

# Utvikling av regionale (kortdistanse) transportmodeller i Norge

Jens Rekdal, Møreforskning Molde AS  
Anne Madslie, Transportøkonomisk Institutt

## Introduksjon

Samferdselsmyndighetene i Norge har engasjert Transportøkonomisk Institutt (TØI) og Møreforskning Molde AS (MFM) til å etablere et sett med landsdekkende kortdistanse transportmodeller som blant annet skal benyttes i arbeidet med nasjonal transportplan (NTP). Grunnlaget for arbeidet er blant annet en nasjonal reisevaneundersøkelse gjennomført i Norge i 2001 (RVU2001), og en rekke andre geografisk fordelte data fra samme tidsrom. RVU2001 og de øvrige data er stedfestet til grunnkretser som geografisk enhet. Det er vel 13500 slike grunnkretser i Norge. RVU2001 omfatter intervjuer med ca 21000 personer. I tillegg har PROSAM (samarbeidet for bedre transportprognoser i Osloregionen) i samme periode intervjuet 8800 personer bosatt i Oslo og Akershus fylke (PRVU01). Denne undersøkelsen er også benyttet i modellestimeringen. Samferdselsmyndighetene har i Samarbeid med SINTEF, også utviklet et sett regionale nettverksmodeller og som skal levere og motta data fra transportmodellene.

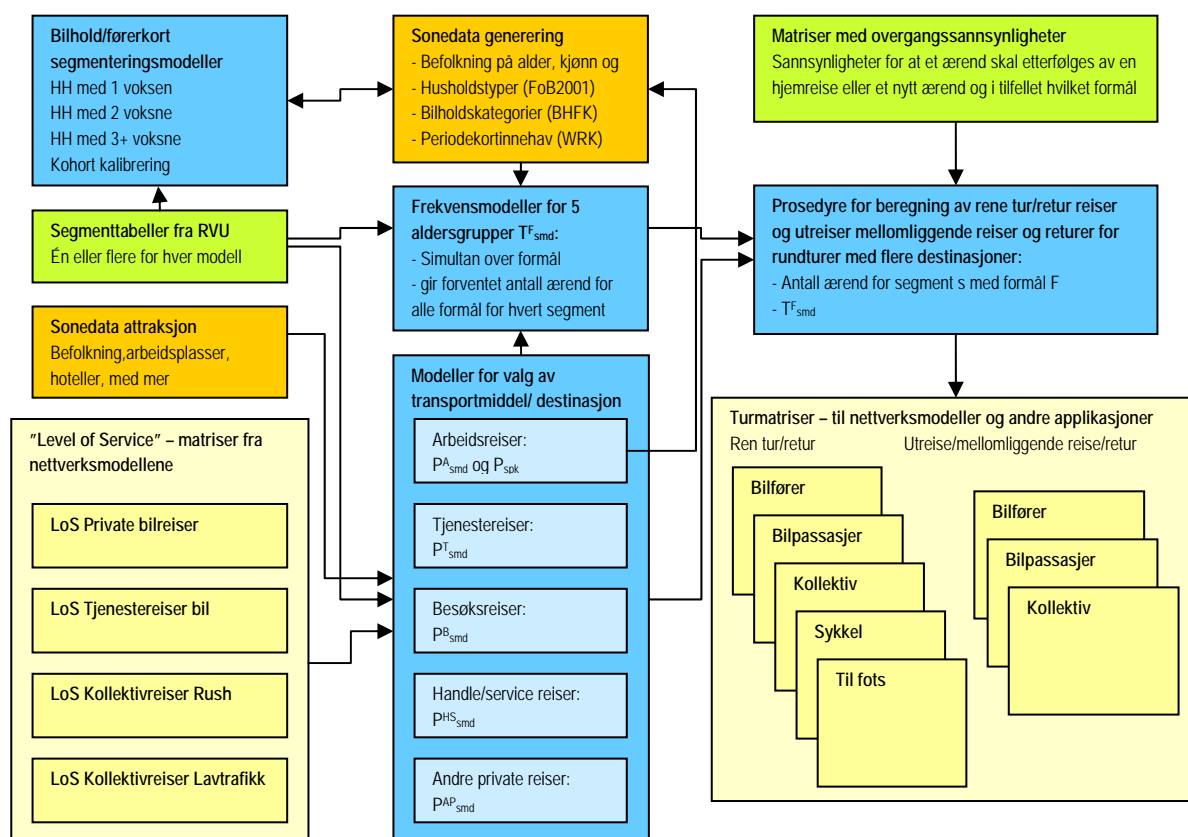
Arbeidsopplegget for prosjektet forutsatte at man skulle estimere felles modeller for hele landet, som videre skulle implementeres og tilpasses spesifikt i 5 landsdeler som tilsvarer Vegvesenets regioninndeling. Dette ble bestemt fordi datagrunnlaget i de to RVUer til sammen ikke var stort nok til å estimere egne modeller for hver av de 5 landsdelene, og fordi det dessuten er vesentlig mindre ressurskrevende å estimere et felles sett modeller. En ulempe er at man ved å estimere felles modeller må gi avkall på å håndtere forhold som er spesifikke for hver landsdel og som dessuten kan variere betydelig internt i hver landsdel (kø/reisetidspunkt, med mer). For å bedre ivareta behovet for samferdselsmyndighetenes daglige planlegging er programmeringen av modellene uavhengig av soneinndelingen. Dette gjør at en planlegger relativt lett kan definere og spesifisere mindre modellområder enn en region, for eksempel fylker eller grupper av kommuner, og kjøre modellene på disse. Dermed kan man unngå å kjøre modellen på en hel region med 5000-6000 soner hvis man skal studere et prosjekt eller en prosjektpakke med en klar geografisk avgrensning.

## Oversikt over modellsystemet

En modellkjøring vil vanligvis starte med ta man tar ut såkalte "Level of Service" (LoS) matriser fra nettverksmodellene. Det dreier seg her om matriser med informasjon om reisetider og kostnader mellom alle soner i modellområdet og for alle transportmåter. Disse spesifiseres som input data sammen med de øvrige data som inngår i en modellkjøring, demografiske *sonedata til generering*, *sonedata for attrahering av turer* og ulike *segmenttabeller* fra RVU2001 som gir gjennomsnittsverdier for befolkningssegmenter når det gjelder variable det ikke finnes sonedata for. Modellkjøringen starter med *bilholds- og førerkortmodellene* (BHFk-modeller) som segmenterer de demografiske data (befolkning i grunnkretser på alder, kjønn og husholdstype) etter bilhold og førerkortinnehav. Det er egne separat estimerte BHFk-modeller for tre ulike husholdstyper, og befolkningen fordeles på 5 segmenter etter biltilgang, avhengig av om personen selv har førerkort og hvor mange biler det er i husholdet i forhold til antall førerkort. Modellene er kalibrert slik at den rene tidseffekten i førerkortinnehavet ivaretas. BHFk-modellene er først og fremst viktige i langsiktige trafikkprognoser.

De 5 modellene for valg av transportmiddel og destinasjon (MD-modeller) er bosteds- og rundtur baserte og benytter bl.a. segmenteringen fra BHFk-modellene direkte. MD-modellene er simultant estimert på valg av transportmiddel og destinasjon, og de fleste er multinomiske logitmodeller. I MD-modellen for arbeidsreiser ivaretas også valg mellom periodekort og enkeltbillett for kollektivtransport, og dette er formulert slik at informasjonen om periodekortinnehavet kan benyttes direkte av de 4 øvrige reisehensiktene. Modellene for reisefrekvens (TF-modellene) er logitmodeller estimert i kombinasjon med Hurdle-poisson. Modellene er estimert simultant for alle reisehensikter og gir forventet antall ærend for alle reisehensikter. I frekvensmodellen inngår bl.a. logsummer fra de 5 MD-modellene som variable, hvilket innebærer at reisefrekvensene påvirkes av transporttilbudet og biltilgjengeligheten. Siden MD-modellene er basert på rundturer med bare én hoveddestinasjon, mens TF-modellen gir alle ærend som er gjennomført, er det laget en prosedyre som ivaretar rundturer med mellomliggende reiser. Prosedyren samordner informasjonen fra MD-modellene og TF-modellen slik at systemet produserer korrekt antall utreiser, hjemreiser og mellomliggende reiser, basert på matriser med "overgangssannsynligheter". Dette er sannsynligheter for at et ærend med et gitt formål etterfølges av en hjemreise eller et nytt ærend, og i tilfelle hvilket formål. Resultatet av en modellkjøring er turmatriser for rene tur/retur reiser og utreiser, mellomliggende reiser og returer for reiser med flere enn én destinasjon.

Figur 1. Oversikt over modellsystemet



### Segmenteringsmodeller for bilhold og førerkortinnehav

De demografiske data tilgjengelig for modellsystemet fordeler befolkningen i grunnkretsene på alder (5 års intervaller), kjønn og 3 husholdskategorier (hushold med én voksen person fra 18 år, med to voksne personer og med tre eller flere voksne personer). Bilholdet og førerkortinnehavet i husholdene

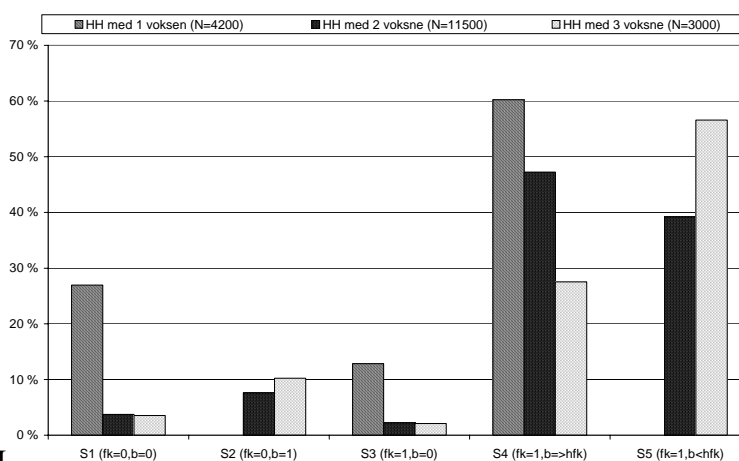
er et svært viktig aspekt ved individers muligheter til å foreta visse typer reiser, samtidig som utviklingen av biltilgjengeligheten over tid sannsynligvis er en avgjørende faktor for veksten i biltrafikken. Vi ønsker derfor at våre transportmodeller skal ivareta effekter av endret biltilgang og førerkortinnhav. Hensikten med disse modellene er dermed å segmentere befolkningen ytterligere for å få informasjon om befolkningsgruppene biltilgjengelighet. Vi har valgt å lage modeller som segmenterer befolkningen på alder og kjønn i sonene videre inn i 5 gjensidig utelukkende segmenter med ulik biltilgang. Resultatene fra disse modellene kan benyttes direkte i MD-modellene. De fem segmentene er:

- S=1: Personer uten førerkort og ingen biler i husholdet (ikke tilgang til bil som fører, dårlig tilgang til bil som passasjer).
- S=2: Personer uten førerkort, men med en eller flere biler i husholdet (bare biltilgang som bilpassasjer).
- S=3: Personer med førerkort, men uten biler i husholdet (dårlig tilgang til bil)
- S=4: Personer med førerkort og minst like mange biler som førerkort i husholdet (full biltilgang)
- S=5: Personer med førerkort og færre biler enn førerkort i husholdet (delvis/god biltilgang)

Som vi ser, har personer som tilhører segment 1 og 2 ikke førerkort, og personer som tilhører segment 1 og 3 har ikke bil. For personer med førerkort, er totalt antall førerkort og totalt antall biler i husstanden avgjørende for biltilgangen. Data fra RVU2001 er benyttet som estimeringsgrunnlag, sammen med data som beskriver respondentenes bostedsgrunnkrets. Materialet er splittet i tre deler avhengig av antall personer i husholdet som har fylt 18 år.

Figur 2. Fordeling av personer på BHFK-segmenter etter husholdstype

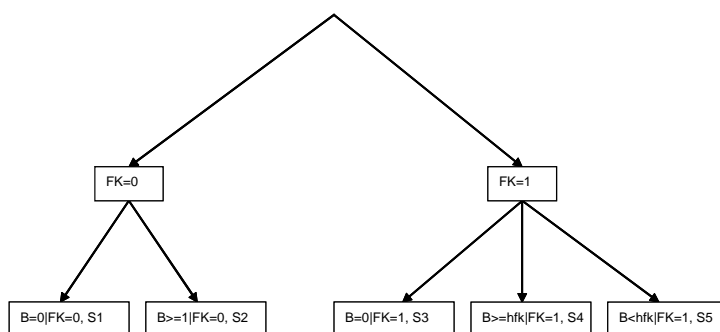
Figur 2 viser at fordelingen på de 5 bilhold/førerkortsegmentene varierer betydelig mellom de tre husholdstypene. I hushold med én voksen er andelen med full biltilgang 60 %, mens den er knappe 50 % i hushold med 2 voksne og knapt 30 % i hushold med 3 og flere voksne. I hushold med én voksen person er dessuten andelen uten biltilgang nær 30 % (eldre og yngre aleneboende). I hushold med 3 og flere voksne er andelen med delvis biltilgang nær 60 %.



Den metoden som er valgt innebærer å definere situasjonen som tre atskilte problemer, førerkortinnhavet, biltilgang for personer med førerkort, og biltilgang for personer uten førerkort. Samtidig er modellene for disse tre situasjonene estimert simultant. Førerkortinnhavet er en binær situasjon ( $fk=1$  eller  $fk=0$ ). Biltilgang for personer med førerkort representeres ved tre gjensidig utelukkende alternativer, ingen biler ( $b=0$ ), bil men færre biler enn førerkort i husholdet ( $b>0, b<hfk$ ), og bil og like mange eller flere biler enn førerkort i husholdet ( $b>0, b>=hfk$ ). Når vi for personer uten førerkort ikke trenger å ta hensyn til antall førerkort og antall biler blir situasjonen som vist i figur 3. For hushold med bare én voksen person vil ikke S2 og S5 være aktuelle, da det i slike hushold ikke er naturlig å ha bil hvis man ikke har førerkort (S2), og fordi det maksimalt kan være én person med førerkort i slike hushold (S5). For hver av de tre husholdskategoriene er det estimert modeller som fordeler befolkningen på de 5 bilhold/førerkortsegmenter.

Figur 3. Struktur i modeller for bilhold og førerkortinnehav

Modellene er estimert med en maximum likelihood prosedyre kodet i GAUSS. Modellene er spesifisert med "nyttefunksjoner" som består av alternativspesifikke konstanter og variable som beskriver bostedssone, individuelle kjennetegn og husholdskaraktistika. De



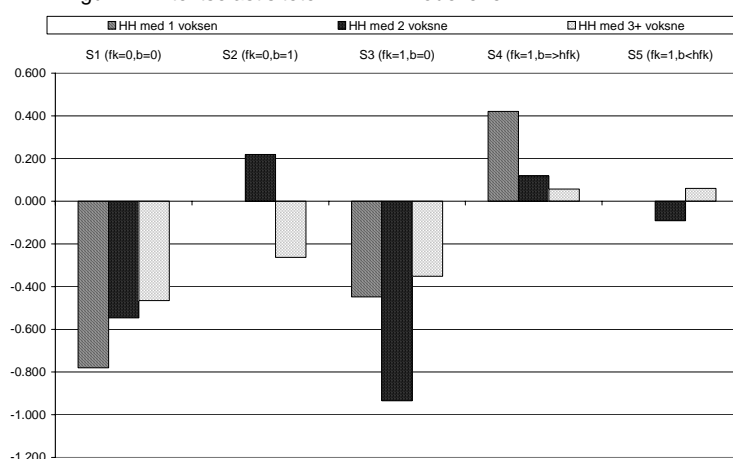
viktigste variablene er alder, kjønn, familietype, befolkningstetthet og inntekt. Alder og kjønn inngår både formulert som dummyvariable og som kontinuerlige variable.

Vi har ikke hatt tilgang til variable som beskriver parkeringsforholdene ved bostedet direkte. Befolkningstetthet er benyttet som en proxyvariabel i forhold til dette, formulert både som kontinuerlige variable og som ulike dummyvariable. Tanken er at det i områder med høy befolkningstetthet ofte kan være høyere generaliserte kostnader knyttet til bilhold enn i områder med lav tetthet. Denne variabelen vil imidlertid også fange opp andre faktorer, som bedre kollektivtilbud, kortere avstander til aktiviteter, med mer, og dermed mindre behov for bil i områder med høy tetthet. Det er en klar sammenheng mellom befolkningstetthet og både bilhold og førerkortinnehav i våre modeller. I områder med høy befolkningstetthet er sannsynligheten for full biltilgang og førerkort lavere, og sannsynligheten for delvis eller ingen biltilgang høyere enn ellers.

Husholdsinntekt inngår i estimeringen som kontinuerlige variable. Det er forutsatt at den marginale bilen har en fast årlig kostnad på NOK 10000.- som trekkes fra husholdsinntekten. Den marginale bilen er ofte mye eldre enn gjennomsnittsbilen og har dermed lavere verdi. Eldre biler har sjelden kaskoforsikring. I alle modeller har inntektsvariabelen et tillegg for storbyområder (de 4 største byene i Norge), som gjør at inntektseffekter dempes i disse områdene.

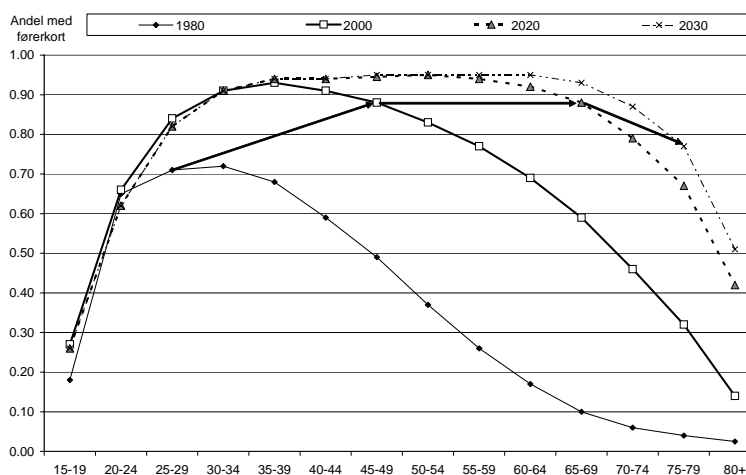
Figur 4. viser inntektselastisitetene i de tre modellene beregnet med utgangspunkt i fordelingen på hushold og segmenter i datagrunnlaget for estimeringen (prosentvis forskyvning mellom segmentene som følge av 1 % økning i inntektene). Størrelsen på elastisitetene er svært avhengig av fordelingen på segmentene i utgangspunktet (jfr. fig 2). Det er et generelt trekk at elastisitetene er høye i tallverdi for alternativer med få personer og lave for alternativer med mange personer.

Figur 4. Inntektselastisiteter i BHFk-modellene



Figur 5. Kohorteffekter i førerkortinnehavet. Kvinner

I implementeringen kalibreres modellene mot prognoser for førerkortinnehavet i år 2010, 2015, 2020, 2025 og 2030. For kvinner er disse rene tidseffektene er illustrert i figur 5. Pilene mellom kurvene antyder forløpet i førerkortandelen blant kvinner mellom 25 og 29 år i 1980. I 2030 vil disse være mellom 75 og 79 år og andelen med førerkort vil være høyere enn i 1980. Når modellene skal benyttes til langsiktige prognoser vil kalibreringen sørge for å ivareta disse effektene.



### Modeller for valg av transportmiddel og destinasjon

**Datamaterialet** til estimeringen av disse modellene er basert på tre ulike kilder. *Reisevanedata* (fra RVU2001 og PRVU01) gir informasjon om valgt transportmåte og valgt destinasjon for rundturer fra bosted til én hoveddestinasjon og tilbake til bosted, om kjennetegn ved informantene som har gjennomført rundturene, og om informantenes hushold. *Sonedata* gir informasjon om hva som befinner seg av attraksjoner (arbeidsplasser, befolkning, med mer) i mulige destinasjoner, og *LoS-data* (Level of Service) gir informasjon om reisetider og kostnader forbundet med besøk i de ulike destinasjoner. All informasjon som er benyttet er stedfestet til grunnkretser som geografisk enhet. Grunnkretser er den mest detaljerte standard geografiske inndelingen i Norge (ca 13500 stk i hele landet). Siden modellene vil reflektere kvaliteten på de tre datakilder er det lagt en del ressurser på å "vaske" dataene.

Utgangspunktet for reisevanene i estimeringsgrunnlaget er informantenes redegjørelse av alle gårsdagens reiser, med valgt transportmiddel, valgt destinasjon, reisetidspunkt, reisetid, med mer, for alle bevegelser utenfor bostedet. Dette skal gi opphav til en sammenhengende sekvens med reiser som, for de aller fleste informantene, starter i bostedet om morgenen og ender i bostedet om ettermiddagen eller kvelden. Ved gjennomgangen av datamaterialet ble det identifisert en del usammenhengende reisesekvenser og andre feil, som enten ble rettet opp eller forkastet (17 % av de ca 93000 delreisene ble forkastet). De sammenhengende reisesekvensene som starter og ender i bostedet ble gjort om til rundturer med én eller flere destinasjoner. Det ble i alt ca 35000 slike rundturer, hvorav vel 75 % bare har én destinasjon. Rundturer med to og flere destinasjoner ble behandlet spesielt for å finne ut om én av destinasjonene kunne benyttes som hoveddestinasjon. Rundturer som "lignet" på en ren tur/retur reise ble akseptert til å inngå i estimeringsgrunnlaget, men spesifisert som en tur/retur, mens mer kompliserte rundturer ble forkastet. Forkastingen var basert på visse kriterier knyttet til tid tilbrakt på hver destinasjon, avstand fra bosted til hver destinasjon, og faktisk reiseavstand for hele rundturen sammenliknet med avstand til og fra bostedet til potensielle hoveddestinasjoner. Ca 35 % av rundturene med to og flere destinasjoner ble akseptert og disse ble tilordnet en variabel som indikerer at én eller flere sekundære destinasjoner er tatt bort fra disse rundturene. Resten av rundturene ble vurdert som for komplisert og forkastet. Årsaken til den relativt konservative behandlingen av disse rundturene er knyttet til problemene ved å spesifisere korrekte LoS-data til de rundturene som har flere destinasjoner når disse ikke likner på en ren tur/retur.

Når observasjoner med manglende stedfesting, uaktuelle transportmåter og reisehensikter, og andre dataproblemer er fjernet omfatter estimeringsgrunnlaget i alt ca 23000 rundturer. For hver rundtur trekkes det et tilfeldig utvalg på maksimalt 249 destinasjoner fra mengden soner innenfor 10 mil fra bostedet, som brukes som alternativer til den destinasjon som faktisk er valgt, og korresponderende sonedata for attraksjon og LoS-data kobles til den valgte og de samlede destinasjoner. Det er estimert bosteds- og rundturbaserte modeller for de fem reisehensiktene arbeidsreiser, tjenestereiser, handle/servicereiser, private besøksreiser og andre private reiser. Estimeringsarbeidet har vært en relativt omfattende prosess med prøving og feiling, feilsøk og korrigeringer av datamaterialet. Vi har fremdeles et inntrykk av at spesielt kollektivtransport er relativt skjematisk behandlet i nettverkene, i hvert fall i deler av landet. På den andre side har hver korrigering som er gjennomført og som har gitt nye LoS-matriser gitt relativt moderat påvirkning på modellestimatene.

**Modellene** er estimert for transportmåtene bilfører, bilpassasjer, kollektivtransport, sykkel og gang (til fots). Flest reiser er gjort som bilfører, og denne andelen varierer fra 47 % (besøksreiser) til 69 % (tjenestereiser). Det er færrest sykkelreiser i materialet, fra 3 % for handle/servicereiser til 8 % for besøksreisene. Det er videre slik at en relativt stor andel av reisene ikke har kollektivtransport tilgjengelig (hele 60 % for noen reiseformål), noe som tyder på at kollektivnettet ikke er så detaljert kodet som man skulle ønsket. I datamaterialet er det signifikante forskjeller mellom kvinner og menn både når det gjelder transportmiddelfordeling og avstandsfordeling. Relativt sett er kvinner oftere bilpassasjerer enn menn og de benytter hyppigere kollektivtransport eller går til arbeid. I tillegg reiser de i gjennomsnitt kortere enn menn, dette er spesielt tydelig for arbeidsreisene og tjenestereisene.

For alle reiseformål har vi testet ut ulike trestrukturer. Generelt har vi fått som resultat at varianter med destinasjonsvalget over transportmiddelvalget fikk en logsumkoeffisient som ikke var signifikant forskjellig fra 1. Variantene med transportmiddelvalget over destinasjonsvalget hadde en noe bedre log-likelihood-verdi enn de multinomiske modellene, men kostnadskoeffisienten var betydelig lavere i tallverdi og signifikans, noe som gav betydelig høyere implisitte tidsverdier. I og med at vi allerede i utgangspunktet hadde tidsverdier som var i overkant av forventet, valgte vi å benytte multinomiske modeller framfor strukturerte ved valg av transportmiddel og destinasjon. Unntaket er modellen for andre private reiser som er strukturert med destinasjonsvalget over transportmiddelvalget. Denne modellen har vært spesielt vanskelig å estimere med tilfredsstillende resultater. Årsaken er trolig at dette er en relativt lite homogen reisehensikt som bl.a. omfatter fritidsreiser, rekreasjonsreiser, hente/bringe reiser og en del andre mindre hensikter.

Modellen for arbeidsreiser har en noe mer komplisert struktur enn de andre modellene, da vi her har valgt å estimere transportmiddel/destinasjon og periodekortinnhav simultant. Den enkle (og mest vanlige) måten å håndtere periodekort for kollektivtransport på, er uten tvil å benytte månedskortpris/antall reisedager som kostnad for dem som har periodekort (til alle aktuelle destinasjoner). Dette er en praksis som har flere ulemper, hvorav den viktigste er at man ikke får tatt hensyn til at kollektivtransport gjerne vil oppfattes som gratis på registreringsdagen. De som har periodekort vil også gjerne bare være sporadiske brukere av bil på arbeidsreisen og det er derfor rimelig å anta at de ikke kan utnytte rabattmuligheter når det gjelder eventuelle bompenger og ferger.

I modellen for arbeidsreiser har vi derfor valgt å behandle reiser med periodekort som et eget nest med en egen logsumparameter. Dette innebærer at vi får 5 ekstra nyttefunksjoner, én for hver reisemåte. De øvrige nyttefunksjoner multipliseres også med samme logsumparameter for å bringe dem på samme

nivå (et "dummy-nest"). Dette er nødvendig for å få korrekte generiske parametre. Kostnaden for en arbeidsreise ved bruk av periodekort settes til månedskortpris/22. Denne kostnaden legges til alle transportmåter under nestet med periodekort, og tilordnes en egen parameter. Kollektivtransport har kun denne kostnaden under nestet med periodekort, mens de øvrige transportmåter i tillegg er tilordnet de ordinære kostnadene forbundet med disse. Internt i nestet for reiser med periodekort vil dermed kollektivtransport fremstå som gratis, og mellom nestene vil kollektivtransport være billigere med periodekort enn uten. Samtidig vil det være dyrere å reise med andre transportmåter hvis man har periodekort. Det viser seg at parameteren til prisen på periodekort blir høyere enn parameteren for andre kostnader (en krone anvendt på månedskort "vurderes" til kr 1.30), noe som kan skyldes at kjøp av periodekort innebærer et relativt stort kontantutlegg, samtidig som det alltid vil være en viss usikkerhet knyttet til hvor mange reiser det faktisk vil benyttes for (pga sykdom osv). Det er flere fordeler ved å operere med et eget nest for periodekort, bla:

- Valgbetingelsene blir så langt mulig korrekt spesifisert og vi bryter samtidig en del av korrelasjonen mellom reisetid og kostnad. Prisen på periodekort påvirker destinasjonsvalget og valget mellom kort / ikke kort for en gitt destinasjon, men for dem som velger periodekort fremstår kollektivtransport som gratis i forbindelse med transportmiddelvalget.
- Vi unngår problemet med spesifisering av kostnaden for kollektivtransport for alternative destinasjoner og for alternative transportmidler siden reiser med periodekort er egne alternativer.
- Vi kan differensiere kostnaden for bilførere og bilpassasjerer etter hvorvidt disse transportmåter benyttes sporadisk eller ikke. Dette bidrar også til å bryte opp korrelasjoner.
- Vi får automatisk en modell for periodekortinnhav som kan benyttes i forbindelse med øvrige transportmidler dersom arbeidsreisemodellen kjøres først.
- Det blir mulig å analysere effekter av å endre prisdifferansen mellom enkeltbillett/klippekort og periodekort med modellen. Prisforskjellen vil påvirke både kortinnhav og reisemiddelvalg.

I alle modellene finner vi at kvinner har høyere tidsverdier enn menn, om enn i varierende grad. Størst forskjell er det i modellene for tjenestereiser, der tidsverdiene for kvinner er nesten det dobbelte som for menn. For besøksreiser og handle/servicereisene finner vi forskjellige tidsverdier kun for bilførere, mens vi for kategorien "andre reiser" også har det for bilpassasjerer. For tjenestereiser og arbeidsreiser finner vi forskjeller mellom kvinner og menn både for bilførere, bilpassasjerer og kollektivreisende. Differensiering av tidsparametrene ble i utgangspunktet forsøkt som en proxy for inntektsforskjeller, men forskjellen ble motsatt av hva man da skulle forvente. Det er flere mulige forklaringer på dette:

- Kvinner har gjennomgående et strammere tidsbudsjett enn menn, noe også tidsverdistudier bekrefter.
- Kvinner oppfatter gjennomgående ulempen ved å kjøre bil som større enn det menn gjør.
- Kvinner kan gjennomgående ha bedre muligheter til å skaffe arbeid nær hjemmet enn menn. I så fall vil forskjellen i tidsparametre reflektere at søners attraktivitet som mål for arbeidsreiser for hhv menn og kvinner ikke er godt nok spesifisert.
- Kvinner og menn har ulike typer tjenestereiser, hvor kvinners reiser generelt er kortere.
- Gjennomsnittsavstanden til valgte reisemål generelt er kortere for kvinner enn menn.

Generelt er det slik at de implisitte tidsverdiene som er beregnet virker noe høye for alle modellene, med unntak av arbeidsreiser. Viktigste grunnen til dette kan være at reisetidene som er brukt ved estimeringen er beregnet i en nettverksmodell hvor det ikke forekommer kapasitetsproblemer. I tillegg er det korrelasjon mellom tid og kostnad som gir en tendens til at en av parametrene fanger opp den samlede effekten av avstand. Tidsverdiene er spesielt høye for kategorien andre private reiser, som trolig henger sammen med at reiseformålet er lite homogent, og spenner fra lange rekreasjonsreiser til korte hente/bringe reiser. Hovedtyngden av disse reisene foregår imidlertid i helgene og dummyvariable i interaksjon med tidsvariablene for bil reduserer tidsverdiene i helgene betraktelig noe som kan ha sammenheng med at folk generelt har slakkere tidsbudsjett i helgen.

Tabell A Implisitte tidsverdier i modellene (NOK/t)

|                                    | Arbeid | Tjeneste | Besøk | Handle/serv. | Annet          |
|------------------------------------|--------|----------|-------|--------------|----------------|
| <b>Bilfører (CD):</b>              |        |          |       |              |                |
| Mann                               |        | 135      | 63    | 66           | 198            |
| Kvinne                             |        | 242      | 89    | 87           | 253            |
| Tillegg storby                     |        | 58       |       |              |                |
| Tillegg helg                       |        |          | -19   |              | -84(m), -77(k) |
| Mann under 50 år                   | 56     |          |       |              |                |
| Mann, 50 år +                      | 73     |          |       |              |                |
| Kvinne, under 50 år                | 84     |          |       |              |                |
| Kvinne, 50 år +                    | 107    |          |       |              |                |
| Tillegg destinasjon Oslo i rush    | 32     |          |       |              |                |
| <b>Bilpassasjer (CP):</b>          |        |          |       |              |                |
| Bilpassasjer                       |        |          | 61    | 55           |                |
| Mann                               | 94     | 64       |       |              | 198            |
| Kvinne                             | 154    | 147      |       |              | 253            |
| Tillegg storby                     |        | 58       |       |              |                |
| Tillegg helg                       |        |          | -19   |              | -84(m), -77(k) |
| <b>Kollektiv (PT):</b>             |        |          |       |              |                |
| Ombordtid                          |        |          | 33    | 28           | 92             |
| Ombordtid, mann                    | 39     | 56       |       |              |                |
| Ombordtid, kvinne                  | 49     | 125      |       |              |                |
| Tillegg helg                       |        |          |       |              | -32            |
| Gangtid                            | 43     | 168      | 61    | 57           | 133            |
| Ventetid ved 5 minutters ventetid  | 60     | 369      | 70    | 84           | 124            |
| Ventetid ved 30 minutters ventetid | 24     | 150      | 29    | 34           | 51             |
| Ventetid ved 60 minutters ventetid | 17     | 106      | 20    | 24           | 36             |
| Overgang (NOK pr overgang)         |        | 39       |       | 16           |                |

Ved estimeringen av modellene har vi hatt visse problemer med å treffe med observert avstandsfordeling. Dette er vanskelig når en modell skal dekke hele avstandsintervallet fra noen meter til 20 mil for en tur/retur reise, samtidig som nyttefunksjonene er lineære i koeffisienter og variable (spesielt når avstandsfordelingen i datamaterialet ikke er tilnærmet lineær). Tendensen er at modellene underpredikerer korte og lange reiser og overpredikerer de mellomlange. Vi har testet ut en del ikke-lineære transformasjoner av tidsvariablene, trinnvis lineære spesifikasjoner, samt ulike interaksjoner mellom variable uten at det gav tilfredsstillende resultater sett i sammenheng med den økning i regnetid som slike formuleringer vil medføre. Vi har derfor lagt inn enkelte avstandsdummys i modellspesifikasjonen, noe som har ført til en markert bedring av modellenes evne til å gjenspeile datamaterialets avstandsfordeling.

Vi vil ikke gå inn på alle variable som har vist seg signifikant i de ulike modellene, men nevner noen fellestrekk for modellene:

- Ventetid i kollektivtransport er en variabel som var problematisk i forbindelse med estimeringen. Dette skyldes at man i nettverksmodellen, hvor denne variabelen beregnes, får ut halvparten av tiden mellom avgangene, summert over antall påstigninger (ev. en maksimal ventetid pr påstigning) og ikke det som er folks faktiske ventetid. Vi testet ut et utall varianter, som åpen og skjult ventetid, ventetid pr påstigning, ulike lineære transformasjoner, sammenslått gangtid og ventetid osv, for å få ventetiden signifikant. Vi endte opp med en formulering basert på kvadratroten av total ventetid. Ved dette får man redusert effekten av lange ventetider, samtidig får man en ventetidsvekt som varierer med størrelsen på ventetiden.
- En dummy som angir at turen har minst én sekundær destinasjon (flere besøk underveis) øker sannsynligheten for bruk av bil.
- De fem ulike kategorier for førerkort/bilhold i husholdningen er viktige forklaringsvariable i alle modellene. Generelt fører god eller full biltilgang til økt sannsynlighet for bruk av bil som fører, samtidig finner vi at det i hushold med færre biler enn førerkort ofte er kvinnene som taper "kampen om bilen".
- Det er lagt inn enkelte dummys for alder og type region (storby osv), men dette er forsøkt begrenset for å unngå altfor mange segmenter i de endelige modellene.
- Vi har begrenset med informasjon om parkeringsforhold knyttet til den enkelte reise. I stedet har vi generert noen indekser som sier noe om arbeidsplass tettheten i destinasjonene, med det for øye at disse skal si noe om hvor vanskelig (og kanskje dyrt) det er å finne parkeringsplass. For arbeidsreiser og tjenestereiser finner vi at de to høyeste arbeidsplass tetthetene reduserer tilbøyeligheten til å kjøre bil, med størst ulempe knyttet til indeksen for høyest arbeidsplass tetthet. Denne kategorien arbeidsplass tetthet virker også negativt inn på bilbruken for handle/servicereisene og andre private reiser, mens vi ikke finner noen slik sammenheng for besøksreisene.



## Modeller for valg av reisefrekvens

Datagrunnlaget for estimeringen av disse modellene er personfilen fra RVU2001 påkodet informasjon om antall ærend foretatt registreringsdagen for de 5 reisehensikter det er estimert MD-modeller for og her er skolereiser også tatt med. Skolereisene er ellers behandlet relativt skjematisk i modellsystemet, med gravitasjonsmodeller som fordeler barn og ungdom i grunnskolealder til nærmeste barne- eller ungdomsskole. Fordi kollektivtilbudet for skolereisene ikke er kodet i nettverkene er det for disse reisene ikke laget noen modell for valg av transportmiddel. Skolereisene tas imidlertid med i frekvensmodellene.

Frekvensmodellene gir forventet antall ærend for hver reisehensikt. Konseptet går i korthet ut på å estimere en vektet logitmodell som ivaretar fordelingen på ærend med ulike reisehensikter i kombinasjon og simultant med en hurdle poisson modell som gir forventet antall ærend totalt. Poisson modeller benyttes ofte i situasjoner hvor man er interessert i fordelingen på antall eller beholdningen av noe (1,2,3,4,5...). Hurdle-poisson benyttes når "0" også inngår i denne fordelingen. Med dette konseptet spesifiseres én nyttefunksjon for hver reisehensikt som formuleres med alternativspesifikke konstanter og variable som beskriver kjønn, alder, familietype, bostedstype, samt en logsum fra korresponderende MD-modell. I estimeringen vektet nyttefunksjonene med antall ærend som er gjennomført for hver reisehensikt. Inntekt inngår ikke direkte i disse nyttefunksjonene, men kommer inn fra bilholdsmodellene via logsummene fra MD-modellene. Koblingen mellom den vektete logitmodellen og hurdle poisson skjer gjennom formulering av en samlet logsum fra logitmodellen for fordeling på ærend med ulike reisehensikter. Denne samlede logsum inngår i fordelingsfunksjonen i Hurdle-poisson modellen, uten parameter for antall ærend lik 0, og med en parameter for antall ærend større enn null. Hurdle-poisson modellen gir dermed forventet antall ærend for de 6 reisehensikter totalt, mens fordelingen av ærendene på reisehensikter skjer med logitmodellen.

Estimeringsarbeidet startet med forsøk på å formulere én samlet modell for alle befolkningssegmenter. Dette krevde imidlertid en dyptgående segmentering med dummyvariable for ulike aldersgrupper. Det ble derfor estimert uavhengige modeller av samme type for 5 ulike aldersgrupper (13-24 år, 25-34 år, 35-54 år, 55-66 år og 67 år og mer). Parameteren knyttet til logsummene fra MD-modellene har størst tallverdi i modellene for de yngste og spesielt de eldste aldersgruppene. Siden logsummene i stor grad reflekterer transportressursene til segmentene (egne transportressurser så vel som generell tilgjengelighet til attraktive destinasjoner med ulike transportmåter), reflekterer dette at de midlere aldersgrupper har et mindre elastisk transportbehov enn de yngste og eldste. Situasjonen når det gjelder tilgjengelighet er altså mer avgjørende for reisefrekvensene for de yngste og eldste. Tabell B viser modellenes fordeling på ærend for de 6 typer av reisehensikter som er definert. Totalt for hele landet treffer modellene mot datamaterialet med under 1 % avvik, mens det er noe større avvik når man går inn i de 5 regionene. Dette har gjort det nødvendig å introdusere regionale dummyvariable. På nasjonalt nivå ser vi at situasjonen er svært ulik mellom aldersgruppene og dette er også noe av bakgrunnen for at inndelingen av materialet i aldersgrupper viste seg nødvendig.

Tabell B Fordeling på ærend av ulike reisehensikter etter modell/aldersgruppe. Virkedøgn, hele landet.

|                | 13-24 | 25-34 | 35-54 | 55-66 | 67+  | Totalt |
|----------------|-------|-------|-------|-------|------|--------|
| Arbeid         | 12 %  | 26 %  | 31 %  | 28 %  | 3 %  | 24 %   |
| Tjeneste       | 3 %   | 9 %   | 11 %  | 9 %   | 1 %  | 8 %    |
| Handle/service | 19 %  | 24 %  | 24 %  | 29 %  | 53 % | 26 %   |
| Besøk          | 15 %  | 10 %  | 7 %   | 11 %  | 14 % | 10 %   |
| Andre private  | 26 %  | 29 %  | 27 %  | 23 %  | 30 % | 27 %   |
| Skole          | 24 %  | 3 %   | 1 %   | 0 %   | 0 %  | 6 %    |

## Prosedyre for beregning av ”mellomliggende” turer

Modellene for valg av transportmiddelvalg er estimert på data som reflekterer rene tur/retur reiser til én hoveddestinasjon. Mer komplekse rundturer med flere destinasjoner er enten forenklet (hvis turen likner på en ren tur/retur reise) eller forkastet fra datamaterialet for estimeringen av disse modellene. Dette er gjort med tanke på å spesifisere mest mulig korrekte LoS-data for reisene til estimeringen av MD-modellene. I materialet for estimeringen av frekvensmodellene er imidlertid alle ærend tatt med, også ærend som er gjennomført gjennom komplekse rundturer med flere destinasjoner. En respondent kan i tillegg ha gjennomført flere rundturer registreringsdagen. I datagrunnlaget for MD-modellene, hvor enheten er rundturer, vil dette gi opphav til flere observasjoner fra slike respondenter, mens alle ærend i materialet for frekvensmodellene er summert over virkedagen.

Hvis resultatene fra MD-modellene og frekvensmodellene ble benyttet direkte ville modellsystemet produsert for mange delreiser. Dette ville vært ekvivalent med å anta at hvert ærend gir opphav til én tur/retur reise. På den andre siden vil et datamateriale for frekvensmodellene bare basert på de rene rundturer slik de er definert i materialet for MD-modellene gitt for få turer, i og med at turer med to og flere destinasjoner enten er forenklet til turer med én hoveddestinasjon, eller forkastet helt fra datamaterialet. På denne bakgrunn er det utarbeidet en metode som i prinsippet skal gi et tilnærmet korrekt antall delreiser, og metoden er basert på følgende tre elementer:

- Modeller for transportmiddel og destinasjonsvalg, som er estimert på data som reflekterer rene tur/retur reiser for de definerte reisehensikter
- Frekvensmodell som gir antall ærend for de definerte reisehensikter
- Matriser med ”overgangssannsynligheter” som gir sannsynligheten for at en reise med hensikt ”i” etterfølges av en hjemreise eller fortsetter til en annen destinasjon med formål ”j” og deretter hjemreise.

Matrisene med overgangssannsynligheter kan beregnes med utgangspunkt i RVU-data. For alle delreiser som ikke er hjemreiser registreres reisehensikten og hensikten for neste reise. Alle slike par telles opp og legges i en tabell, slik at sannsynlighetene kan beregnes. I vårt tilfelle får tabellen 6 rader og 7 kolonner fordi vi har 6 reisehensikter pluss hjemreiser (og dermed 42 mulige kombinasjoner av reisehensikter for reiser og påfølgende reiser inkl hjemreiser).

Metoden er lettest å illustrere gjennom et enkelt eksempel. Vi kan tenke oss en situasjon hvor bosatte i en sone, S1, har tre mulige reisemål S1, S2 og S3 og kan gjennomføre to typer ærender A og B. Vi antar at frekvensmodellen gir 100 ærend av type A og 50 ærend av formål B, og at destinasjonsvalgmodellen gir  $PA(S1,S2,S3) = (0.1, 0.5, 0.4)$  og  $PB(S1,S2,S3) = (0.2, 0.2, 0.6)$ .

Tabell C Hvis vi antar at alle ærend gjennomføres som rene tur/retur reiser:

| fra\til | 1   | 2  | 3  | sum |
|---------|-----|----|----|-----|
| 1       | 40  | 60 | 70 | 170 |
| 2       | 60  | 0  | 0  | 60  |
| 3       | 70  | 0  | 0  | 70  |
| sum     | 170 | 60 | 70 | 300 |

Hvis alle reiser gjennomføres som rene tur/retur reiser finnes antall utreiser og returer ved å multiplisere antall ærend for hvert formål med vektoren av destinasjonssannsynligheter for formålet og summere resultatene fra de to formålene (tabell C). Dette er ekvivalent med å kombinere resultatene fra frekvensmodellene og MD-modellene direkte. Sett at vi vet at 30 av ærendene med formål A etterfølges av et ærend med formål B. Dette betyr at bare 70 av ærendene med formål A er rene tur/retur reiser og at bare 20 av ærendene med formål B er bostedsbaserte.

Tabell D Hvis vi antar at alle reisene bare har ett formål:

| fra/til | 1   | 2  | 3  | sum |
|---------|-----|----|----|-----|
| 1       | 28  | 54 | 52 | 134 |
| 2       | 54  | 0  | 0  | 54  |
| 3       | 52  | 0  | 0  | 52  |
| sum     | 134 | 54 | 52 | 240 |

Hvis vi later som om at alle reisene bare har ett formål, slik vi har gjort det i prepareringen av dataene for MD-modellene, forsvinner 30 av ærendene med formål B (tabell D). Dette betyr at 60 delreiser blir borte (tur/retur). Dette er ekvivalent med å benytte materialet fra MD-modellene til å etablere datamaterialet for frekvensmodellene, og ikke et materiale som reflekterer alle ærend som er gjennomført. Den tilgjengelige informasjonen kan imidlertid benyttes til å gi et mer korrekt bilde av reisemønsteret.

Vi vet at 70 av ærendene med formål A og 20 av turene med formål B er rene tur/retur reiser. Dette gir opphav til 180 delreiser tur/retur. (tabell E)

Tabell E Rene tur/retur reiser:

| fra/til | 1   | 2  | 3  | sum |
|---------|-----|----|----|-----|
| 1       | 22  | 39 | 40 | 101 |
| 2       | 39  | 0  | 0  | 39  |
| 3       | 40  | 0  | 0  | 40  |
| sum     | 101 | 39 | 40 | 180 |

Det er 30 mellomliggende reiser og matrisen  $[PA(S1,S2,S3)' * PB(S1,S2,S3)]$  kan benyttes som et rimelig anslag på fordelingen på start og målpunkter for disse reisene (tabell F).

Tabell F Mellomliggende reiser:

| fra/til | 1   | 2   | 3    | sum  |
|---------|-----|-----|------|------|
| 1       | 0.6 | 0.6 | 1.8  | 3.0  |
| 2       | 3.0 | 3.0 | 9.0  | 15.0 |
| 3       | 2.4 | 2.4 | 7.2  | 12.0 |
| sum     | 6.0 | 6.0 | 18.0 | 30.0 |

Utreiser og returer for de mellomliggende reisene finnes ved å snu marginalene fra den mellomliggende matrisen som utreiser og returer fra/til sone 1 (tabell G).

Tabell G Utreise/hjemreise for mellomliggende reiser:

| fra/til | 1  | 2  | 3  | sum |
|---------|----|----|----|-----|
| 1       | 9  | 15 | 12 | 36  |
| 2       | 6  | 0  | 0  | 6   |
| 3       | 18 | 0  | 0  | 18  |
|         | 33 | 15 | 12 | 60  |

Totalt antall turer, finnes til slutt ved å summere de tre matrisene over, og resultatet blir en matrise som er symmetrisk på marginalene, men ikke element for element (tabell H).

Tabell H Totalt antall turer:

| fra/til | 1     | 2    | 3    | sum   |
|---------|-------|------|------|-------|
| 1       | 31.6  | 54.6 | 53.8 | 140.0 |
| 2       | 48.0  | 3.0  | 9.0  | 60.0  |
| 3       | 60.4  | 2.4  | 7.2  | 70.0  |
| 0       | 140.0 | 60.0 | 70.0 | 270.0 |

Dette eksempelet illustrerer tankegangen ved det konsept som er utviklet og implementert i modellsystemet. Komplexiteten øker når man skal ta hensyn til vesentlig flere mulige soner, flere typer ærend og reisehensikter, flere transportmåter, og en stor mengde befolkningssegmenter. Fordelen ved metoden er at det tar svært liten regnetid i forhold til andre tilnæringsmetoder for å få beregnet mellomliggende reiser (for eksempel mer eksplisitt modellering av sekundære destinasjoner i rundturene). Ulempene er i første rekke at fordelingen av mellomliggende reiser blir avhengig av sannsynlige destinasjoner med bostedet som utgangspunkt, og ikke ut fra det sted som er sannsynlig startpunkt for disse reisene, og at transportkvalitet ikke påvirker omfanget og fordelingen av de mellomliggende reisene direkte, men kun via utreise/retur for reiser uten mellomliggende reiser.