formel (1) og alle de videre konsekvenser av den – inkludert formlene (2) og (3), sjøl om portefølje-effekten ikke var definert på det tidspunktet Eppen skreiv – utledes på grunnlag av avisgutt-problemet. I tillegg får man altså også definert den optimale leveringssikkerhetspolitikken p. Forutsetningen for overgangen fra løsninga på avisgutt-problemet til konsekvensene av sentralisering av lageret er imidlertid at h og k er konstanter som ikke varierer over selskapene.

## 2.4 Kommentar om effekten av sentralisering på behovet for sjåfører

Innledningsvis var vi inne på at det samme resonnementet som her er gjennomført mht. bussparken, også kunne vært gjennomført mht. sjåførene. I enhver uke vil noen av dem være sjuke, og det gir grunnlag for å holde seg med et ”sikkerhetslager” av sjåfører, som imidlertid kan reduseres ved sentralisering. Dette behovet oppstår imidlertid bare hvis det er dyrt eller umulig å skaffe vikarer i perioden vi ser på. Med vikarer mener vi da folk som ikke koster selskapet noe når de ikke benyttes. Hvis vikarer er billig og mulig, trenges inget ”sikkerhets-lager”, og det er følgelig heller ingen stordriftsfordeler ved sentralisering. Hvis midlertidige overføringer fra andre selskaper blant den mengden av selskaper vi betrakter, er billig og mulig, vil det heller ikke være noen gevinst ved sentralisering.

Vi ville ikke visst hvordan vi skulle håndtere problemstillingen med behov for sikkerhetslager av *både* busser og sjåfører. Derfor vil vi anta at det er mulig å holde seg med vikarer.

## 3. Betydningen av statistiske stordriftsfordeler

Om de effektene vi har behandlet her, betyr mye eller lite i praksis, er et empirisk spørsmål som vi ikke har sett på. Men gitt at de betyr en del, er det likevel ikke gitt at kollektivselskaper bør fusjonere av den grunn. Et alternativ er åpenbart å arbeide med å endre de forutsetningene som ligger til grunn for våre konklusjoner. Det er mulig å inngå et samarbeid om materiell og mannskap uten fusjonering. Dersom det er spesielle krav til utstyr eller mannskap som gir opphav til de prohibitivt høye transaksjonskostnadene ved å skaffe ekstra ressurser etter behov midt i perioden, er det mulig å endre disse kravene. Dersom en buss ikke behøver å ha selskapets farger, er det åpenbart enklere å skaffe erstatningsmateriell i et påkommende tilfelle. Alternativt kan myndighetene kreve at alt materiell i kollektivtrafikken i en by skal ha samme farger og standarder.

Dersom myndighetene ikke bryr seg med å redusere transaksjonskostnadene ved å skaffe ekstra resurser raskt i påkommende tilfeller, eksisterer en reell fare for at oppsplitting av kollektiv-tilbudet i en by på mange selskaper gjennom anbud vil ha negative effekter. Disse effektene kan enten komme i form av lavere servisgrad (flere innstilte ruter) eller i form av høyere enhetskostnader. Effektene kan delvis motvirkes gjennom den tendensen vi observerer til at det dannes store kollektivselskaper som har virksomhet i mange byer. Det vil da være mulig å overføre ressurser fra by til by internt i selskapene, men til en viss kostnad (”transport-kostnadene” i Das og Tyagi (1997)).

Det er grunn til å understreke at i et system med anbud på enkeltruter eller enkeltområder i en by, vil det måtte være myndighetenes ansvar å ta hensyn til disse effektene. Situasjonen som myndighetene skaper, kan være en av de følgende.

1. Myndighetene tillater ikke eksistende operatører å by på nye ruter eller områder, eller inn-fører regler som diskriminerer eksisterende operatører. Alle *nye* operatører som leverer inn bud, vil da ha samfunnsøkonomisk unødig høye kostnader, eller vil legge seg på et samfunnsøkonomisk uakseptabelt servisnivå. I den grad det siste forhindres gjennom kvalitetskrav, får vi det første.

2. Myndighetene tillater eksisterende operatører å by, med resultat at disse vil ha større mulig-het enn de andre til å vinne anbudet og opprettholde sin monopolstilling.

3. Anbydere vil arbeide for en kopling av rutene eller områdene, og stille krav om at de får flere områder som betingelse for å opprettholde sine bud. Det oppstår fare for samrøre mellom myndighetene og selskapene. Alternativt vil vi få en fusjonering av selskaper som har vunnet anbud i samme by, etter anbudsrundene.

4. Myndighetene fjerner eller reduserer de statistiske stordriftsfordelene gjennom tiltak som letter mulighetene for billig og enkel tilførsel av ekstrakapasitet når uventede behov for det oppstår. Dette kan være felles standarder for alt materiell, og andre tiltak som letter fram-veksten av et fritt marked for ekstramateriell. Alternativt kan myndighetene kreve at alt materiell leies inn av selskapene fra en felles pool. (Andre former for stordriftsfordeler, som fordeler ved å kombinere ruter, vil imidlertid være vanskeligere å ”flytte ut” fra selskapenes tilpasning).

I den engelske jernbanereformen er det opprettet en håndfull selskaper som skal leie ut rullende materiell. Her ser vi på den ene sida at innsparingsmulighetene ved å ha bare ett slikt selskap er ignorert til fordel for konkurransehensynene. Dette må bety at det totalt sett holdes for mye jernbanemateriell på lager. På den andre sida er dette nå en stordriftsfordel som foreligger som potensiale blant selskapene som leier ut materiell, og ikke blant operatørselskapene. Den vil derfor ikke påvirke operatørselskapenes tilpasning ved innlevering av anbud. Dersom selskapene som leier ut materiell også kan leie ut til hverandre, vil potensialet for stordriftsfordeler blant disse selskapene til en viss grad være realisert. Jeg veit ikke om det er slik.

Den typen av problemer vi har berørt, kan også vise seg å være relevant på andre områder enn samferdsel, for eksempel når det gjelder søppeltømming.

## Litteratur


Chang, P. og C.T. Lin (1991): On the effect of centralization on the expected costs in a multi-location newsboy problem. *Journal of the Operational Research Society* **42** side 1025-1030.

Das, C. og R. Tyagi (1997): Role of inventory and transportation costs in determining the optimal degree of centralization. *Transportation Research* **33E** (3) side 171-179.

Eppen, G.D. (1979): Effects of centralization on expected costs in a multi-location newsboy problem. *Management Science* **25** (5) side 498- 501.

Maister, D.H (1976): Centralization of inventories and the "square root law". *International Journal of Physical Distribution and Materials Management* **6** side 124-134.

Tallon, W.J (1993): The impact of inventory centralization on aggregate safety stock: the variable supply lead time case. *Journal of Business Logistics* **14** (1) side 185-203.

Zinn, W., M. Levy og D.J. Bowersox (1989): Measuring the effect of inventory centralization/decentralization on aggregate safety stock: The "square root law" revisited. *Journal of Business Logistics* **10** (1) side 1-14.

TØI er et anvendt forskningsinstitutt som mottar basisbevilgning fra Norges forskningsråd og gjennomfører forsknings- og utredningsoppdrag for næringsliv og offentlige etater. TØI ble opprettet i 1964 og er organisert som uavhengig stiftelse.

TØI utvikler og formidler kunnskap om samferdsel med vitenskapelig kvalitet og praktisk anvendelse. Instituttet har et tverrfaglig miljø med rundt 90 høyt spesialiserte forskere.

Instituttet driver forskningsformidling gjennom TØI-rapporter, artikler i vitenskapelige tidsskrifter, bøker, seminarer, samt innlegg og intervjuer i media. TØI-rapportene er gratis tilgjengelige på instituttets hjemmeside www.toi.no.

TØI er partner i CIENS Forskningssenter for miljø og samfunn, lokalisert i Forskningsparken nær Universitetet i Oslo (se www.ciens.no). Instituttet deltar aktivt i internasjonalt forskningssamarbeid, med særlig vekt på EUs rammeprogrammer.

TØI dekker alle transportmidler og temaområder innen samferdsel, inkludert trafikksikkerhet, kollektivtransport, klima og miljø, reiseliv, reisevaner og reiseetterspørsel, arealplanlegging, ITS, offentlige beslutningsprosesser, næringslivets transportbehov og generell transportøkonomi.

Transportøkonomisk institutt krever opphavsrett til egne arbeider og legger vekt på å opptre uavhengig av oppdragsgiverne i alle faglige analyser og vurderinger.

**Postadresse:**
Transportøkonomisk institutt
Gaustadalléen 21
0349 Oslo
Norge

E-post: toi@toi.no

**Kontoradresse:**
Forskningsparken
Gaustadalléen 21

Telefon: 22 57 38 00

Hjemmeside: www.toi.no