



Det här verket har digitaliserats vid Göteborgs universitetsbibliotek och är fritt att använda. Alla tryckta texter är OCR-tolkade till maskinläsbar text. Det betyder att du kan söka och kopiera texten från dokumentet. Vissa äldre dokument med dåligt tryck kan vara svåra att OCR-tolka korrekt vilket medför att den OCR-tolkade texten kan innehålla fel och därför bör man visuellt jämföra med verkets bilder för att avgöra vad som är riktigt.

This work has been digitized at Gothenburg University Library and is free to use. All printed texts have been OCR-processed and converted to machine readable text. This means that you can search and copy text from the document. Some early printed books are hard to OCR-process correctly and the text may contain errors, so one should always visually compare it with the images to determine what is correct.



Rapport

R36:1992

**Hushållsbaserade trafikmodeller
för konsekvensanalyser i flera
dimensioner**

**Metodik, uppläggning och resultat-
sammanfattning**

**Staffan Widlert
Staffan Algers**

V-HUSETS BIBLIOTEK, LTH



15000

400129236

Byggforskningsrådet

R36:1992

**HUSHÅLLSBASERADE TRAFIKMODELLER
FÖR KONSEKVENSANALYSER I FLERA DIMENSIONER**

Metodik, uppläggnig och resultatsammanfattning

Staffan Widlert
Staffan Algiers

Denna rapport hänför sig till forskningsanslag 820571-2
från Bygghforskningsrådet till Stockholms läns landsting,
Stockholm.

LUNDS TEKNISKA HÖGSKOLA
VÄG- OCH VATTENBYGGNAD
BIBLIOTEKET

REFERAT

Syftet med projektet har varit att utveckla ett modellsystem som:

- avser bilinnehav, trafikgenerering, områdesval och färdmedelsval
- innehåller samband mellan dessa olika val
- beaktar interaktioner inom hushållet som påverkar de olika valen
- avser alla typer av resänder
- beaktar restriktioner för möjligheterna att välja olika alternativ
- kan utnyttjas för att beskriva konsekvenser av olika åtgärder för olika socio-ekonomiska kategorier
- täcker hela Stockholms län

Modellsystemet tar sin utgångspunkt i teori för individens val mellan diskreta alternativ. I projektet används i allmänhet strukturerade logitmodeller där olika valsituationer beskrivs på olika nivåer i sammanhängande modellstrukturer. Den viktigaste datakällan utgörs av en resvaneundersökning som ger information om hur ett urval hushåll valt att resa under en viss mättdag.

Viktiga hypoteser har varit att trafiksystemet och tillgängligheten har betydelse för fler valbeslut än vad som tidigare kunnat visas empiriskt, att flera viktiga val rör hela hushållet och att hushållssamband därför bör modelleras explicit. Generellt har dessa centrala hypoteser bekräftats.

I Bygghörskningsrådets rapportserie redovisar forskaren sitt anslagsprojekt. Publiceringen innebär inte att rådet tagit ställning till åsikter, slutsatser och resultat.

Denna skrift är tryckt på miljövänligt, obektt papper.

R36:1992

ISBN 91-540-5488-5
Bygghörskningsrådet, Stockholm

Innehållsförteckning

	sida
Förord	7
1. Sammanfattning.....	11
Summary in English	17
2. Inledning	23
2.1 Avhandlingens uppläggning	23
2.2 Trafikmodellsystemets roll i trafikplaneringsprocessen	23
Att definiera trafikplaneringsproblemet	24
Att generera alternativa handlingsvägar	24
Att utvärdera alternativa handlingsvägar	25
Att besluta om handlingsväg	25
Trafikmodellsystemets roll	25
2.3 Syftet med avhandlingsarbetet.....	26
3. Utgångspunkter för trafikmodellsystemet.....	27
3.1 Syftet med ett trafikmodellsystem	27
3.2 Teoretiska utgångspunkter	27
En allmän modell för individers resbeteende	27
En modell för diskreta val	30
Konsistens med samhällsekonomisk analys	33
3.3 Hittillsvarande trafikmodellsystem.....	39
Svenska modellsystem	40
Utländska modellsystem.....	40
3.4 Brister i hittillsvarande modellsystem.....	41
Allmänna brister.....	41
Brister i trafikmodellsystem för Stockholms län	41
3.5 En allmän specifikation av ett trafikmodellsystem.....	42
Ett hushållsbaserat trafikmodellsystem	42
Interaktioner inom hushållet	42
Valdimensioner	43
Interaktion mellan olika resbeslut	44
Restriktioner	44
Oberoende variabler	44
Ärendeuppdelning	45
Numerisk metod.....	45
Ej behandlade aspekter	45
Modellsystemets avgränsning till Stockholms län.....	47

4.	Val av modelltyp.....	49
4.1	Sannolikhetsbaserad beteendeteori.....	49
4.2	Den multinomiala logitmodellen.....	52
	Härledning	52
	Den multinomiala logitmodellens egenskaper.....	53
4.3	Den strukturerade logitmodellen	57
4.4	Alternativa ansatser	62
	Probitmodellen.....	62
	Alternativa funktionsformer	63
5.	Estimeringsteknik.....	65
5.1	Val av skattningsmetod.....	65
	Maximum likelihood-metoden	65
	Några viktiga egenskaper hos maximum likelihoodestimatorn.....	67
5.2	Hypotestestning	68
	t-testet.....	69
	Likelihood ratio-testet	70
5.3	Tester av modellstrukturen.....	71
	Oberoende av irrelevanta alternativ	71
	Smakvariation.....	72
	Konstant varians	72
5.4	Risker med hypotesprövning i svårupprepade urval	73
5.5	Modellutveckling och statistiska tester i praktiken	73
5.6	Beräkningstekniska problem vid estimering av trädmodeller.....	74
5.7	Några speciella estimeringsproblem.....	79
	Storleksvariabler	79
	Sampling av alternativ.....	83
5.8	Kausalitet och korrelation	88
5.9	Tvärsnittsdata	88
6.	Datamaterial	91
6.1	Inledning	91
6.2	Resvaneundersökning	91
	Planering och uppläggning	91
	Huvudundersökningen	93
6.3	Bearbetning av resvanedata.....	97
	Huvudresa och huvudärende.....	97
	Hushållsturer.....	99
	Huvudfärdsätt	100
6.4	Trafiksystemdata	101
	Områdesindelning	101
	Nätdata	101
	Parkeringsdata.....	103
6.5	Områdesdata.....	104
	Ärendeindelning.....	106
	Storleksuppgifter.....	107

6.6	Mätproblem	109
	Nätdata.....	109
	Kostnadsuppgifter	109
7.	Validering	111
7.1	Validering baserad på estimeringsdata.....	111
7.2	Validering baserad på externa datakällor.....	111
	Samtida data.....	112
	Data för andra tidsperioder	113
8.	Ett modellsystem för Stockholms län	115
8.1	Arbetsresor	115
	Struktur.....	115
	Sekundära destinationer.....	117
	Färdmedelsval, bilallokering och resfrekvens	118
	Destinationsval och val av bilnehav.....	118
8.2	Skolresor.....	119
	Struktur.....	119
	Destinations- och färdmedelsval.....	120
8.3	Tjänsteresor	120
	Struktur.....	120
	Val av destination, färdmedel och restyp	121
	Val av resfrekvens.....	122
8.4	Inköpsresor	122
	Struktur.....	124
	Resultat.....	125
	Frekvensval.....	126
	Individ- och restypsval	126
	Färdmedels- och destinationsval.....	126
8.5	Service- och rekreationsresor	126
	Resultat.....	127
	Frekvensval.....	128
	Individ- och restypsval	128
	Färdmedels- och destinationsval.....	128
8.6	Besöksresor.....	128
	Struktur.....	129
	Resultat.....	130
	Frekvensval.....	131
	Individ- och restypsval	131
	Färdmedels- och destinationsval.....	131
8.7	Jämförelse av resultaten för olika resänder	131
	Resfrekvens.....	131
	Restyps- och individfördelning.....	132
	Destinationsval.....	133
	Färdmedelsval	133
8.8	Prognossystem	135

9. Överföring	137
9.1 Begreppet överförbarhet	137
9.2 Tidigare erfarenheter	139
9.3 Stockholmssystemets överförbarhet	139
10. Slutsatser och generella kommentarer	141
10.1 Allmänna slutsatser	141
10.2 Modellsystemets svaghet och styrka.....	141
Processen	141
Modellsystemet.....	142
11. Referenser	145

Förord

Resvaneundersökningen 1986/87 (RVU) är ett samarbetsprojekt mellan Regionplane- och trafikkontoret vid Stockholms läns landsting (Rtk) och AB Storstockholms Lokaltrafik (SL), med ekonomiskt stöd från Transportforskningsberedningen och Bygghörsrådet. Resvanedata har samlats in dels genom intervjuer med ca 6 500 individer i nära 3 000 hushåll (hushållsundersökningen), dels genom postenkäter med svar från drygt 30 000 individer (individundersökningen). Intervjumaterialet har under en längre tid använts vid trafikplanering och trafikanalys. Projektet hade redan från början två huvudsyften. Det ena var att ge ett underlag för deskriptiva analyser, dvs. ge en möjlighet att beskriva hur resandet i Stockholms län faktiskt ser ut. Det andra syftet var att ge underlag för att utveckla ett system av trafikmodeller. När väl datainsamlingen var genomförd har arbetet med de två olika huvudsyftena utförts i separata projekt med separat bemanning. De deskriptiva resultaten redovisas i rapporter från Rtk och SL (Mouwitz 1991, Tomth 1992).

Det ena huvudsyftet med projektet har således varit att utveckla ett system av trafikmodeller för Stockholms län. Den här föreliggande första delen av slutrapporten redovisar metoden som använts och de delar av uppläggningsen av projektet som är gemensam för de olika modeller som utvecklats. Slutrapportens första del syftar även till att ge en helhetsbild av de tankar och hypoteser som styr utformningen av projektet och sammanfattar också de principiellt viktigaste resultaten från de utförda analyserna. Rapporten har skrivits gemensamt av Staffan Algers och Staffan Widlert. De detaljerade resultaten från utvecklingen av modellerna för arbets-, skol- och tjänsteresor redovisas i slutrapportens andra del. Analyserna i andra delen har utförts av Staffan Widlert som också skrivit rapporten. Analysen av inköps-, besöks-, service- och rekreationsresor har utförts av Staffan Algers och redovisas av honom i slutrapportens tredje del. De tre modellrapporterna utgör två doktorsavhandlingar vid institutionen för trafikplanering vid Kungliga Tekniska Högskolan i Stockholm.

Projektet är mycket komplext och har en ovanligt stor omfattning. Resultatet har blivit det mest omfattande och avancerade system av modeller för individers och hushålls resbeteende som oss veterligen hittills utvecklats någonstans i världen. Den ovanligt långa projekt-tiden motiverar en återblick på projektets tillkomst och genomförande.

De första tankarna på ett nytt modellsystem för Stockholms län uppkom redan 1977 då Trafikkontoret vid Stockholms läns landsting organiserades om och en ny utredningsavdelning bildades. Avdelningen nyrekryterades i stort sett fullständigt och flera av oss som då anställdes där hade tidigare arbetat med modellutveckling. Redan vid det första arbetsplaneringsmötet innan vi hade börjat på den nya arbetsplatsen stod det klart att vi behövde genomföra en ny stor resvaneundersökning och att denna skulle användas för att utveckla modeller som var mer heltäckande - både i termer av de resärenden som skulle behandlas och av de valbeslut som skulle täckas in. Redan tidigt markerades behovet av en ny undersökning i kontorets flerårsbudget. Avsikten var att undersökningen skulle påbörjas hösten 1984.

Stora resvaneundersökningar är dyrbara att genomföra. De första kostnadsberäkningarna pekade på en totalbudget på 7 mkr. Ett omfattande arbete fick därför läggas ner på finansieringen av projektet. Eftersom projektet innehöll betydande moment av ren forskning söktes medel från Statens Råd för Byggnadsforskning (BFR) och från Transportforskningsdelegationen (nuvarande Transportforskningsberedningen, TFB). I samband med dessa ansökningar utfördes en detaljerad planering av det modellsystem som skulle utvecklas. En betydande del av uppläggningsarbetet tog form under ett par intensiva veckor på en hösttom skärgårdsö. Redan då stod det klart att modellerna skulle bli komplicerade - vi kom tillbaka från skärgården med en modellstruktur som inte ens vår mest fanatiska modellanhängare trodde skulle vara möjlig att förverkliga.

Arbetet redovisades i promemorieform (Algers och Widlert 1982). Forskningsråden godkände i princip forskningsansökningarna samma år. Därefter följde en tid av diskussioner med andra forskare inom och utom landet. Dessa resulterade i en reviderad uppläggning som beskrevs i projektets första planeringsrapport (Algers och Widlert 1983). Hösten 1983 genomfördes den första av totalt tre provundersökningar. Den sista utfördes försommaren 1985. Huvudundersökningen påbörjades våren 1986 och pågick ett år med avbrott för sommar och jul.

Fram till och med våren 1986 utförde vi planeringen av analysarbetet vid landstingets trafikkontor. Planeringen av datainsamlingen utfördes gemensamt av trafikkontoret och SLs avdelning för trafik- och marknadsundersökningar. Sommaren 1986 flyttade vi själva till Utvecklingsstaben på SL. Därefter har arbetet med modellsystemet helt försiggått på SL.

När väl data började komma in finslipades förslaget till uppläggning av modellarbetet. Detta slutgiltiga förslag redovisades i Planeringsrapport 5 som kom i december 1986 (Algers och Widlert 1986a). Rapporten har fungerat som en detaljerad arbetsplan för modellutvecklingen och den uppläggning som beskrivs i rapporten har kommit att stå sig mycket väl arbetet igenom. Att omvandla de data som samlats in till en form som passar för modellerna och att ta fram trafiknät och andra bakgrundsuppgifter visade sig vara betydligt mer tidskrävande än vi hade väntat oss. Först hösten 1988 lämnade de första modellerna datorn.

När datamaterialet var klart för analyser började en lång period av problem med det program som användes för analyserna. Programmet - ALOGIT från Hague Consulting Group i Holland - hade använts i en lång rad tidigare tillämpningar. I samband med arbetet på vårt projekt - och delvis bekostat av projektet - utvecklades estimeringsprogrammet så att det kunde klara av att samtidigt estimeras flera nivåer i så kallade strukturerade logitmodeller. Detta gjorde det möjligt att estimeras väsentligt mer komplexa modeller än tidigare. De modeller vi skisserat var också större (mätt i t.ex. antalet alternativ i modellen) och mer komplicerade än några tidigare modeller. Avigsidan visade sig snabbt i form av ett antal programproblem som endast gjorde utslag vid så komplicerade modeller. Även när dessa problem så småningom var helt lösta tog modellutvecklingen betydande tid. Först under våren 1992 kunde estimeringsarbetet avslutas definitivt.

Parallellt med arbetet på detta projekt har vi arbetat med andra forskningsprojekt inom området vilka fungerat som förstudier av vissa aspekter på modellsystemet. Med hjälp av data från Göteborg utvecklade vi dels enkla färdmedelsvalsmodeller (Algers, Tegnér och Widlert, 1984), dels bilinnehavs- och bilallokeringsmodeller (Algers och Widlert, 1986b). Dessa senare modeller testade direkt olika hypoteser som vi ville bringa klarhet i inför

RVU-projektet. I projektet "Logitmodellen - användbarhet och generaliserbarhet" användes tidigare insamlade svenska datamaterial för att studera ett stort antal olika frågeställningar (Algers, Colliander och Widlert 1987, Algers och Widlert 1987b). Bland annat prövades samtidig estimering av flera nivåer i en logitmodell för första gången. Dessutom prövades bland annat olika strategier för att koda på alternativ, olika sätt att formulera attraktivitetsvariabler, betydelsen av att arbeta med olika strukturer för olika valbeslut, aggregeringsfelens betydelse, hanteringen av gång- och cykelalternativ och alternativa definitioner av valgrupper i modellerna. Resultaten kom i hög grad att påverka uppläggningsen av RVU-projektet.

De många åren har gjort det möjligt att successivt och även internationellt diskutera uppläggningsen och resultat med olika forskare. Vid sex tillfällen har uppläggningsen och/eller resultat presenterats vid internationella konferenser (Algers och Widlert 1985, 1987a, Algers 1985, Algers, Daly och Widlert, 1990, 1991 och 1992). Fortfarande återstår mycket att publicera. Ett antal ytterligare 'papers' kommer därför att skrivas de närmaste åren. Inom landet har uppläggningsen och resultat presenterats vid ett stort antal tillfällen, bland annat vid flera av Väg- och trafikinstitutets forskardagar. Dessutom har Forskargruppen för Trafikberäkningsmodeller, vars verksamhet delvis finansierats av TFB, utgjort en referensgrupp som fortlöpande diskuterat uppläggningsen och resultat under projektets utveckling. Även landstingets trafikanalysgrupp, som innehåller representanter för olika regionala organ, har successivt tagit del av projektets resultat.

Detta förord skrivs sent på hösten 1991 vid en ny frivillig förvisning till en blåsig, regnig och hösttom skärgård. Valet av författarlya illustrerar ytterligare en faktor som påverkat tidsplanen under hela projektet. Att bedriva forskningsarbete parallellt med vanligt arbete vid en offentlig myndighet är inte utan sina komplikationer. Den dagliga verksamheten ställer ofta krav som måste åtgärdas genast och det är svårt att tillräckligt prioritera det långsiktiga arbetet. Behovet av att fly undan blir därför till slut akut.

Ett ytterligare skäl till att projektet har tagit 15 år från första idé till färdig rapport är att det har handlat om "tillämpad forskning på grundforskningsnivå". Vi har hela tiden haft möjligheten att välja den ambitionsnivå vi själva önskat. Under arbetets gång har betydande tid använts för att utforska uppslag som senare har visat sig vara sidospår, som inte kommit till praktisk användning. Detta har lämnat oss med glädjen av att ha fått arbeta med ett projekt där inte tids- och budgetrestriktioner ständigt stoppat alla nya idéer. Idag när vi båda övergått till konsultverksamhet och projekttiden oftast inrymmer ungefär lika många veckor som detta projekt tagit år framstår denna period som en tid vi inte skulle velat vara utan, men som vi knappast vill ha igen.

Så här vid slutet av vårt arbete vill vi också tacka ett stort antal personer. Framför allt vill vi rikta ett tack till Andrew Daly från Hague Consulting Group. Han har spelat en stor roll som rådgivare och lösare av estimeringsproblem under hela projektet och har i praktiken kommit att fungera som en extra handledare åt oss. Andrew har dessutom kontinuerligt fått vidareutveckla och korrigera den programvara för att estimeras modeller som utnyttjats för att få denna att klara av de stora modeller vi önskat bygga. Vi vill också tacka våra chefer under de här åren: Olle Nordell, Göran Tegnér, Bo Malmsten och Bosse Wallin. Ni har inte bara låtit oss hållas alla dessa år - alltså oftast har ni till och med uppmuntrat oss! Utan Görans entusiasm och ihärdiga arbete för att klara finansieringen av projektet hade detta aldrig blivit av.

Det praktiska arbetet med projektet har involverat många personer. Kerstin Blomqvist, Kjell Falkborn och flera andra har fått slita sina hår vid databearbetningen. Gunilla Rosenqvist hade det praktiska ansvaret för hushållsdelen av resvaneundersökningen och var den som såg till att kvaliteten blev så hög som den faktiskt blev. Gunilla Rosenqvist och Kjell Jansson har även medverkat i arbetet med estimeringen av modellerna för val av sekundära destinationer. Jan Mouwitz som är ansvarig för avdelningen för marknadsundersökningar vid SL har inte bara lagt ner ett stort arbete på uppläggningsen av undersökningen, han har också fått bära sin personals vånda över alla våra galna idéer. Monika Fredriksson och Paula Darke arbetade med provundersökningarna och uppläggningsen av både individ- och hushållsundersökningen. De svarade dessutom för genomförandet av individundersökningen.

Vi vill också varmt tacka vår handledare Åke Claesson, samt alla de forskare inom och utom landet som gett oss idéer och som diskuterat våra egna under åren.

Ett särskilt tack vill vi rikta till Lars-Göran Mattsson som påpekat åtskilliga fel i formlerna och som också lämnat en rad synpunkter på resultaten.

Tack också Ulla, Börje och Kjell. Ni har tåligt läst våra regelbundet återkommande nya versioner av avhandlingen och hjälpt oss att leta fel och inkonsistenser.

Ni, och alla ni andra som deltagit i arbetet, har gjort projektet möjligt. Ansvaret för eventuella återstående fel ligger naturligtvis helt hos oss själva.

Nu Lena, Pia och fröna kommer avgörandets ögonblick: kommer ni att få se mer av oss framöver eller inte? Tack för att ni har väntat!

Staffan & Staffan

1. Sammanfattning

Avhandlingen redovisar arbetet med ett modellsystem för persontrafikresorna i Stockholms län. Denna första del beskriver metoden och det datamaterial som använts, samt sammanfattar vissa resultat. De empiriska resultaten från modellarbetet redovisas detaljerat i separata rapporter.

Syftet med avhandlingsarbetet

Syftet med avhandlingsarbetet är att testa hypoteser om betydelsen av olika faktorer och kopplingar på ett modellsystem som omfattar alla typer av personresor. Målet har varit att modellsystemet ska:

- avse bilinnehav, trafikgenerering, områdesval och färdmedelsval
- innehålla samband mellan dessa olika val
- beakta interaktioner inom hushållet som påverkar de olika valen
- avse alla typer av resänder
- beakta restriktioner för möjligheterna att välja olika alternativ
- kunna utnyttjas för att beskriva konsekvenser av olika åtgärder för olika socio-ekonomiska kategorier
- täcka hela Stockholms län

Avhandlingsarbetet syftar i första hand till att vara ett empiriskt forskningsarbete. Detta innebär att vi avgränsar oss till att använda den teoribildning som är accepterad inom de discipliner som metodiken hämtats från.

Utgångspunkter för trafikmodellsystemet

Modellsystemet tar sin utgångspunkt i teori för individers val mellan diskreta alternativ (dvs. val mellan ett ändligt antal alternativ). Ansatsen förutsätter att individerna är nytto-maximerande och att de är rationella, dvs. att de gör medvetna och överlagda val utifrån sina egna värderingar av alternativen. En viktig önskemål är att resultaten ska vara konsistenta med den ekonomiska analysapparatur som normalt används vid samhällsekonomiska analyser.

Bilnehavet är ett långsiktigt val med stor betydelse för resbeteendet. Vår hypotes är att bilnehavet dels beror av hur det påverkar tillgängligheten för flera hushållsmedlemmar, och dels beror av hela hushållets ekonomi, varför vi formulerar det som ett hushållsval. I detta avseende är således allt resande hushållsbaserat.

Även för att kunna hantera interaktioner inom hushållet måste modellsystemet i princip vara hushållsbaserat. Det innebär inte, att alla modeller är hushållsbaserade - när det gäller skolresor och tjänsteresor dominerar interaktion mellan individer i hushållet och exogena

aktörer, varför dessa resor behandlas som individbaserade (dock med det hushållsbaserade bilinnehavet som utgångspunkt).

Gängse teoribildning utgår från *individuell* nyttomaximering. Som framgått utgår vi ofta från *hushållet* som beslutande enhet, vilket innebär att vi maximerar någon slags funktion av individuella nyttor - bara individer kan uppleva nytta, och hushållet är ju i detta avseende endast en organisationsform. Avhandlingen diskuterar hur denna process kan ses som en *förhandling* inom hushållet där den slutliga överenskommelsen utgör den för alla parter bästa möjliga lösningen i en given situation. Den representerar därför den för varje individ maximala nyttan, givet den restriktion som beroendet mellan olika hushållsmedlemmar ger. Man skulle därför kunna tala om en beslutsfunktion (eller en välfärdsfunktion för hushållet), vilken dels har individuella nyttor som argument och dels avspeglar hushållets förhandlingssituation.

Denna beslutsfunktion pekar ut vilket alternativ individerna i hushållet kommer överens om, och innebär samtidigt att ingen individ kan förbättra sin situation genom att efterfråga ett annat alternativ - utom genom att säga upp det "kontrakt" som hushållet innebär. Detta synsätt innebär alltså att vi ansätter en beslutsfunktion för hushållet, som är konsistent dels med individuell nyttomaximering och dels med de restriktioner - och möjligheter - som hushållet innebär.

Val av modelltyp

Eftersom vi inte säkert kan beräkna nyttan av de olika alternativ som individerna och hushållen väljer mellan behandlas nyttorna som stokastiska variabler. Den stokastiska komponenten i nyttorna orsakas av sådana faktorer som att vi inte kan observera alla attribut som beskriver alternativen, att det finns ej observerade skillnader i individernas preferenser och erfarenheter, att det förekommer mätfel i våra indata, att vi måste använda proxyvariabler för att beskriva vissa faktorer, etc.

Olika modeller kan härledas utifrån olika antaganden om fördelningen hos denna stokastiska komponent. I vårt fall används *logitmodellen* som förutsätter att de stokastiska komponenterna har en speciell fördelning (den så kallade Gumbelfördelningen). I projektet används i allmänhet *strukturerade logitmodeller* där olika valsituationer beskrivs på olika nivåer i sammanhängande modellstrukturer (t.ex. beskrivs valet av destination, färdmedel och resfrekvens på olika nivåer). Inom varje sådan nivå förutsätts de stokastiska komponenterna vara oberoende och lika fördelade. Detta krav behöver ej vara uppfyllt mellan nivåerna.

Modellerna skattas med maximum likelihood-metoden.

Datamaterial

För att skatta modellsystemet krävs betydande mängder indata. I detta projekt används uteslutande uppgifter om hur individer och hushåll faktiskt betett sig när de stått inför en valsituation (det vill säga uppgifter om "Revealed Preferences"). Den viktigaste delen utgörs därför av en *resvaneundersökning* som ger information om hur ett urval hushåll valt att resa under en viss mätdag. Modellerna bygger på en undersökning som omfattar inter-

vjuer med ca 3 000 hushåll i Stockholms län. Undersökningen genomfördes under 1986 och 1987. Huvudundersökningen föregicks av omfattande provundersökningar. Stora ansträngningar lades ned på att få ett datamaterial av hög kvalitet. Svarsprocenten blev också hög (80%) och partiella bortfall av hushållsmedlemmar eller av svar på enstaka frågor var sällsynta.

Utöver uppgifter om resandet under en viss dag krävs även uppgifter om vad hushållen kunde gjort istället, det vill säga vilka alternativ de stod inför. Delvis får vi denna information från resvaneundersökningen (där vi kan se om hushållet har tillgång till bil, om personen ifråga har körkort, etc.). En ytterligare viktig källa till information om de olika alternativen utgörs av *utbudsregister* som ger information om de möjligheter som finns att utföra olika aktiviteter i olika målområden. För detta syfte används framför allt uppgifter från folk- och bostadsräkningarna.

Tillgängligheten till olika målområden enligt attraktivitetsregistren ges i sin tur av *trafiksystemdata* som huvudsakligen hämtas från trafiknät som kodats för projektet.

För att få en bild av möjligheterna att parkera i olika områden har ett särskilt *parkeringsregister* byggts upp för projektet.

Ett modellsystem för Stockholms län

Projektet innebär bland annat att ett system av trafikmodeller byggts upp. Separata modellstrukturer har skattats för följande slag av resärenden:

- arbetsresor
- skolresor
- tjänsteresor
- inköpsresor
- resor för service och rekreation
- besöksresor

De olika modellstrukturerna täcker delvis in olika typer av valsituationer. De allra flesta behandlar färdmedels-, destinations- och frekvensval. Dessutom förekommer bilinnehavsmodeller, modeller för hur hushållet fördelar användandet av bilen mellan de olika hushållsmedlemmarna om flera konkurrerar om den, modeller för hur reskedjor bildas och för hushållets beslut om vem som ska göra vissa typer av ärenden, etc. Ambitionen har varit att bygga upp ett så heltäckande system som möjligt för det icke yrkesmässiga resandet, både med avseende på studerade resärenden och behandlade valbeslut.

De viktigaste resultaten från modellarbetet

De viktigaste resultaten är:

- ett modellsystem som beaktar kopplingen mellan olika hushållsmedlemmar, olika resärenden och olika typer av resor
- ett modellsystem som gör det möjligt att beskriva konsekvenser av olika åtgärder i flera dimensioner

Huvudresultatet uttrycks i form av de skattade parametervärdena vilka återfinns i de två delrapporterna som redovisar det empiriska arbetet.

Ur modellerna kan bland annat tidsvärden och vikter härledas. Samtliga tidsvärden, utom värdena för tjänsteresor, ligger i intervallet 15-25 kronor per timme och dessa tidsvärden är inte signifikant skilda från varandra. Tjänsteresornas tidsvärden uppgår till 100 - 130 kronor per timme. Resultaten ansluter väl till resultaten från tidigare studier. Även de vikter för olika tidskomponenter som kan beräknas ur modellerna ansluter relativt väl till tidigare resultat.

I rapporterna beskrivs de modeller som utvecklats. Modellerna har dock även byggts in i ett prognosystem för att göra det lättare att utvärdera effekten av olika åtgärder. Prognossystemet innebär att modellerna körs med hjälp av befolkningsdata för de 850 områden som länet delats in i. Befolkningen i varje område delas in i ett stort antal olika grupper vilka definieras olika för olika resänder. Modellerna tillämpas därefter för varje sådan grupp i varje område för respektive resänder.

Överföring

De begränsade erfarenheterna av överföring av hela modellsystem tyder på att det är möjligt att tillämpa de modeller som utvecklats i detta projekt för andra orter. Överföringen innebär normalt någon form av anpassning av modellerna. Tidigare erfarenheter visar att trafikstandardvariabler, det vill säga parametrar för olika restidskomponenter och kostnader, är överförbara till andra platser. Även parametrar för socioekonomiska variabler går i allmänhet att överföra. Överförbarheten hos de parametrar som direkt speglar geografiska förhållanden är sannolikt sämre.

Modellerna i detta projekt är dock betydligt mer komplicerade än de som tidigare överförts och många av de samband som modellerna täcker in kan se olika ut på olika orter. Modellerna behandlar explicit interaktioner och restriktioner inom hushållen. Sådana samband kan förväntas vara rimligt överförbara inom landet, men det är ännu en öppen fråga hur väl överförbara sambanden är mellan olika länder.

Slutsatser

Vårt arbete har styrts av ett antal hypoteser om vad som påverkar individernas och hushållens beteende. De viktigaste av dessa formulerades redan ganska exakt i form av förslag till modellstrukturer i planeringsrapporter som gavs ut i början av arbetet med projektet. Hypoteserna handlar om att trafiksystemet och tillgängligheten har betydelse för fler valbeslut än vad som tidigare kunnat visas empiriskt, att flera viktiga val rör hela hushållet och att hushållssamband därför bör modelleras explicit.

Generellt kan vi konstatera att våra centrala hypoteser har bekräftats. Detta tar sig uttryck i att de modellstrukturer som erhållits överensstämmer mycket väl med dem som föreslogs i planeringsrapporterna. Parametrar som direkt speglar om våra hypoteser styrkts eller ej är bland annat de logsumparametrar som för vidare tillgänglighetens betydelse genom modellstrukturerna. Dessa parametrar har kunnat estimeras framgångsrikt. Det är särskilt

uppmuntrande att detta varit möjligt trots att modellernas komplexitet (främst antalet alternativ) tvingat oss att dela strukturerna på flera ställen och estimerade dessa delar sekventiellt. Att logsumparametrarna blivit signifikanta även i de delmodeller som behandlar hushållsinteraktioner visar på existensen av dessa val och att de påverkas av trafiksystemets utformning.

Andra resultat som visar att modellarbetet lyckats är att modellerna innehåller de viktigaste förklaringsvariablerna som vi utifrån tidigare genomförd forskning har anledning tro påverkar trafikanternas beteende, samt att de dessutom innehåller ett vitt spektrum av förklaringsvariabler vars betydelse för beteendet kan förklaras a priori.

Modellsystemet är oss veterligen det mest fullständiga som hittills utvecklats någonstans i världen. Systemet behandlar en stor del av det totala resandet och framför allt behandlas fler valbeslut än vad som tidigare varit möjligt. Systemet har sin främsta styrka i att det behandlar individernas och hushållens resande mer fullständigt och realistiskt än vad som tidigare varit fallet. Samtidigt är det ett system som kan användas - och redan används - för praktisk trafikplanering.

Även om modellsystemet är mer fullständigt och realistiskt än tidigare utvecklade system innebär den realiserade strukturen avsevärda förenklingar i beskrivningen av beteendet. Dessa förenklingar gäller bland annat bildandet av reskedjor, kopplingen mellan olika resärenden och beskrivningen av destinationsvalet. Styrande för de förenklingar som gjorts har varit kravet på att systemet ska vara möjligt att använda praktiskt.

Summary in English

This work forms two theses that documents the work with a model system for personal traffic in the county of Stockholm. This first part describes the method and the data, and summarises some of the results. The empirical results from the modelling work are described in full detail in separate parts, which also contain English summaries.

The purpose of the thesis

The purpose of the thesis is to test hypotheses about the importance of different factors and relationships on a model system that covers all types of person trips. The purpose has been that the model system should:

- cover car-ownership, traffic generation, choice of destination and mode
- deal with connections between those choices
- consider interactions within the households affecting different choices
- cover all trip purposes
- consider restrictions on the possibilities to choose different alternatives
- be possible to use for analysis of distribution effects for different groups of people
- cover the whole county of Stockholm

The primary aim of the thesis is to be an empirical study. This implies that we use already known and generally accepted theory.

Basis of the model system

The model system is based on theory for individuals' choice between discrete alternatives. It is assumed that the individuals are utility-maximising and rational, i.e. that they make deliberate choices, based on their own valuation of the alternatives. An important prerequisite is that the results should be consistent with economic theory and its use in project evaluation.

Car ownership is a long term decision with important implications for travel behaviour. Our hypothesis is that car ownership on one hand is affected by the accessibility it gives the whole household, and on the other hand depends on the economy of the whole household. Therefore, the car ownership model is formulated as a household model. In this respect, it could be argued that all travel is household based.

Household based models are also in principle necessary if interactions within the households are to be modelled. This does not mean that all models are household based - for school trips, business trips, and social trips, interactions between individuals and exogenous actors dominate. Those trip purposes are modelled with individual-based models (but with the household based car ownership as a basis).

The generally accepted theory assumes *individual* utility maximisation. As stated above, we often regard the *household* as the decision making unit. This implies that we maximise some function of individual utilities - only individuals can experience utility, and the household is only a form of organisation. In the thesis, we discuss how this process can be viewed as a *negotiation* within the household, where the final agreement represents the best possible solution in a given situation. For every household member it represents the maximum utility given the restriction imposed by the dependence between the different household members. This can be regarded as a decision function, or a social utility function for the household, which has individual utilities as arguments and which also reflects the negotiation situation in the household.

This decision function shows which alternative the household will agree upon. It also implies that no member of the household can improve his situation by demanding another alternative - without at the same time cancelling the "contract" that the household implies. In summary, we assume a decision function that is consistent with both individual utility maximisation and the restrictions - and possibilities - that forming a household involves.

Type of model

We cannot with certainty calculate the utility of the different alternatives that individuals and households choose between. These utilities are therefore treated as random variables. The random component in the utilities is caused by the fact that we cannot observe all attributes that characterise the alternatives, that there are unobserved differences between different individuals' preferences and experiences, that there are measurement errors in the data, that we have to rely on proxy variables to describe some factors, etc.

Different models can be derived from different assumptions about the distribution of the random component. We use the *logit model* which assumes that the random component has a specific distribution (the Gumbel distribution). Most models are *structured logit models* where different choices are modelled on different levels in a structure (where for example the choices of destination, mode and frequency are modelled at different levels). Within each such level, the random components are assumed to be identically and independently distributed. This does not have to be the case between the levels.

The models are estimated with maximum-likelihood techniques.

Data

A large amount of data is needed to estimate the model system. In this project, only data about the individuals' and the households' *actual* behaviour are used ("Revealed Preference-data"). The most important part is a travel study that gives information about all trips made by all members of the household of at least 12 years of age, during one day. The sample for the travel study consisted of 3 000 households in the county of Stockholm, and the study was carried out during 1986 and 1987. The main study was preceded by three large pilot studies. Great effort was put into the data collection to ensure a high quality. The final response rate was 80% and very little partial non-response occurred (for

example missing members in the households, or non-response for certain questions in the forms).

Beside information about all trips during one day, information was also collected about how they could have travelled instead, that is, information about the alternatives. Partly, this information was collected in the travel study (where we can see if the household has access to a car, if each person has a driving license, etc.). One further important source of information about the alternatives was a *supply register* which gave information about possibilities to perform different activities in different areas. For this purpose, mainly census data were used.

The accessibility to different areas by different modes was given by *network data*, which were mainly taken from networks coded specifically for this project.

To get a picture of the possibilities to park in different areas, a special parking register was built up for the project.

A model system for the county of Stockholm

In the project, a comprehensive system of traffic models has been built up. Separate structures have been estimated for the following trip purposes:

- work
- school
- business
- shopping
- service and recreation
- social trips

The different model structures cover different types of choice situations. Most of them cover choice of mode, destination and frequency. In addition to those, there are models for car ownership, models for the allocation of the car between different household members, models for how trip chains are formed and models describing which member of the household who is likely to make different trips. The ambition has been to build a system for personal traffic that is as comprehensive as possible, both in terms of included trip purposes and in terms of included choice situations.

Main results from the modelling work

The most important results are:

- a model system that takes into account interdependencies between different household members, different trip purposes and different types of trips
- a model system that makes it possible to describe consequences of different measures in several dimensions

The main results consist of estimated parameter values which can be found in the two parts describing the empirical work.

From the models, it is possible to calculate values of time and weights for trip time components. All in-vehicle time values - except those for business trips - are in the interval 15-25 SEK per hour. The time value for business trips is in the interval 100-130 SEK per hour. The result is fairly consistent with previous studies. This is also true for the weights for different time components.

The models have also been implemented in a forecasting system. The models are used with population data for the 850 zones in the county. The population in each zone is divided into a large number of groups which differ for different purposes. The models are applied for each such group in each zone for each trip purpose.

Transferability

The limited experience of transferring complete systems of models indicates that it is possible to use the models that have been estimated in this project in other geographical areas. Such model transfers typically involve some adjustment of the models. Previous experience shows that parameters for level-of-service (time and cost parameters) typically are transferable. Also parameters for socio-economic groups are usually possible to transfer. Parameters that reflect characteristics of geographical areas are less transferable.

The models in this project are much more complicated than the models that have been transferred in previous projects. Many of the interactions that the models describe might differ between different areas. Restrictions that are explicitly modelled might be transferable within the country, but might be different in different countries.

Conclusion

Our work has been guided by a number of hypotheses about factors that affect the travel behaviour of individuals and households. The most important ones were formulated in the form of model structures in planning reports that were written at the beginning of the work. Important such hypotheses are that the traffic system and the associated accessibility affect more choice decisions than has been possible to show up to now, and that several important choices concern the whole household and should therefore be modelled as household choices.

In general, our hypothesis have been confirmed. This is shown by the fact that the final model structures agree closely with the ones in the planning reports. Parameters that are central for this conclusion are logsum parameters that take account of the influence of accessibility through the structure. These parameters have mostly been successfully estimated. This is especially encouraging since we have been forced to estimate some parts of the structures sequentially.

Other successful results are that the models include the most important explanatory variables, according to previous research, and that they include a wide spectrum of relevant variables.

As far as known to the authors, this model system is the most elaborate transportation model system in the world in terms of integrated behavioural dimensions. The system deals with a large proportion of all trips, and it covers more trip related decisions than has previously been possible. The system has its major strength in that it deals more realistically with travel behaviour than previous systems. At the same time, it is a system that can be used for practical planning - and in fact already is.

Although the system is more comprehensive and realistic than previous systems, the chosen structure involves substantial simplifications of actual behaviour. Those simplifications are for example found in the area of trip chaining, connections between different trip purposes and in the treatment of destination choice. Guiding for the simplifications that have been made is the fact that the model system must be possible to use for practical purposes.

2. Inledning

2.1 Avhandlingens uppläggning

Avhandlingen består av tre delar. I första delen redovisas avhandlingsarbetets syfte och metodmässiga grund. I andra delen redovisas de trafikmodeller som avser arbets-, tjänste-, och skolresor, och i tredje delen redovisas trafikmodellerna för inköps-, service-, rekreations- och besöksresor.

Första delen innehåller i kapitel 1 en sammanfattning av hela avhandlingsarbetet. I kapitel 2 redovisas dels en bakgrund till avhandlingsarbetet, och dels avhandlingsarbetets syfte. I kapitel 3 diskuteras de krav som ställs på utvecklandet av ett trafikmodellsystem för Stockholm, dels från vetenskapliga utgångspunkter och dels med avseende på trafikmodellsystemet i sig. I kapitel 4 diskuteras den valda modelltypen, medan kapitel 5 innehåller en redovisning av tekniken för att skatta den valda modellen. I kapitel 6 redovisas det datamaterial som legat till grund för modellarbetet, och i kapitel 7 diskuteras valideringen av modellsystemet.

Kapitel 8 innehåller en sammanfattning av de empiriska resultaten av modellarbetet. Möjligheterna att utnyttja resultaten för andra orter diskuteras i kapitel 9. Avslutningsvis redovisas slutsatser och allmänna kommentarer i kapitel 10.

En närmare redovisning av innehållet i andra respektive tredje delen återfinns i respektive del.

2.2 Trafikmodellsystemets roll i trafikplaneringsprocessen

Avhandlingsarbetets titel är "Hushållsbaserade trafikmodeller för konsekvensanalyser i flera dimensioner". Med detta avser vi ett system av kvantitativa trafikmodeller för personresor, som ger möjlighet att beräkna effekterna på resandet av olika åtgärder, inom eller utom trafiksystemet. Syftet med avhandlingsarbetet är att utveckla ett sådant trafikmodellsystem, som på ett mer utvecklat sätt än vad som hittills skett fyller denna uppgift.

Avhandlingsarbetet har utförts inom ramen för ämnet trafikplanering. Vi vill därför inledningsvis försöka klargöra sammanhanget mellan vårt avhandlingsarbete och trafikplaneringsprocessen i stort. För att göra det måste vi först beskriva olika delar i trafikplaneringsprocessen, och sedan relatera avhandlingsarbetet till relevanta delar.

Inom all planering måste i allmänhet följande fyra steg behandlas:

- att definiera ett problemområde
- att generera olika alternativa handlingsvägar
- att utvärdera dessa handlingsvägar
- att besluta vilken handlingsväg som ska väljas

Genomförandet av planförslaget ligger med detta synsätt utanför planeringen, även om möjligheterna att genomföra planförslaget naturligtvis måste beaktas som en del av planeringsförutsättningarna. Med handlingsvägar avses enstaka åtgärder eller hela åtgärds paket.

Samhällsplaneringen är också en politisk process, vilket innebär att de olika stegen i olika grad får en politisk styrning. De politiskt viktigaste delarna av samhällsplaneringen är problemformuleringen och beslut om vilka handlingsvägar som ska väljas.

Att definiera trafikplaneringsproblemet

När det gäller trafikplanering är det första steget, problemdefinitionen, delvis redan klart - trafiken utgör på något sätt problemområdet. Med trafik avser vi här alla former av förflyttningar, av personer eller gods. Det återstår naturligtvis att sedan beskriva på vilket sätt trafiken är ett problem.

Det mest uppenbara problemet är trafikens kostnadssida - såväl samhället som den enskilda har stora utgifter för trafik av olika slag. Kostnader förekommer också i form av tidsåtgång, olyckor, miljöpåverkan, estetisk påverkan, barriäreffekter etc. Samtidigt finns det en intäktssida - trafiken möjliggör ju att olika aktiviteter kan komma till stånd, vilket har ett värde för enskilda, näringslivet och offentliga verksamheter.

Intäktssidan utgörs således av att personer eller gods kan sammankopplas med aktiviteter som har en annan rumslig fördelning. Detta innebär också att lokaliseringen av dessa aktiviteter utgör ett annat problemområde som har ett nära samband med trafikplaneringen.

Vi kan således konstatera att trafikplaneringen har att hantera en kostnads- och en intäktssida, samt att det finns ett nära beroende mellan trafik och lokalisering.

Att generera alternativa handlingsvägar

Det andra steget i planeringsprocessen, att generera alternativa handlingsvägar, innebär att finna olika sätt att minska kostnaderna och/eller öka intäkterna. Detta inrymmer såväl tekniska som administrativa och ekonomiska åtgärder. Möjligheterna att generera intressanta alternativ beror i stor utsträckning på kunskaperna om samband mellan åtgärder och effekter i olika avseenden.

De olika handlingsvägarna kan vara enstaka åtgärder eller hela åtgärds paket, analyserade i kombination med förutsättningar som inte är åtgärdsberoende. Exempel på sådana åtgärder och förutsättningar kan vara:

- förhållanden utanför trafikplaneringen
- utveckling av disponibel inkomst
- energipriser
- bebyggelseutveckling
- hushållsstruktur

övergripande trafikplaneringsåtgärder

- taxepolitik
- generella kollektivtrafikåtgärder, som t.ex. ökad turtäthet
- parkeringspolitik
- skattepolitik, t.ex. resavdrag
- restriktioner mot biltrafiken, t.ex. bilavgifter

specifika utbudsförändringar

- nya, ändrade eller indragna kollektivtrafiklinjer
- nya väglänkar
- busskörfält

Att utvärdera alternativa handlingsvägar

Det tredje steget - att utvärdera alternativa handlingsvägar - innebär dels att beskriva konsekvenserna av olika åtgärder i intressanta dimensioner, och dels att - så långt möjligt - göra en sammanvägning av för- och nackdelar. Denna sammanvägning görs normalt med en etablerad teknik, nämligen samhällsekonomisk kostnads/intäktsanalys (SKI). Även detta ställer krav på kunskaper om samband mellan åtgärder och effekter på resandet.

För att beräkna många av effekterna är det dock inte tillräckligt att känna sambandet mellan åtgärder och antalet resor, eftersom antalet resor kan ge upphov till ytterligare effekter som kan vara av stor betydelse. Exempel på detta är trafikbuller och avgasutsläpp. För att beakta sådana effekter krävs kunskaper om samband mellan antalet resor av olika slag och andra effekter.

Att besluta om handlingsväg

Det sista steget i trafikplaneringsprocessen innebär att besluta om vilka åtgärder som ska genomföras, vilket ofta är ett rent politiskt beslut. Detta kan ställa krav på en intressentorienterad redovisning av de olika handlingsvägarna, vilket ställer ytterligare krav på kunskaper om sambandet mellan åtgärder och resande.

Trafikmodellsystemets roll

Som påpekats ovan, är kunskaper avseende sambandet mellan olika handlingsvägar och resandet en god hjälp när det gäller att finna effektiva alternativa handlingsvägar för att lösa olika trafikproblem. De är dock framför allt en förutsättning för att över huvud taget kunna göra en seriös utvärdering av olika handlingsvägar, under olika yttre förutsättningar.

Dessa kunskaper, formulerade som kvantitativa samband, kallas ofta trafikmodeller. För att kunna behandla ett visst problemområde - exempelvis personresor i tätort - behövs ofta flera olika modeller, som sammantagna utgör ett verktyg. Ett trafikmodellsystem är således ett antal sammanhörande kvantitativa modeller för att beräkna hur resmönstret påverkas av olika faktorer.

Det är uppenbart att förekomsten av och kvaliteten på trafikmodellsystem är avgörande för möjligheterna att belysa centrala frågeställningar inom trafikplaneringen.

2.3 Syftet med avhandlingsarbetet

Bristerna i hittillsvarande modellsystem har gjort det till en angelägen uppgift att förbättra såväl tillgång som kvalitet på trafikmodellsystemen. Det är vår förhoppning att avhandlingsarbetet ska ge ett påtagligt bidrag i detta avseende, genom att för persontrafikresor i tätort skapa ett "Hushållsbaserat trafikmodellsystem för konsekvensbeskrivning i flera dimensioner".

Syftet med avhandlingsarbetet är mer konkret att testa våra hypoteser om betydelsen av olika faktorer och kopplingar på ett modellsystem som omfattar alla personresor. Vi har, av skäl som redovisas nedan, varit tvungna att göra ett antal avgränsningar av olika slag. Dessa har inneburit att vi i det empiriska arbetet försökt utveckla ett trafikmodellsystem som

- avser bilinnehav, generering, områdesval och färdmedelsval
- innehåller samband mellan dessa olika val
- beaktar interaktioner inom hushållet som påverkar de olika valen
- avser alla slags resänder
- beaktar restriktioner på möjligheterna att välja olika alternativ
- kan utnyttjas för att beskriva konsekvenser av olika åtgärder för olika socio-ekonomiska kategorier
- täcker hela Stockholms län

Avhandlingsarbetet avser i första hand att vara ett empiriskt forskningsarbete. Detta innebär att vi avgränsar oss till att använda den teoribildning som är accepterad inom de discipliner som metodiken hämtats från. En ytterligare avgränsning som vi gjort är att utelämna implementeringen av modellsystemet.

3. Utgångspunkter för trafikmodellsystemet

3.1 Syftet med ett trafikmodellsystem

Vårt syfte är att beskriva kvantitativa samband mellan resmönstret och de olika faktorer som detta påverkas av. Utgångspunkten är att resmönstret utgörs av människornas beteende i trafiksystemet (med avseende på de för oss intressanta dimensionerna). Det är således människornas resbeteende som ska modelleras, för att vi ska kunna göra någon utsaga om hur resmönstret förändras om en viss åtgärd vidtas.

För att trafikmodellsystemet ska kunna behandla både monetära kostnader och andra variabler - exempelvis restider - måste det innehålla parametrar som avspeglar resmönstrets känslighet för såväl de monetära kostnaderna som de andra variablerna. Detta innebär att modellerna innehåller en viss värdering av exempelvis restid. Om detta sätt att värdera restid och andra variabler är konsistent med den samhällsekonomiska utvärderingsmetoden kan trafikmodellsystemet även utnyttjas för att framställa en del av underlaget till den samhällsekonomiska analysen.

Syftet med ett trafikmodellsystem bör därför vara att:

- ge kvantitativ information om hur resmönstret påverkas av olika åtgärder, med utgångspunkt i människornas resbeteende
- ge kvantitativ information om värderingen av olika åtgärder

3.2 Teoretiska utgångspunkter

Avhandlingsarbetet är av tillämplig natur. Detta innebär att avsikten inte främst är att utveckla nya teorier på området. Vår utgångspunkt när det gäller arbetet med att utveckla ett trafikmodellsystem utgörs därför huvudsakligen av den befintliga teoribildningen. Det egentliga forskningsarbetet avser att utveckla ett trafikmodellsystem, i vilket flera av de restriktioner på beskrivningen av resbeteendet som funnits i hittillsvarande modellsystem, och som kan förväntas ha stor betydelse, har släppts.

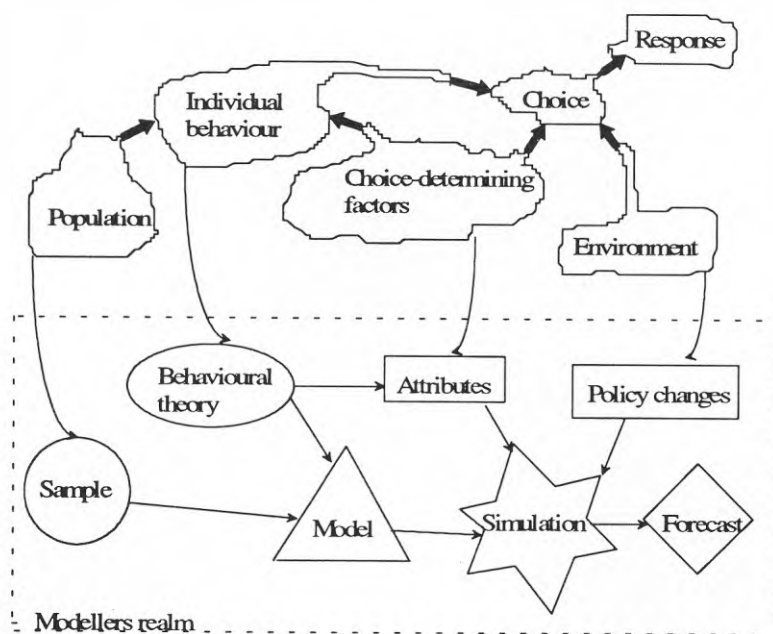
En allmän modell för individers resbeteende

Personresandet kan beskrivas som antalet förflyttningar R med ett visst färdmedel f längs en viss rutt r mellan två punkter i och j vid en viss tidpunkt t , dvs. R_{frijt} . Mängden av alla R_{frijt} kan vi kalla resmönstret. Resmönstret förutsätts vara ett resultat av individers val av handlingsalternativ under ett antal förutsättningar av olika slag.

Det är mekanismerna för detta val som är föremålet för vår analys, dels för att vi ska kunna förstå mekanismerna som sådana och dels för att vi ska kunna simulera dem med andra förutsättningar, och på så sätt kunna beräkna effekterna på resandet av olika åtgär-

der. Förhållandet mellan verkligheten och våra försök att beskriva den kan bla. representeras i följande figur, hämtad från Ortuzar och Willumsen (1990):

Figur 3.1 Verklighet och modell



Figuren är uppdelad i två parallella delar, varav den övre är "lösare i konturerna" för att illustrera att verkligheten bara delvis är känd, medan den undre delen illustrerar vårt sätt att beskriva verkligheten.

Forskningen kring resbeteendet utgår sedan länge från att det är individers val som måste beskrivas för att resbeteendet ska kunna modelleras på ett realistiskt sätt. En central fråga är naturligtvis hur mekanismerna för dessa val ska beskrivas. I de modellsystem som hittills kommit i praktiskt bruk, behandlas individernas val av olika resalternativ på ett starkt förenklat sätt, vilket mött kritik från olika håll.

Det är egentligen inte förvånande att en sådan situation uppstått. Befintliga modellsystem har i regel utvecklats av ingenjörer/ekonomer, vilka betonat kvantitativ analys, även till priset av starka förenklingar. Inom andra discipliner, främst kanske kulturgeografien, har man varit mer intresserad av hur individers handlingsmöjligheter påverkas av trafiksystemet, och därför kommit att mer betona betydelsen av olika restriktioner.

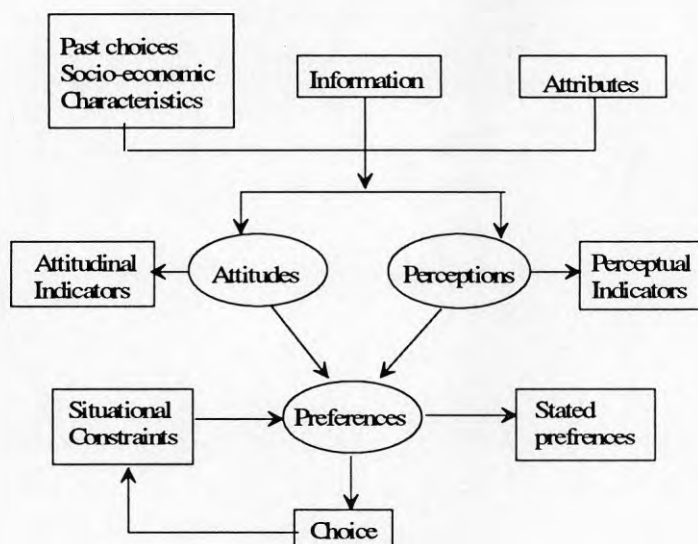
Kritiken mot hittillsvarande modellsystem avser egentligen flera aspekter. En aspekt är det faktum att själva valprocessen egentligen inte behandlas alls, utan enbart resultatet av processen, dvs. valet. En annan aspekt är att det egentligen är valet av aktiviteter som bör behandlas, snarare än att bara se till resandet vilket ju endast är ett resultat av valet av aktiviteter. En tredje aspekt är att kopplingarna mellan olika individer inte beaktas.

När det gäller att göra modeller av själva valprocessen - dvs. den process som består av

individernas utvärdering av de olika handlingsmöjligheterna, finns det idag ingen färdig metodik som kan användas i större tillämpningar. Detta gäller också modeller för val av aktiviteter. Forskning pågår dock sedan länge inom detta område. Aktivitetsbaserade ansatser kan föras tillbaka till Hägerstrand (1970). TSU-gruppen vid universitetet i Oxford har varit drivande inom detta område (se Jones et al 1980). För en state-of-the-art översikt, se Axhausen och Gärling (1992).

Vi har därför begränsat oss till att utgå från den ansats, som innebär att vi modellerar valen direkt. En sådan ansats kan beskrivas med följande figur, hämtad från Ben-Akiva och Boccara (1987):

Figur 3.2 Individens val



I denna figur bestäms individens val av restriktioner och preferenser. Dessa bestäms av attityder och perceptioner, vilka i sin tur bestäms av individens socioekonomiska förhållanden, tidigare erfarenheter, information och faktiska förhållanden. Av figuren framgår vidare att attityder, perceptioner och preferenser kan ta sig uttryck på olika sätt - i form av attitydindikatorer, perceptionsindikatorer samt uppgivna preferenser (Stated Preferences).

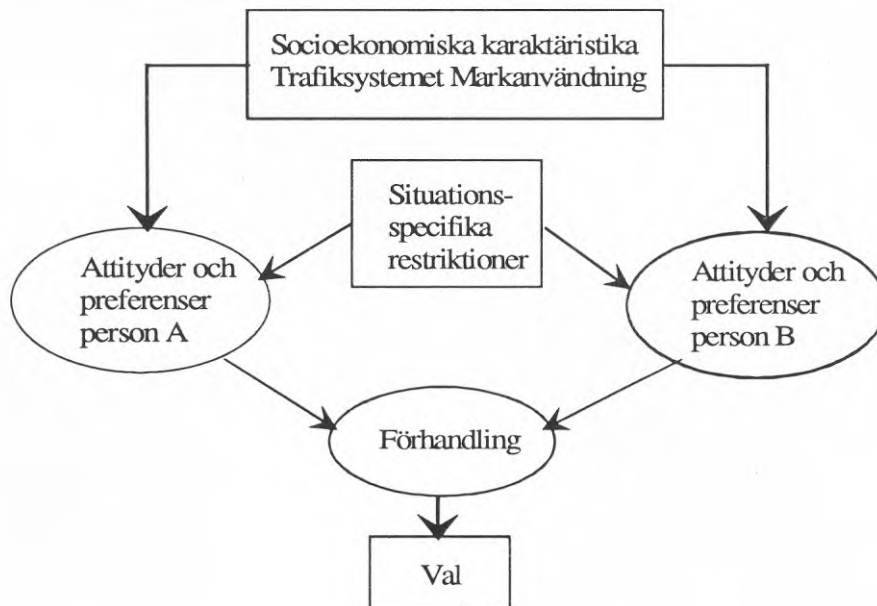
Figuren visar på såväl problem som möjligheter. Trots att vi begränsar oss till faktiska val är det naturligtvis fortfarande fråga om att beskriva en komplicerad mekanism. Samtidigt är det möjligt att med hjälp av information om de olika indikatorerna och uppgivna preferenser åstadkomma en mer realistisk modell. Metoder för detta slags analys har emellertid kommit fram efter det att utformningen av vår undersökning fastställts (se Ben-Akiva och Boccara 1987, eller Morikawa, Ben-Akiva och McFadden 1990).

Vi begränsar oss därför ytterligare till att formulera valet enbart som en funktion av olika uppmätta attribut. De olika latent (icke mätbara) variablerna som attityder och preferenser kommer då att ligga implicit i de kvantitativa samband som vi kan identifiera. Även om vi begränsar oss i detta avseende kan vi öka realismen i flera andra avseenden - dels genom

att identifiera flera valdimensioner, och dels genom att behandla valen som hushållsval, vilket gör det möjligt att beakta kopplingar inom hushållet.

I nedanstående figur redovisas den paradig som vi generellt har ansatt i detta arbete:

Figur 3.3 Ansatt paradig



Vi har således definierat trafikbeteendet som en valprocedur där flera individer interagerar, vilken resulterar i ett valbeslut avseende alla dessa individer.

En modell för diskreta val

De olika alternativ som individerna väljer mellan är i trafiksammanhang praktiskt taget alltid av diskret natur. Det är exempelvis knappast möjligt att bestämma färdväg eller destination längs någon kontinuerlig skala. Modeller för diskreta val har utvecklats och utnyttjats inom många discipliner, bla. inom matematisk psykologi (Thurston 1927).

Vår beskrivning av teorin för valmodeller i följande avsnitt följer Ben-Akiva och Lerman (1985) samt Ortuzar och Willumsen (1990). För en utförligare behandling av ämnet hänvisas i första hand till dessa referenser.

Individernas val kan ses som resultatet av en sekventiell beslutsprocess som innebär att valsituationen definieras, att olika alternativ genereras, att alternativens egenskaper utvärderas, att ett valbeslut fattas och sedan genomförs i praktiken. En teori för individernas val måste därför definiera följande olika element:

- beslutsenheten
- alternativ
- alternativens attribut
- beslutsregeln

Beslutsenheten kan vara varje enhet som fattar ett beslut, dvs. en individ, ett hushåll, ett företag, etc. I vårt fall koncentreras intresset både till individer och hushåll. I framställningen nedan skriver vi oftast *individens* val, men principerna är allmängiltiga.

Per definition förutsätter vi att varje val sker från en uppsättning ömsesidigt uteslutande alternativ. Vi kan tänka oss att det finns en universell alternativmängd som innehåller alla alternativ för det aktuella valet. Individen gör sedan sitt val från en mindre delmängd av dessa alternativ. Vi kan kalla denna för valmängden. Valmängden avgörs av vilka alternativ som är fysiskt tillgängliga för varje individ, vilka alternativ individen känner till, samt av individens restriktioner (tidsrestriktioner, budgetrestriktioner, etc).

Alternativmängden kan vara kontinuerlig eller diskret. I det kontinuerliga fallet kan det handla om hur mycket av en viss vara som ska köpas. I trafiksammanhang är valmängden istället normalt diskret, dvs. valsituationen består i att välja mellan ett begränsat antal avgränsade alternativ.

Ett alternativs attraktivitet kan beskrivas genom en vektor av attribut för alternativet. Mät-skalan kan vara *ordinal* (t.ex. snabbaste alternativet) eller *kardinal* (t.ex. alternativets restid är 30 minuter). Om alternativen är homogena (t.ex. mjölk och smör) och alternativerna bara är en vektor av mängder så reduceras vektorn av attribut enbart till mängderna. I de valsituationer som är aktuella i detta projekt är alternativen påtagligt heterogena, olika beslutsenheter har olika valmängder, de värderar olika attribut och de kan värdera samma attribut hos ett visst alternativ olika. Det är då mer naturligt att beskriva alternativen med dess olika attribut än bara med mängden.

Valet mellan alternativen i valmängden förutsätter en beslutsregel. Beslutsregeln beskriver hur beslutsenheten behandlar informationen om alternativen och kommer till ett beslut. En rad olika sådana beslutsregler kan tänkas. Ben-Akiva och Lerman (1985) klassificerar reglerna i följande fyra kategorier:

1. Dominans. Alternativet väljs om det är bättre än alla andra alternativ för minst ett attribut och inte sämre för något attribut. I praktiken kan denna beslutsregel oftast inte ange vilket alternativ som väljs eftersom alternativen ofta är bättre för något attribut men sämre för andra.
2. Tillräcklighet. Alternativet väljs om det är tillräckligt bra för alla attribut. För varje attribut finns således en viss nivå som måste uppfyllas för att alternativet ska komma ifråga. Beslutsregeln kan även innebära att alternativ elimineras om de inte är tillräckligt bra för ett visst attribut. Även för denna regel gäller att den inte med säkerhet kan ange vilket alternativ som väljs - flera alternativ kan vara tillräckligt bra.

3. Lexikografiska regler. Individens rangordnar attributen efter deras betydelse och väljer det alternativ som är bäst för det attribut som är viktigast. Om detta inte leder till att ett visst alternativ kan pekats ut (flera alternativ är lika bra för det viktigaste attributet) väger individen även in det näst viktigaste attributet, och så vidare.

4. Nyttva. I denna klass av beslutsregler förutsätts utbytbarhet mellan attributen. Detta innebär att ett alternativs attraktivitet som kännetecknas av en vektor av attribut kan reduceras till ett enda värde. Alternativets attraktivitet kan således uttryckas som en funktion av dess attribut. Beslutsregeln förutsätter att brister i ett visst attribut kan kompenseras av fördelar i andra attribut. Individen förutsätts välja det alternativ som maximerar hans nytta.

En kombination av tillräcklighetskriteriet och det lexikografiska kriteriet är känt som "elimination by aspects". Denna beslutsregel innebär att individen börjar med det viktigaste attributet och eliminerar de alternativ som inte är tillräckligt bra för detta attribut. Om mer än ett alternativ då återstår betraktas det näst viktigaste attributet på samma sätt, etc.

Det är rimligt att tro att olika individer i verkligheten kan följa olika beslutsregler och att olika regler kan väljas i olika sammanhang. Beslutsreglerna är inte heller ömsesidigt uteslutande. Om individen nyttomaximerar kommer naturligtvis dominerande alternativ att föredras. Beslutsregler som bygger på tillräcklighet kan till stor del fångas in även om nyttomaximering förutsätts, t.ex. genom att vissa alternativ inte är tillgängliga om restiden är alltför lång, eller genom icke-linjära nyttofunktioner där "dåliga" attribut ges extra hög negativ vikt. Nyttomaximeringskriteriet utesluter inte heller att beslutsprocessen i sig kan vara sekventiell för vissa beslut, t.ex. så att de sämsta alternativen sorteras bort på ett tidigt stadium. Ett sådant beteende kan antingen ses som en tillämpning av tillräcklighetskriteriet eller nyttomaximeringskriteriet.

Frågan om vilka beslutsregler som är mest realistiska har belysts i ett arbete av Payne, Bettman och Johnson (1988), där ett antal olika beslutsstrategier studerades i praktiska experiment. Resultatet visar dels att valet av beslutsstrategier beror av vilken tidspress beslutssituationen innehåller, och dels att försökspersonerna i situationer utan tidspress tenderade att tillämpa vad som kallas en vägd additiv kompensatorisk strategi. Denna innebär att man för alla alternativ gör en sammanvägning av attributen för respektive alternativ och väljer det som får högst värde, vilket motsvarar ett nyttomaximeringsbeteende.

Young och Brown (1983) utvecklade både modeller som bygger på "elimination by aspects" och logitmodeller som bygger på nyttomaximering för färdmedelsvalsanalyser på ett datamaterial från Melbourne. Båda ansatserna resulterade i modeller med samma variabelinnehåll. De statistiska skillnaderna var små, men med marginellt bättre resultat för logitmodellen. Även skillnaden i förmåga att förutsäga valet i olika delar av det använda datamaterialet var liten - också här gick de små skillnaderna i logitmodellens favör. Modellerna innebar dock väsentliga skillnader i känslighet för ändrad restid - logitmodellen var betydligt känsligare och dess parametrar innebar ett klart högre tidsvärde. Det var dock ej möjligt att avgöra vilket resultat som var mest korrekt.

Vi har valt att basera vårt arbete på en nyttomaximerande valprocess, vilket stöds eller åtminstone inte motsägs av de erfarenheter som redovisats ovan. Det är rimligt att anta att de flesta resbeslut fattas utan större tidspress, även om undantag naturligtvis förekommer. Nyttomaximeringsteorin innebär också att valprocessen beskrivs på ett sätt som lämpar sig för matematisk analys och statistiska tillämpningar. De samhällsekonomiska analysmetoderna, för vilka trafikmodellerna är en viktig datakälla, baseras också på nyttomaximeringsteori, vilket utvecklas vidare nedan.

Modellerna vi utvecklar förutsätter att individerna är rationella. Med detta avses att individerna gör medvetna och överlagda val utifrån sina egna värderingar av alternativen. Mer precist förutsätter vi att valen är *konsistenta* och *transitiva*. Kravet på konsistens innebär att vi förutsätter att individen kommer att göra samma val om han två gånger ställs inför exakt samma valsituation och förutsättningarna även i övrigt är lika. Kravet på transitivitet innebär att om alternativ A föredras framför alternativ B och alternativ B föredras framför alternativ C så föredras alternativ A även framför alternativ C.

Konsistens med samhällsekonomisk analys

Ett syfte med modellsystemet är att ge underlag för en samhällsekonomisk utvärdering. I detta sammanhang spelar antaganden om tidsvärderingen en stor roll. Trafikmodellerna innehåller implicita tidsvärden genom att modellerna innehåller parametrar för tids- och kostnadsvariabler. Det är därför intressant att diskutera under vilka förutsättningar som sådana tidsvärden är konsistenta med den ekonomiska analysapparaten som normalt används vid samhällsekonomiska analyser.

Effekterna av en trafikinvestering kan när det gäller personresor beröra såväl konsumenter (privata resor) som företag (tjänsteresor). Dessa kategorier kan också beröras samtidigt eftersom tjänsteresor kan ske såväl inom som utom normal arbetstid. I detta avsnitt ska de teoretiska förutsättningarna diskuteras för såväl privatresor som tjänsteresor.

Privata resor

För att beräkningar av restidsvinsternas betydelse ska vara av värde som en del i en ekonomisk analys krävs att de kan föras tillbaka på ekonomisk teori. Detta kräver bl.a. en explicit behandling av tidsbegreppet. Försök att införa tidsbegreppet explicit i den ekonomiska analysapparaten har gjorts sedan länge. Enligt Bruzelius (1978) tycks de första försöken ha gjorts redan på slutet av 1800-talet. Bruzelius har i sin doktorsavhandling 1978 formulerat en teori (baserad på en modell av DeSerpa) för hur tidsbegreppet bör föras in i den välfärdsteoretiska analysapparaten. Framställningen nedan är hämtad ur detta avhandlingsarbete.

I traditionell neoklassisk mikroteori antas individen maximera en nyttofunktion, vars argument utgörs av varor och tjänster av olika slag, under bivillkoret att summan av utgifterna inte får överskrida inkomsten. Detta problem kan formuleras på följande sätt:

$$\text{Max } U(x_1, \dots, x_n, t_w)$$

$$\text{u.b. } \Sigma p_i x_i - w t_w - I \leq 0$$

där U är nyttofunktionen, x_i ($i=1, \dots, n$) är kvantiteterna för olika varor, t_w är mängden arbetade tidsenheter, p_i är priset på respektive vara, w är lönen per tidsenhet och I är inkomst utöver arbetsinkomst.

Eftersom tid inte ingår explicit i denna formulering ges ingen vägledning för hur tid ska behandlas i ekonomisk analys. För att råda bot på detta problem måste således begreppet tid föras in explicit i analysen.

Detta kan göras genom att såväl ren fritid som tiden för konsumtion av olika varor och tjänster förs in i nyttofunktionen, samt att tidsåtgången i konsumtionen och den totala tillgängliga tiden förs in som restriktioner:

$$\text{Max } U(x_1, \dots, x_n, l, t_1, \dots, t_n, t_w)$$

$$\text{u.b. } \Sigma p_i x_i - w t_w - I \leq 0$$

$$l + \Sigma t_i + t_w - T = 0$$

$$q_i x_i - t_i \leq 0 \quad i = 1, \dots, n_m$$

$$q_i x_i - t_i = 0 \quad i = n_m + 1, \dots, n$$

där t_i ($i=1, \dots, n$) är den totala tidsåtgången för olika varor, l är fritid, q_i är den *minsta* tidsåtgång som per enhet krävs för en viss vara, och T är den totala tillgängliga tiden för perioden i fråga. I nyttofunktionen ingår nu också tidsåtgången för respektive vara, vilket också innebär att en enhets tidsförändring kan påverka nyttofunktionen olika mycket för olika varor.

Det finns två bivillkor som avser tidsåtgången. Det första (som avser de n_m första varorna) innebär att restriktionen inte alltid är bindande, vilket innebär att konsumenten kan låta tidsåtgången vara större än vad som egentligen krävs, medan den andra alltid är bindande. I det förra fallet är restriktionerna alltid exogent givna (t.ex. i form av hastighetsgränser), medan de i det andra fallet kan vara såväl exogena (en kollektivtrafiktidtabell) som endogena (man har fritt valt en hastighet).

Denna modell kan formuleras om på ett sådant sätt att den från matematiska utgångspunkter inte skiljer sig från den traditionella nyttomaximeringsansatsen. Det innebär att de grundläggande slutsatserna från denna teori inte påverkas av en utvidgning med tidsbegreppet.

Med utgångspunkt från denna modell blir det således möjligt att analysera hur förändringar i tidsåtgången för olika varor eller tjänster - t.ex. en arbetsresa - påverkar nyttonivån för en given individ. Denna analys kan ske i tidsvärdes- och konsumentöverskottstermer.

Tidsvärdesbegreppet

Vi kan definiera begreppet "marginellt tidsvärde" som det som en konsument är villig att betala för en marginell restidsminskning vid en viss inkomstnivå. Bruzelius (1978) uttrycker detta i termer av modellen ovan genom att först formulera om den som en Lagrangefunktion:

$$L = U(x_1, \dots, x_n, t_1, \dots, t_n, t_w) - \lambda (\sum p_i x_i - w t_w - I) - \mu (L + \sum t_i + t_w - T) \\ - \sum c_i (q_i x_i - t_i) - \sum c_i (q_i x_i - t_i)$$

där $i = 1, n$ resp $i = n+1, m$ i den undre raden (och avser de två olika typerna av bivillkor för tidsåtgången)

Genom att därefter utnyttja villkoren för optimum kan ett uttryck för den marginella substitutionskvoten för restid i förhållande till pengar ställas upp. Lagrangemultiplikatorerna för en restidsrestriktion samt för budgetrestriktionen kan då utnyttjas, eftersom de representerar den marginella nyttan av restid respektive pengar:

$$\frac{c_i}{\lambda} = \frac{\mu}{\lambda} - \frac{\delta U}{\delta t_i}$$

Det marginella värdet av en minskad restid är således lika med värdet av mer tid totalt ("tiden som resurs") minus värdet av tiden i den aktuella aktiviteten (resan). Det marginella tidsvärdet beror således av alla modellparametrar - såväl inkomst som lön, priser och tidsåtgång för konsumtion av olika varor.

Ett konstant tidsvärde

Som antytts ovan, är det grundläggande syftet med en ekonomisk analys av restidseffekterna i samband med en transportinvestering att beräkna effekterna med avseende på förändringen i konsumentöverskottet. I praktiken tillämpas oftast ett konstant tidsvärde i samband med konsumentöverskottsberäkningar. För att förutsättningen om ett konstant tidsvärde ska resultera i korrekta beräkningar av konsumentöverskottet krävs emellertid starka antaganden om konsumenternas beteende.

Det kan visas (Bruzelius 1978) att nyttofunktionen i detta fall måste vara linjärt additiv och att efterfrågan måste uttryckas som en funktion av endast den s.k. generaliserade kostnaden (dvs. summan av pris och den med tidsvärdet multiplicerade restiden). En viktig konsekvens härav är att tidsvärdet förutsätts oberoende av inkomsten. Nyttofunktionerna innehåller i praktiska tillämpningar även andra slags variabler än restid och reskostnad, som exempelvis kön och bilkonkurrens. Dessa bör snarast ses som förenklade segmenteringsvariabler och proxyvariabler för alternativtillgång.

Antagandet att tidsvärdet (eller snarare den marginella nyttan av pengar) är oberoende av inkomsten har på senare tid ifrågasatts av bla. Jara-Diaz och Videla (1989), och Hau (1987) och också varit föremål för analys med svenska data från RVU 86/87 (Ramjerdi 1990). Problemet kan endast delvis lösas genom att dela upp populationen i inkomstsegment - antagandet om oberoende av pengar kommer ändå att gälla inom varje segment.

Jara-Diaz utgår i sin analys från en Lancaster-baserad modell (där tiden är ett av en varas många attribut, och där tidsbudgetrestriktionen saknas). Något försök att introducera inkomstberoende utifrån den mer explicita tidsallokeringsmodellen förefaller inte ha gjorts. Konsekvenserna av att beakta inkomstberoendet enligt Jara-Diaz metod blir mer komplicerade modellspecifikationer. Ramjerdis (1990) test på data från resvaneundersökningen i Stockholm 1986/87 visade inte på något samband mellan den marginella nyttan av pengar och inkomst.

Vi har genomgående valt att inte fördjupa oss i sambandet mellan inkomst och tidsvärde. Detta innebär att de tidsvärden som kan härledas är genomsnittliga värden oavsett inkomsten. Detta innebär inte att man vid tillämpning av modellsystemet måste förutsätta att tidsvärdena är konstanta, men däremot att kunskaperna om hur de utvecklas måste sökas på annat håll - exempelvis genom speciella tidsvärdesanalyser. Man bör hålla i minnet att en prognos över framtida tidsvärden också måste baseras på prognoser av annat än själva inkomstutvecklingen (som exempelvis utvecklingen av arbetstidens längd och fördelning - tidsbudgetrestriktionen påverkar ju också tidsvärdet).

Hushållsbaserad nyttomaximering

Den teoribildning som redovisats ovan utgår från *individuell* nyttomaximering. Som framgått ovan utgår vi ofta från *hushållet* som beslutande enhet, vilket implicit innebär att vi maximerar någon slags funktion av individuella nyttor - bara individer kan uppleva nytta, och hushållet är ju - i detta sammanhang - endast en organisationsform. Frågan är då om detta angreppssätt är konsistent med den samhällsekonomiska analysapparaten, vilken - i likhet med den överväldigande delen av ekonomisk teori - utgår från individuell nyttomaximering.

Med utgångspunkt från den ovan redovisade individuella nyttomaximeringsansatsen kan olika tidsvärden skattas. De skattade tidsvärdena kan sedan tillämpas för att utvärdera exempelvis olika investeringsalternativ. Vid denna tillämpning rangordnas olika investeringar efter den totala nyttan, summerad över alla individer. För att en investering ska vara samhällsekonomiskt lönsam, krävs att åtminstone någon individ får det bättre, medan ingen individ får det sämre (Paretokriteriet). Detta innebär att de som i praktiken skulle få det sämre måste kompenseras. Efter kompensation måste nyttan överstiga kostnaderna. (I praktiken genomförs praktiskt taget aldrig kompensationer projektvis, bla. med motiveringen att kompensation kan ske på sikt och i andra projekt eller genom andra åtgärder.)

Denna procedur innebär att man inte behöver göra några *normativa* interpersonella nyttojämförelser. För att kunna göra sådana krävs väldefinierade samhälleliga nyttofunktioner (se exempelvis Malinvaud 1972), vilka vanligen saknas.

Om det emellertid i verkligheten är så att flera individer är inblandade i ett val av handlingsalternativ (som exempelvis frågan om vem som ska ta bilen till arbetet eller vem som ska utföra ett inköp) kommer en modell som enbart baseras på en av dessa individer att kunna innehålla en bias till följd av den ofullständiga behandlingen. Det krävs då att den individuella nyttomaximeringsansatsen utvidgas till att beakta dels de restriktioner som beroendet mellan de olika hushållsmedlemmarna ger, och dels att valet sker med hänsyn till flera individers nyttor och preferenser.

Beroendet mellan olika individer kan behandlas konsistent genom att låta modellen avse val mellan alla de alternativ som är inbördes beroende. Exempel på detta är valet av färd-sätt till arbetet, där alternativen innebär att samma bil endast kan användas av en person, om de inte samåker.

Beaktandet av att valet berör flera personers nyttonivå kan ske genom att nyttorna för de enskilda personerna ingår som argument i avbildningen av den beslutsprocess inom hushållet vilken leder fram till valet av ett visst alternativ. Exempel på detta är valet av vem som ska göra en inköpsresa, där alternativen beskrivs i form av resuppofteringen för respektive hushållsmedlem.

Att slutligen beakta att valet av alternativ är beroende av preferenserna inte bara hos den individ som berörs av ett visst alternativ, utan hos alla individer i hushållet, kan ske genom att formulera valet just som ett hushållsval. Den målfunktion som vi då kan estimeras, och som är den som bäst beskriver hushållets beteende, är ett uttryck för den process inom hushållet som leder fram till de observerade valen. En målfunktion av detta slag skulle också kunna kallas för hushållets välfärdsfunktion, eller en samhällelig nyttofunktion i miniatyr (se exempelvis de V. Graaf 1967, Cigno 1991).

Det kan diskuteras om denna målfunktion är uttryck för en enda persons preferenser vilka påtvingats övriga hushållsmedlemmar, om den är uttryck för identiska preferenser hos alla hushållsmedlemmar eller om den är uttryck för ett förhandlingsspel inom hushållet mellan hushållsmedlemmar med olika preferenser och förhandlingsstyrka. De två första alternativen kan ses som specialfall av det sista alternativet, vilket är mest generellt.

Hushållsbaserade modeller förutsätter således förekomsten av interpersonella nyttojäm-förelser, vilka är en uppenbar del i förhandlingprocessen inom hushållet. Här ligger ett potentiellt problem - är tidsvärden som skattats med utgångspunkt från hushållsbaserade valmodeller är konsistenta med den ovan diskuterade tillämpningen av tidsvärdena?

Det man vill undvika i den ovan beskrivna utvärderingsprocessen är *normativa* interperso-nella nyttojäm-förelser. Genom att skatta hushållsmodeller för vi dock inte in något norma-tivt inslag i utvärderingsprocessen - givet att valen i verkligheten är hushållsval kommer vi i stället att undvika den felspecificering vi gör om vi skattar individuella tidsvärden utifrån felaktiga premisser när det gäller individens beslutssituation. Den information om interper-sonella nyttojäm-förelser som vi utnyttjar när vi tillämpar hushållsmodellerna är inte nor-mativ, utan empiriskt underbyggd eftersom vi *skattar* formen på de interpersonella nytto-jäm-förelser som görs inom hushållet.

Det som sker i den förhandlingsprocess som vi antytt ovan - om vem som ska få ta bilen, eller om vem som ska utföra ett inköp - är att flera individer når fram till en överenskom-melse, som (förutom faktiska omständigheter) baseras dels på individernas egna preferen-ser (inklusive hänsynstagande till andra personer), och dels på deras förhandlingsstyrka.

Denna överenskommelse är därmed den för alla parter bästa *möjliga* lösningen i en given situation. Den representerar därför den för varje individ maximala nyttan, givet den restriktion som beroendet mellan olika hushållsmedlemmar ger. Hushållets målfunktion pekar ut vilket alternativ individerna i hushållet kommer överens om, och innebär samti-digt att ingen individ kan förbättra sin situation genom att efterfråga ett annat alternativ - utom genom att säga upp det "kontrakt" som hushållet innebär.

Synsättet innebär alltså att vi ansätter en beslutsfunktion, som är konsistent dels med individuell nyttomaximering, dels med de restriktioner - och möjligheter - som hushållet innebär. Det skattade tidsvärdet blir därmed konsistent med de avvägningar som leder till att välfärden för hushållets medlemmar - vägd på det sätt som hushålls"kontraktet" innebär - maximeras.

När det sedan gäller att beräkna nyttoeffekter kan nyttan beräknas för respektive individ med utgångspunkt från hushållets målfunktion och summeras på vanligt sätt. Vi lämnar alltså egentligen inte paradigmen om individuell nyttomaximering, utan försöker i stället göra den mer realistisk genom att föra in några av de restriktioner som den grymma verkligheten faktiskt innehåller.

Någon mer formaliserad teori avseende tidsvärdering och nyttomaximering med flera individer har, såvitt vi vet, inte utvecklats. Inom disciplinen "economics of the family" har dock problemet med hushållbaserad nyttomaximering behandlats (se exempelvis Cigno 1991), vilket skulle kunna utgöra en utgångspunkt för en mer formaliserad analys. Vi får här emellertid begränsa oss till att utgå från det mer intuitiva resonemang som här förts.

Tjänsteresor

Tjänsteresor innebär som antytts ovan dels kostnader för företaget - en tjänsteresa medför (förutom de monetära kostnaderna) kostnader i form av att arbetstid ersätts av restid (som kan vara mer eller mindre produktiv), och dels kostnader för individen genom att fritid ersätts av restid. En restidsinbesparing kan således minska dessa kostnader.

När det gäller att skatta en modell för tjänsteresor, är frågan om vems nyttofunktion det är som ska gälla ett uppenbart problem. När det gäller tjänsteresor inom Stockholm län, vilka normalt utförs under dagen, kan det vara rimligt att anta att arbetsgivarens preferenser är styrande under arbetstid, medan arbetstagarens preferenser är styrande under övrig tid.

Hensher (1989) har definierat värdet av en restidsinbesparing för tjänsteresor på följande sätt:

$$TV = (1 - r - pq) * MP + (1-r) * vw + r * vl + MPF$$

där

- TV = värdet av inbesparad tid för tjänsteresor
- r = andelen av inbesparad restid som används till fritid
- p = andel inbesparad produktiv restid
- q = relativ produktivitet för restid som används till arbete
- MP = marginell produktivitet för arbetskraften
- vw = det monetära värdet för den anställde av restid jämfört med att vara på arbetet
- vl = det monetära värdet för den anställde av restid jämfört med fritid
- MPF = värdet av ökad produktivitet genom minskad trötthet

Modellen innehåller således en del som utgörs av företagets kostnadsinbesparing, och en del som utgörs av individens nyttoökning.

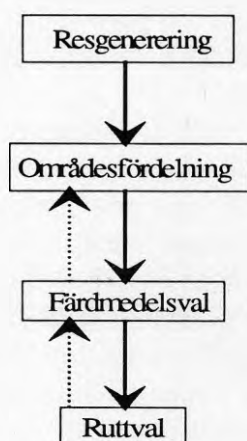
När vi skattar en modell för tjänsteresor, kommer parametrarna ovan att avspeglas i modellparametrarna. Det är således önskvärt att så långt möjligt hålla isär företagets och individens värderingar.

Våra modellskattningar bygger på företagets kostnader för resorna, och skiljer inte på om resan äger rum under arbetstid eller fritid. De flesta arbetsresor är arbetsplatsbaserade, och restiden avser därför restid med produktiv alternativ användning. Detta bör också i stor utsträckning gälla tjänsteresor som utförs på väg mellan hemmet och arbetsplatsen eller tvärtom. Bostadsbaserade tjänsteresor är en minoritet. Det är därför knappast någon grov förenkling att för arbetsresor inom länet inte särskilja restid under fritid. De resulterande restidsvärdena kommer därför att i huvudsak avspegla arbetsgivarens värdering av restid.

3.3 Hittillsvarande trafikmodellssystem

En vanlig utgångspunkt när det gäller trafikmodeller för persontrafik är den klassiska fyrstegsmodellen, vilken kan beskrivas på följande sätt:

Figur 3.4 Fyrstegsmodellen



Delmodellerna beskriver, så när som på tidpunktsvalet, de resbeslut som bestämmer resmönstret. Modeller som beskriver dessa val har utvecklats sedan lång tid tillbaka med olika tekniker.

Estimeringsteknikens utveckling mot att tillämpa s.k. logitanalys på disaggregerade data (dvs. data för enskilda individer eller hushåll) har givit möjligheter till att införa ett beroende mellan de olika delmodellerna. Samtidigt har detta inneburit att ruttvalet kommit att behandlas separat, eftersom logitanalysens krav på oberoende mellan olika alternativ är svåra att uppfylla för ruttvalsproblemet. En bidragande orsak är att vissa ruttvalsalgoritmer är deterministiskt orienterade.

Svenska modellsystem

De möjligheter som finns för att inom Sverige analysera dessa val inom ramen för ett sammanhängande modellsystem på tätortsnivå utgörs av ett antal modeller, som har sitt ursprung på olika orter, men som genom ett inskallningsförfarande (se Algers, Colliander och Widlert 1987) anpassats till orter i Sverige (Widlert 1989, Nylander 1992) och Norge (Nylander och Widlert 1991). Modellsystemet har också använts i Västerås och Uppsala.

När det gäller regionala resor har samma modeller genom ett enklare förfarande anpassats till resande på regional nivå för resor mellan tätorter i hela Sverige. Slutligen har ett modellsystem estimerats för långväga resor (Algers 1991). Samtliga dessa modellsystem avser generering, destinationsval samt färdmedelsval. Modellen för inomregionala resor innehåller dock inget beroende mellan generering och övriga val.

En viktig faktor i ett trafikmodellsystem utgörs också av bilinnehavet. Detta har hittills modellerats utan samband med trafikmodellsystemet i övrigt, med något undantag (Algers 1974).

Utländska modellsystem

Det första mer sammanhängande trafikmodellsystemet som baseras på disaggregerade logitmodeller torde vara det modellsystem som 1974-1977 utvecklades för Metropolitan Transportation Commission i San Francisco (CSI 1980). Detta modellsystem byggde på en undersökning från 1965 omfattande 30 000 hushåll, och innehöll 21 submodeller för 4 res-ärenden. I samband med detta projekt genomfördes en omfattande genomgång av skattningmetoder och en validering av systemet (McFadden och Talvitie 1977). Modellsystemet ansågs i slutet av 70-talet som föråldrat, och man beslöt då att uppdatera det (Kollo och Purvis 1988). Denna uppdatering innebar också genomförandet av en "small sample survey" 1981, omfattande 6 200 hushåll (dvs. ungefär dubbelt så stor som vår undersökning). Uppdateringen innebar flera förenklingar i modellstrukturen bla. beroende på att undersökningen inte innehöll tillräckligt många kollektivresor.

Hollands nationella modell, som är det mest avancerade trafikmodellsystem som finns idag, består av individbaserade logitmodeller för generering, områdesval, färdattsval och tidpunktsval (Daly och Gunn 1985). Ruttvalet hanteras genom deterministiska assignment-algoritmer.

Andra exempel på större modellsystem utgörs av Zeudvleugel-systemet i Holland (Daly och van Zwam 1983).

3.4 Brister i hittillsvarande modellsystem

Allmänna brister

De hittills inom eller utom Sverige utvecklade trafikmodellsystemen saknar många av de kopplingar som man kan tänka sig förekommer i verkligheten. Vår hypotes är att flera av dessa kopplingar är av betydelse för resbeteendet. Avsaknaden av dessa kopplingar utgör därför potentiella brister i ett trafikmodellsystem. Följande brister har vi ansett mest angelägna att åtgärda:

- Hittillsvarande modeller har avsett individer, utan någon koppling till de restriktioner och möjligheter som interaktioner med andra hushållsmedlemmar ger.
- Bilinnehavet har behandlats avskilt från behandlingen av resmönstret.
- Restriktioner utanför hushållet har inte beaktats, vilket innebär att man inte rensat bort de alternativ som inte kan realiseras inom ramen för hushållets och individernas tidsbudget.
- Möjligheten att genomföra reskedjor har endast beaktats utan återkoppling till övriga resbeslut.

Dessa brister förekommer i varierande grad i olika modellsystem. Inget modellsystem är fritt från alla brister.

Beskrivningen av resbeteendet kan naturligtvis göras mer realistisk i en rad olika avseenden. Ben-Akiva och Lerman (1985) delar in utvecklingsmöjligheterna vid beskrivningen av resbeteendet i 4 olika områden:

- teorin för resbeteende
- datainsamlingen
- modellstrukturen
- estimeringen

Inom varje område anges ett antal olika aspekter, bland annat interaktion mellan hushållsmedlemmar och bestämning av valmängden. Dessa behandlas också i vår ansats.

Brister i trafikmodellsystem för Stockholms län

När det gäller tillämpningar i Stockholms län, har bristerna i trafikmodellsystemen varit mycket stora, även jämfört med existerande modellsystem. Den enda modell som fanns i bruk vid avhandlingsarbetets början var en färdmedelsvalsmodell för arbetsresor, skattad på data för en del av Stockholms län. Vissa försök hade också gjorts att skatta modeller på aggregerade data (Hansen 1980, Algers 1982). Jämfört med ett modellsystem som är hel-

täckande i den meningen att alla slags resor och alla valdimensioner behandlas fanns det således stora brister.

Det är således önskvärt att ett nytt modellsystem för Stockholms län:

- omfattar alla valdimensioner
- omfattar alla slags resärenden
- avser hela Stockholms län
- tillåter uppdelning på olika socioekonomiska kategorier

3.5 En allmän specifikation av ett trafikmodellsystem

Mot bakgrund av den ovan förda diskussionen ansätter vi således ett modellsystem för diskreta val, baserat på nyttomaximering. Från denna utgångspunkt kan en allmän specifikation göras av trafikmodellsystemet. Specifikationen återspeglar också de avgränsningar som vi valt att göra. Mer detaljerade specifikationer görs i samband med det empiriska arbetet, vilket redovisas i slutrapportens andra och tredje delar.

Ett hushållsbaserat trafikmodellsystem

Bilnehavet är ett långsiktigt val med stor betydelse för resbeteendet. Vår hypotes är att bilnehavet såväl beror av hur det påverkar tillgängligheten till olika målpunkter för flera av hushållsmedlemmarna, som av hela hushållets ekonomi, varför vi formulerar det som ett hushållsval. I detta avseende är således allt resande hushållsbaserat.

Även för att kunna hantera interaktioner inom hushållet måste modellsystemet i princip vara hushållsbaserat. Det innebär inte att alla modeller är hushållsbaserade - när det gäller exempelvis skolresor och tjänsteresor dominerar interaktion mellan individer i hushållet och exogena aktörer, varför dessa resor behandlas som individbaserade resor (dock med det hushållsbaserade bilnehavet som utgångspunkt).

Interaktioner inom hushållet

De olika individerna fattar sina resbeslut inom ramen för sitt hushåll. Detta innebär dels att det finns ett beroende mellan de olika individernas val, och dels att det i viss utsträckning finns en substituerbarhet mellan olika individer när det gäller att genomföra en resa.

Färdmedelsvalet utgör ett exempel på ett viktigt beroende mellan individer. Om det finns en bil i hushållet men två personer med körkort som båda samtidigt vill göra en resa till olika målpunkter, så kan bara en person ta bilen. Denna typ av beroende har tidigare normalt hanterats genom att reducera benägenheten att välja bilalternativet med hjälp av en variabel som beskriver graden av konkurrens om bilen, exempelvis som antal bilar per körkort.

Förutom att vara en bristfällig beskrivning av valprocessen kan detta sätt att behandla beroendet också innebära risk för inkonsistens, genom att det inte finns någon garanti för

att summan av dem som väljer bil inte överstiger antalet bilar. I stället bör beroendet behandlas explicit, vilket innebär att färdmedelsvalet beskrivs som ett hushållsval, där båda personernas färdmedelsval behandlas simultant. Detta sätt att behandla beroendet innebär också att man explicit behandlar rollfördelningen mellan män och kvinnor. Såvitt vi vet har detta tidigare behandlats empiriskt enbart av Algers och Widlert (1986b och 1987a).

När det gäller substituerbarheten mellan olika hushållsmedlemmars resande har detta såvitt vi vet inte heller behandlats empiriskt tidigare. Det främsta exemplet på substituerbarhet i detta avseende torde vara de resor som genomförs för att utföra hushållets varutransporter, främst inköpsresorna.

Det är rimligt att tänka sig att det finns flera olika hushållsmedlemmar som var och en skulle kunna utföra ett visst inköp. Frågan om vem som i praktiken kommer att utföra inköpet kan vara ett resultat såväl av personens roll i hushållet som av personens möjligheter att utföra inköpet, i form av tillgänglighet till olika inköpsställen. Denna tillgänglighet kan bero på biltillgång och körkortsinnehav, men också på var personen arbetar och de möjligheter till inköp på vägen till och från arbetet som detta ger.

Valdimensioner

För att kunna beskriva resgenerering, områdesfördelning och färdmedelsval måste givetvis modellsystemet omfatta dessa valdimensioner. Bilinnehavet spelar naturligtvis en central roll, och inkluderas därför som en särskild valdimension. För att kunna beakta resandets fördelning på olika restyper (hushållsbaserade resor, arbetsplatsbaserade resor och reskedjor) införs också en valdimension som beskriver individens val i detta avseende.

Interaktionerna mellan olika hushållsmedlemmar kan också beskrivas som valprocesser. Detta gäller dels de situationer där hushållet väljer vilken kombination av hushållets medlemmar som ska genomföra en given resa, och dels de situationer där biltillgången för arbetsresan fördelas. Vi definierar således en valdimension för individvalet och en valdimension för bilallokeringen.

Modellsystemet kommer därför att omfatta följande val:

Övergripande val:

- bilinnehav

Val som avser resbeslut:

- resfrekvens
- områdesval
- färdmedelsval
- val av reskedja

Val som avser interaktioner inom hushållet:

- individval
- val av vem som använder bilen

Interaktion mellan olika resbeslut

Även om resandet beror av olika resbeslut kan dessa vara inbördes beroende. Om exempelvis turtätheten ökar från ett område till en viss målpunkt ökar benägenheten att åka kollektivt för resor i den relationen. Samtidigt är det rimligt att tänka sig, att antalet personer som reser i den relationen ökar, liksom antalet resor totalt.

En viktig kritik som riktats mot tidigare trafikmodeller har varit att de inte beaktar beroendet mellan de olika trafikbesluten. Speciellt när det gäller ärenden med en hög grad av substituerbarhet på kort sikt - exempelvis inköpsresor - kan dessa beroenden ha mycket stor betydelse.

Restriktioner

Förutom de restriktioner som hushållets egna förutsättningar implicerar spelar också andra restriktioner en viss roll. Ett exempel på sådana restriktioner är de olika hushållsmedlemmarnas arbetstider och affärernas öppettider. Dessa restriktioner innebär att olika alternativa färdvägar eller destinationer kan falla bort, eftersom det helt enkelt inte är möjligt att nå en viss målpunkt med ett visst färdmedel inom rimlig tid.

Modellernas alternativmängd baseras därför på möjligheterna att nå olika destinationer, när detta är tillämpligt. Betydelsen av att beakta restriktioner av detta slag har bl.a. påpekats av kulturgeografer, vilka baserat sina iakttagelser på tidsbudgetstudier (Hägerstrand et al 1974).

Oberoende variabler

Vi har således specificerat ett modellsystem, där de olika valdimensionerna definierats. Alternativmängden i dessa valdimensioner definierar utfallsrummet för den beroende variabeln, dvs. det faktiska valet. Ett viktigt led i analysen är naturligtvis att avgöra vilka oberoende variabler som förklarar dessa val, samt att avgöra styrkan i dessa samband. Denna analys genomförs som en del i det empiriska arbetet, vilket innebär att mer detaljerade hypoteser formuleras i samband med analysen av respektive modell.

En mer övergripande redovisning av olika förklaringsvariabler kan dock göras här. Dessa faktorer kan delas in i följande grupper:

- faktorer som påverkar de uppostringar som är förknippade med resan, som t.ex. reskostnader, restider och bekvämlighet
- faktorer som är förknippade med hushållet och hushållsmedlemmarna, som t.ex. kön, ålder, inkomst etc
- de faktorer som är förknippade med restriktioner för resandet, som t.ex. arbetstider och öppettider i affärer
- de faktorer som är förknippade med nyttan av att genomföra resan, som t.ex. förekomsten av affärer eller arbetsplatser i olika områden

Ärendeuppdelning

De olika valbesluten kan påverkas på olika sätt beroende på resärende. Förklaringen kan vara

- olika tidsrestriktioner, vilket innebär olika tidsvärden (allt annat lika)
- olika resförhållanden, vilket också innebär olika tidsvärden (allt annat lika)
- olika fördelning på hushållsmedlemmar
- olika attraktivitetsmått
- olika hög grad av substituerbarhet mellan målpunkter, vilket innebär olika medelreslängd

I vilken grad man behöver särskilja de olika resärendena är till slut en empirisk fråga.

Numerisk metod

För att kunna formulera en kvantitativ modell krävs någon form av numerisk metod, som är relevant för diskreta beroende variabler. För att numeriskt bestämma modellparametrarna i modellerna krävs också att vi ansätter en viss modellstruktur med en viss uppsättning förklaringsvariabler formulerade på ett visst sätt.

Detta innebär samtidigt att vi formulerar hypoteser om de faktorer som påverkar resbeteendet, liksom om sambanden mellan de olika valbesluten. Dessa hypoteser formuleras i form av hypoteser avseende skattade parametrar. För att hypoteserna ska kunna testas statistiskt, måste den numeriska metoden vara av sådan beskaffenhet att den anger den statistiska osäkerheten i skattningen.

Ej behandlade aspekter

Eftersom resbeteendet är ett mycket komplext fenomen kommer varje modell av resbeteendet att innehålla en rad mer eller mindre betydelsefulla begränsningar. Vi har i vårt arbete försökt angripa vissa av dessa begränsningar. Av olika skäl behandlar vi emellertid inte i vårt modellsystem vissa viktigare aspekter som egentligen krävs för att modellen ska ge en fullständig beskrivning av resmönstret.

Tidpunktsval och ruttval

Resmönstret definierades i 3.2 som antalet resor i en viss områdesrelation med ett visst färdstätt längs en viss rutt vid en viss tidpunkt. I vårt modellsystem behandlas inte val av tidpunkt eller rutt. Såväl tidpunktsvalet som ruttvalet är intressant eftersom de avgör kapacitetsutnyttjandet i trafiksystemet. Av speciellt intresse är därför tidpunktsvalet mer i detalj vid morgon- och eftermiddagsrusningen.

Vi ansåg det emellertid inte möjligt att ta fram ett dataunderlag, som skulle beskriva skillnaderna i trafiksystemet för ett antal korta näraliggande tidsperioder. Därför formulerades modellsystemet ursprungligen med en tidpunktsvalsmodell endast för vissa resärenden. Detta tidpunktsval avsåg då en grov tidsperiodindelning. Det visade sig emellertid mycket resurskrävande att inkludera denna dimension i modellerna. Eftersom modellsystemet ändå

inte skulle innehålla någon tidpunktsvalsmodell för alla resärenden, lät vi denna del utgå, trots att empiriska tester tydde på att sådana modeller skulle kunna estimeras.

Det finns flera skäl till att ruttvalet inte ingår. När det gäller bilresor, är det i och för sig tänkbart att formulera ruttvalet som ett val mellan ett antal olika rutter, vilket har gjorts i olika studier (Antonisse, Daly och Ben-Akiva 1989). För att meningsfullt kunna tillämpa en sådan modell krävs emellertid, att man också beaktar interaktioner mellan olika resenärer, dvs. trängseffekter. Detta innebär, att man måste ha en procedur för att göra en återkoppling till trafiksystemet, och som innebär att man finner en jämviktslösning.

De olika ruttalternativen kommer sannolikt också att vara inbördes beroende på ett sätt som vållar svårigheter med den statistiska metodansats som vi valt. Av dessa skäl används oftast modeller som direkt ger en jämviktslösning av ruttvalsproblemet (se t.ex. Florian 1982). Forskning som syftar till att göra modeller som förutom ruttvalet även innehåller flera andra steg som exempelvis färdmedelsval och destinationsval pågår på olika håll (Boyce et al 1983), men har hittills baserat sig på aggregerade individdata utan hushållskoppling.

När det gäller kollektivresor i tätort kan man ifrågasätta om det överhuvud taget är relevant att diskutera ruttvalet som ett val mellan ömsesidigt uteslutande alternativ. Det skulle bland annat förutsätta att man tvingas ner på en detaljnivå som beskriver enskilda avgångar, vilket ställer alltför stora krav på dataunderlaget. Det är i stället mer relevant att tillämpa en mer aggregerad ansats, där enskilda avgångar aggregeras till frekvenser.

Liksom när det gäller ruttvalet för bil finns olika programvaror utvecklade för ruttvalet i kollektivtrafiken. Sådana programsystem finns också i drift inom Stockholms län.

Även om ett detaljerat ruttval inte lämpar sig för den metodansats vi valt, kan man tänka sig att mer aggregerade ruttvalsproblem kan hanteras på detta sätt. Några sådana försök har dock inte genomförts inom ramen för detta arbete.

Körkortsinnehav

Körkortsinnehavet är väsentligt, dels för att avgöra om bil är ett relevant alternativ, och dels för att avgöra om bilkonkurrens föreligger inom hushållet. En analys (se Planeringsrapport 5) visar, att körkortsinnehavet uppvisar ett introduktionsförlopp. Detta innebär, att äldre generationer har ett lägre körkortsinnehav, men också att det ökar vartefter tiden går. 1971 var körkortsinnehavet bland manliga pensionärer 50 procent i Stockholms län, medan det hade ökat till 83 procent 1986. För kvinnliga pensionärer hade körkortsinnehavet ökat från 16 procent 1971 till 25 procent 1986.

För män mellan 26-45 år var andelen med körkort 95 procent, och för kvinnor 78 procent. Det förefaller därför rimligt att tro, att de flesta individer som kan skaffa sig körkort också kommer att göra det. En körkortsvalsmodell kommer därför huvudsakligen att beskriva när man skaffar sig körkort. Vi har inte bedömt det relevant att utveckla en valmodell för detta, utan förutsätter i stället att körkortsinnehavet skrivs fram i olika köns- och åldersgrupper med ledning av körkortsstatistik när prognosmodellen tillämpas.

Månadskortsinnehav

Förekomsten av färdbevisstyper som månadskort, där man erlägger en viss summa för ett ospecificerat antal resor under en viss tid, gör det svårt att beräkna en kostnad för enskilda resor. När man väl köpt ett månadskort, är ju marginalkostnaden noll kronor per resa. Innehavet av månadskort kan därför ha stor betydelse. Det är rimligt att tänka sig, att köpet av månadskort är en fråga om att ställa den extra kostnaden för att köpa ett kort (jämfört med de kostnader man annars skulle betala för sina kollektivresor) mot den ökade nytta som resandet i situationen med månadskortet skulle innebära.

Eftersom månadskortet är individuellt ställer en månadskortsvalsmodell baserad på förväntat resande stora krav på identifiering av de olika hushållsmedlemmarna i modellsystemet. Modellen blir därför ganska resurskrävande att skatta, och av detta skäl togs en sådan modell inte med.

Modellsystemets avgränsning till Stockholms län

Avsikten är att utveckla ett modellsystem för tillämpning i Stockholms län. Nedan följer en mer exakt definition av tillämpningsområdet.

Population

Den population som modellsystemet ska avse, begränsas till hushåll bestående av personer tillhörande den mantalsskrivna befolkningen i Stockholms län. Detta innebär att modellerna inte omfattar personer som tillfälligt befinner sig i Stockholms län. Hushåll med endast personer över 74 år ingår inte heller, eftersom svårigheterna med att intervjua dessa personer bedömts vara för stora.

Geografisk omfattning

Modellsystemet begränsas till resor inom Stockholms län. Detta innebär att resor över länsgränsen inte kommer med. Avgränsningen har främst gjorts för att resandet över länsgränsen är förhållandevis obetydligt. Dessa resor är delvis av samma karaktär som resorna inom länet, men utgörs också av mer långväga resor. De senare kan när det gäller resbeteendet skilja sig påtagligt från resorna inom länet, eftersom de till stor del utgörs av helgreisor och semesterresor. För sådana resor har modeller skattats i andra sammanhang.

Tidsperiod

Modellsystemet begränsas till att avse resor under vardagar. Denna begränsning har gjorts främst av resursskäl. En anpassning till att avse även övriga dagar är metodmässigt knappast något problem, men kräver ytterligare tid.

Resärende

De begränsningar som gjorts när det gäller resärende innebär dels att vi utesluter de resänderen som medför att man reser över länsgränsen, och dels att vi utesluter de gångförflyttningar som endast avser att gå ut och promenera.

4. Val av modelltyp

Vi har i tidigare avsnitt beskrivit grunderna för de modeller vi använder. Vi har valt en ansats som innebär att vi förutsätter att de beslutsenheter vi studerar (individer och hushåll) är nyttomaximerande och att de i definierade avseenden är rationella. I detta kapitel beskriver vi mer konkret de modeller som valts för projektet. För en utförligare beskrivning av ämnet hänvisas till Ben-Akiva och Lerman (1985).

4.1 Sannolikhetsbaserad beteendeteori

Den första utvecklingen av en sannolikhetsbaserad beteendeteori skedde inom psykologin (Thurston 1927). Behovet uppstod för att förklara experimentella observationer där försökspersoner visade inkonsistenta och icke-transitiva preferenser (Luce och Suppes 1965). Vid experimenten visade det sig att försökspersonerna inte valde samma alternativ vid upprepningar av samma experiment och att förändringar av valmängden gav upphov till val som ej var transitiva. En sannolikhetsbaserad valmekanism introducerades därför för att förklara dessa inkonsistenser.

Vi kan skilja mellan angreppssätt där beteendet i sig antas vara sannolikhetsbaserat och angreppssätt som innebär att beteendet enbart förefaller vara sannolikhetsbaserat för att vi inte kan beskriva beslutsprocessen helt fullständigt.

Angreppssättet i denna studie innebär att de observerade inkonsistenserna i beteendet antas bero på att vi inte kan observera och mäta de verkliga nyttorna helt säkert. Vi antar visserligen att individen alltid väljer alternativet med den största nyttan, men vi kan inte säkert beräkna alternativens nyttor. Dessa måste därför behandlas som stokastiska variabler ("random utility"). Den stokastiska komponenten orsakas av:

- ej observerade attribut
- ej observerade skillnader i individernas preferenser och erfarenheter
- mätfel och ofullständig information
- förekomsten av proxyvariabler

Vi definierar valmängden $C = \{c_1, \dots, c_i, \dots, c_N\}$ och en uppsättning attribut X för dessa N alternativ. En viss individ n har då en valmängd C_n som består av J_n alternativ och som utgör en delmängd av C .

För individ n har varje alternativ $i \in C_n$ en nytta U_{in} . Denna kan skrivas:

$$U_{in} = V_{in} + \varepsilon_{in}$$

V utgör den systematiska delen av nyttan och ε är den stokastiska delen. ε speglar alla de egenskaper hos individen och alternativet som enligt ovan ej kunnat observeras och mätas.

Den systematiska delen av nyttan utgörs normalt av en linjär funktion:

$$V(X_i) = \beta_1 * x_{i1} + \beta_2 * x_{i2} + \dots$$

där β är parametrar som ska estimeras.

Individen väljer det alternativ som maximerar nyttan, det vill säga individen väljer alternativet i om och endast om:

$$U_{in} \geq U_{jn}, \quad \forall j \in C_n$$

det vill säga om

$$V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}$$

Sannolikheten att individ n väljer i ges av:

$$P_n(i) = \Pr[V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}, \quad \forall j \in C_n, i \neq j]$$

eller

$$P_n(i) = \Pr[V_{in} + \varepsilon_{in} \geq \max_{j \in C_n} (V_{jn} + \varepsilon_{jn}), \quad i \neq j]$$

Eftersom fördelningen för den stokastiska delen ε inte är känd, är det inte möjligt att teckna ett analytiskt uttryck för modellen.

Olika antaganden om fördelningen för slumptermerna ε leder fram till olika modeller. Om ε -termerna antas oberoende och identiskt lika Gumbelfördelade (se nedan) erhålls den multinomiala logitmodellen. Antas de istället vara multivariat normalfördelade erhålls probitmodellen.

Skillnaden mellan normalfördelningen och Gumbelfördelningen är liten. I praktiken kan Gumbelfördelningen betraktas som en god approximation till normalfördelningen. Den största skillnaden mellan logit- och probitmodellerna är istället logitmodellens antagande om oberoende och lika fördelning för slumptermerna ε . Skillnaden kan illustreras med en varians-kovariansmatris för ε :

$$\begin{pmatrix} \sigma_{11} & \dots & \sigma_{1j} & \dots & \sigma_{1J} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{j1} & \dots & \sigma_{jj} & \dots & \sigma_{jJ} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{J1} & \dots & \sigma_{Jj} & \dots & \sigma_{JJ} \end{pmatrix}$$

där σ_{jj} = variansen hos ε för alternativ j
 σ_{ij} = kovariansen mellan ε för alternativen i och j

Antagandet om *oberoende* fördelning innebär att alla $\sigma_{ij} = 0$ för $i \neq j$, det vill säga att det inte finns beroenden mellan slumptermerna för olika alternativ. Sådana beroenden kan till exempel uppstå om mätfel för variabler särskilt gäller vissa alternativ eller om vissa individer har avvikande värderingar av vissa grupper av alternativ.

Antagandet om *lika* fördelning innebär att diagonalelementen är lika, det vill säga att alla ϵ har samma varians. Skillnader i varians kan enligt ovan orsakas av ett antal olika faktorer. Även skillnader i värderingar mellan olika individer (skillnader som inte gäller grupper av alternativ) kan ge skillnader mellan diagonalelementen.

Att ϵ är Gumbelfördelad innebär att den har fördelningsfunktionen

$$F(\epsilon) = e^{-e^{-\mu(\epsilon-\eta)}}, \quad \mu > 0$$

och täthetsfunktionen (frekvensfunktionen)

$$f(\epsilon) = \mu e^{-\mu(\epsilon-\eta)} e^{-e^{-\mu(\epsilon-\eta)}}$$

där η är en lägesparameter som anger det värde på ϵ för vilket $f(\epsilon)$ har sitt maximum och där μ är en positiv skalparameter. Fördelningen har följande egenskaper:

1. Fördelningens mest sannolika värde (typvärde, mode) är η
2. Medelvärdet är $\eta + \gamma / \mu$ där γ är Eulerkonstanten ($\sim 0,577$)
3. Variansen är $\pi^2 / 6\mu^2$
4. Om ϵ är Gumbelfördelad med parametrarna (η, μ) , samt V och $\alpha > 0$ är skalära konstanter, så är $\alpha\epsilon + V$ Gumbelfördelad med parametrarna $(\alpha\eta + V, \mu/\alpha)$. Egenskapen innebär att Gumbelfördelningen bevaras vid linjära transformationer.
5. Om ϵ_1 och ϵ_2 är oberoende Gumbelfördelade variabler med parametrarna (η_1, μ) respektive (η_2, μ) , så är $\epsilon^* = \epsilon_1 - \epsilon_2$ logistiskt fördelad:

$$F(\epsilon^*) = \frac{1}{1 + e^{\mu(\eta_2 - \eta_1 - \epsilon^*)}}$$

6. Om ϵ_1 och ϵ_2 är oberoende Gumbelfördelade variabler med parametrarna (η_1, μ) respektive (η_2, μ) , så är $\max(\epsilon_1, \epsilon_2)$ Gumbelfördelad med parametrar:

$$\left(\frac{1}{\mu} \ln(e^{\mu\eta_1} + e^{\mu\eta_2}), \mu\right)$$

7. Om $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_J)$ är J oberoende Gumbelfördelade variabler med parametrarna $(\eta_1, \mu), (\eta_2, \mu), \dots, (\eta_J, \mu)$ så är $\max(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_J)$ Gumbelfördelad med parametrarna:

$$\left(\frac{1}{\mu} \ln \sum_{j=1}^J e^{\mu \eta_j}, \mu\right)$$

4.2 Den multinomiala logitmodellen

Härledning

Om slumptermerna antas vara oberoende, lika fördelade och ha Gumbelfördelning erhålls enligt ovan den så kallade multinomiala logitmodellen. Modellen kan skrivas:

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}}$$

där $P_n(i)$ = sannolikheten att individ n väljer alternativ i
 μ = skalparameter

Modellen kan härledas med hjälp av de egenskaper hos Gumbelfördelningen som redovisats ovan i avsnitt 4.1.

Vi förutsätter att $\eta = 0$ för alla slumpstermer. När inte annat explicit anges gäller detta antagande generellt i kapitel 4. Antagandet innebär inte någon restriktion om alla alternativ har samma värde på η eftersom en godtycklig konstant kan läggas till den systematiska delen av alla nyttofunktioner utan att valsannolikheterna påverkas. Antagandet innebär inte heller någon restriktion om η skiljer sig mellan alternativen under förutsättning att varje nyttofunktion har en alternativspecifik konstant som kan fånga in denna skillnad. Om vi studerar sannolikheten att välja alternativet $i = 1$ erhålls:

$$P_n(1) = \Pr \left[V_{1n} + \varepsilon_{1n} \geq \max_{j=2, \dots, J_n} (V_{jn} + \varepsilon_{jn}) \right]$$

Vi definierar

$$U_n^* = \max_{j=2, \dots, J_n} (V_{jn} + \varepsilon_{jn})$$

Från egenskap 4 och 7 följer att U_n^* är Gumbelfördelad med parametrarna

$$\left(\frac{1}{\mu} \ln \sum_{j=2}^{J_n} e^{\mu V_{jn}}, \mu\right)$$

Med hjälp av egenskap 4 kan vi skriva $U_n^* = V_n^* + \varepsilon_n^*$ där

$$V_n^* = \frac{1}{\mu} \ln \sum_{j=2}^{J_n} e^{\mu V_{jn}}$$

och där ε_n^* är Gumbelfördelad med parametrar $(0, \mu)$. Vårt ursprungliga uttryck ger då:

$$P_n(1) = \Pr[V_{1n} + \varepsilon_{1n} \geq V_n^* + \varepsilon_n^*] = \Pr[\varepsilon_n^* - \varepsilon_{1n} \leq V_{1n} - V_n^*]$$

Enligt ovan är $\eta = 0$ för både ε_n^* och ε_{1n} . Egenskap 5 ger därmed direkt:

$$\begin{aligned} P_n(1) &= \frac{1}{1 + e^{\mu(V_n^* - V_{1n})}} = \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + e^{\mu V_n^*}} \\ &= \frac{e^{\mu V_{1n}}}{e^{\mu V_{1n}} + e^{\ln \sum_{j=2}^{J_n} e^{\mu V_{jn}}}} = \frac{e^{\mu V_{1n}}}{\sum_{j=1}^{J_n} e^{\mu V_{jn}}} \end{aligned}$$

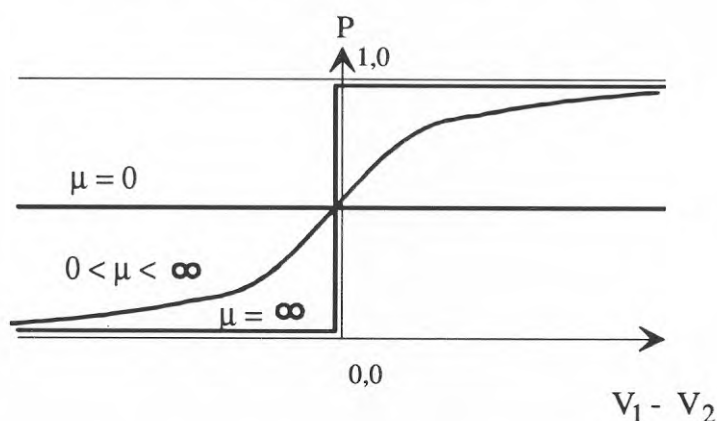
Den multinomiala logitmodellens egenskaper

Enligt avsnitt 4.1 beror storleken på skalfaktorn μ på variansen hos ε :

$$\mu^2 = \frac{\pi^2}{6 * \text{var}(\varepsilon)}$$

Ju mindre variansen i ε är (och ju större således skalfaktorn är) desto brantare blir logitfunktionen. Figuren nedan visar en grafisk representation av sannolikheten att välja ett visst alternativ i en binär modell för val mellan två alternativ. Sannolikheten uttrycks som en funktion av skillnaden i den systematiska delen av nyttofunktionen för de två alternativen:

Figur 4.1 Logitfunktionen vid olika skalfaktorer



Om variansen är 0 och μ således går mot oändligheten blir valet deterministiskt. Modellen pekar då *alltid* ut ett alternativ som valt så snart det är åtminstone något bättre med avseende på den systematiska delen av nyttofunktionen än andra alternativ.

Variansen hos ε har enligt ovan flera komponenter. En del beror på att det inte är möjligt att med modellernas variabler exakt fånga in varje individs värdering av olika faktorer (beroende på att olika individer har olika värderingar samt att variabler som borde ingå i modellen fattas). En annan del beror på mätfel, t.ex. i de restider som ansätts för olika alternativ.

Vid en "perfekt" modell där alla relevanta variabler finns med, där alla individer värderar variablerna lika och där inga mätfel finns blir μ stort och valet kan beskrivas som nära deterministiskt. Vid en mycket dålig modell där variansen är stor blir μ litet och modellens känslighet för variabelförändringar liten. μ kunde således utgöra ett mått på modellens godhet om parametern gick att skatta.

När logitmodellen estimeras kan dock värdet på μ ej separeras. De parametrar som estimeras, β , får istället följande värde:

$$\beta = \beta^0 * \mu$$

där β^0 är parametrarnas sanna värde.

En "bra" modell med högt värde på μ får därigenom också parametrar med relativt sett högre absolutvärden. Parametrarnas storlek utgör därigenom i viss grad ett mått på modellens kvalitet. Naturligtvis har dock varje parameter ett visst "sant" absolutvärde som motsvarar variabelns betydelse (givet att funktionsformen är sann). Parametrarnas storlek påverkas dessutom av mätskalan för variablerna, exempelvis blir kostnadsparametern 100 gånger större om kostnaden mäts i kronor än om den mäts i öre (så att en viss förändring får samma effekt oavsett mätskalan).

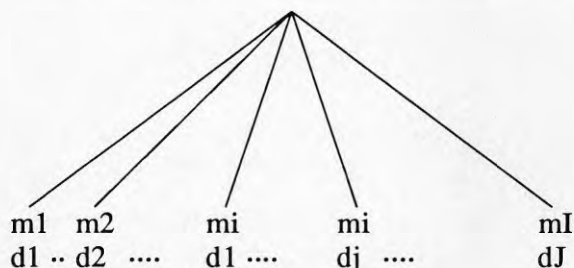
Alternativen i den multinomiala modellen kan antingen vara enkla alternativ som vissa färdmedel, eller sammansatta alternativ som t.ex. kombinationer av färdmedel till olika

destinationer. Om vi betraktar en modell för samtidigt val av färdmedel (m) och destination (d) får den följande utseende om vi utelämnar notationen för individ:

$$P(md) = \frac{e^{\mu V_{md}}}{\sum_{(m',d') \in C} e^{\mu V_{m'd'}}$$

Grafiskt kan modellen illustreras på följande sätt om det finns I färdmedel och J destinationer:

Figur 4.2 Simultan struktur för val av färdmedel och destination



Den multinomiala logitmodellen kallas även ofta för *simultan* logitmodell vilket ter sig naturligt i ovanstående exempel där flera val behandlas *samtidigt* på samma nivå i modellen.

Det kan visas (Ben-Akiva 1973) att denna modell rent matematiskt kan skrivas om som två separata modeller där den första uttrycker sannolikheten att välja ett visst färdmedel *givet* en viss destination, och den andra uttrycker sannolikheten att välja en viss destination. Vi betecknar nedan mängden möjliga färdmedel *givet* att destinationen d väljs med M_d och mängden destinationer med D:

$$P(m|d) = \frac{e^{\mu V_{md}}}{\sum_{m' \in M_d} e^{\mu V_{m'd}}}$$

$$P(d) = \frac{e^{\mu V_d + \ln \sum_m \exp(\mu V_{md})}}{\sum_{d' \in D} e^{\mu V_{d'} + \ln \sum_m \exp(\mu V_{m'd'})}}$$

$$P(md) = P(m|d) * P(d)$$

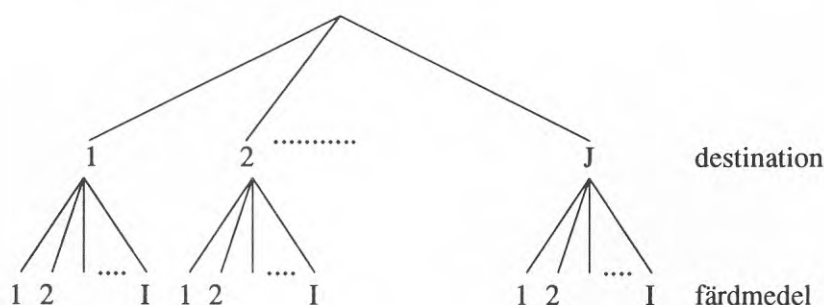
där:

$P(m|d)$ = sannolikheten att välja färd sättet m givet destinationen d

$\ln \sum_m \exp(\mu V_{md'})$ = logaritmen för nämnaren i färdmedelsvalsmodellen = logsumvariabeln

Färdmedelsvalet kan i detta fall tänkas ske på en lägre nivå där destinationsvalet är givet. Grafiskt kan strukturen illustreras på följande sätt:

Figur 4.3 Modell i två nivåer



Färdmedelsvalets inverkan på destinationsvalet överförs mellan modellerna av logsumvariabeln. Logsumvariabeln för över inverkan av tillgängligheten till de olika målpunkterna i form av restider och kostnader vägda med sina parametrar för alla tillgängliga färdmedel. Variabelns värde kommer därmed även att påverkas av vilka alternativ individen har tillgängliga, t.ex. om hushållet disponerar bil eller inte. Dessutom innehåller logsumvariabeln inverkan av alla andra variabler som ingår i färdmedelsvalsmodellen, t.ex. olika könsvariabler, konstanter, etc. För den enskilda observationen kommer variationen mellan destinationsalternativen dock primärt vara betingad av tillgänglighetsskillnader mellan dessa alternativ.

Alternativt kunde vi ha formulerat modellen i exemplet ovan med färdmedelsvalet på den övre nivån och destinationsvalet på den undre.

Den multinomiala logitmodellen har egenskapen att valet mellan två alternativ är oberoende av övriga alternativ ("Independence of Irrelevant Alternatives" ofta förkortat IIA). Denna egenskap framgår om vi tecknar sannolikheten för att välja alternativ i framför alternativ j :

$$\frac{P(i)}{P(j)} = e^{\mu(V_i - V_j)}$$

Egenskaper är förknippad med såväl för- som nackdelar. Fördelarna är främst att modellen kan estimeras genom att valet mellan en delmängd av samtliga relevanta alternativ studeras. Ett exempel på detta är modeller för valet mellan bil och kollektiva färdmedel där övriga färdmedel inte behandlas. När modellen är utvecklad kan effekten av nya alternativ i princip behandlas genom att det nya alternativet adderas till nämnaren i modellen.

Nackdelen är att modellen ger fel resultat om det nya alternativet konkurrerar hårdare med "likartade" alternativ jämfört med mindre likartade. Som exempel härpå brukar anföras fallet med bil och en blå buss som i ett utgångsläge vardera väljs av 50% av trafikanterna. Vi tänker oss att vi introducerar en ny busslinje med samma egenskaper som den gamla, där enda skillnaden är att färgen på den nya bussen är röd. Den nya bussen går från en annan hållplats som ligger lika långt bort från trafikanterna som den gamla och den går dessutom på samma tider som den gamla bussen, dvs. trafikanterna får samma trafikstandard som innan förändringen. Enligt den multinomiala logitmodellen kommer resultatet att bli att en tredjedel väljer respektive alternativ (förutsatt att den nya bussen behandlas som ett nytt alternativ i modellen). Intuitivt förefaller det rimligt att anta att den nya busslinjen drar åt sig fler trafikanter från den gamla busslinjen än från bilalternativet. Exemplet visar på en benägenhet att överskatta alternativ som uppfattas som lika av de som väljer. Problemet har sin rot i att de stokastiska elementen ϵ för de olika alternativen inte uppfyller kravet på oberoende och lika fördelning. De två bussalternativen kan i detta fall förväntas ha gemensamma stokastiska komponenter som skapar ett beroende mellan alternativen.

4.3 Den strukturerade logitmodellen

I många fall är det orealistiskt att anta att alternativen är oberoende. Detta kan gälla alternativen för en viss valsituation, som i fallet med bil och buss tidigare. Mer vanligt är att vi kan vänta oss ett beroende för val av samma typ, till exempel att färdmedelsalternativen och destinationsalternativen har stokastiska element som är inbördes mer lika, men som kan skilja sig mellan de två valsituationerna. Sådana skillnader mellan valsituationer är sannolika eftersom den stokastiska delen av nyttofunktionen enligt tidigare bland annat orsakas av mätfel, saknade variabler och proxyvariabler. Samtliga dessa tre faktorer bör normalt ha olika inverkan för olika valsituationer.

I de fall då det är rimligt att anta att de stokastiska elementen är oberoende och lika fördelade inom en viss valsituation, men detta krav ej är uppfyllt mellan de olika valen kan den strukturerade logitmodellen användas. Modellen kallas i litteraturen även hierarkisk logitmodell, sekventiell logitmodell, trädmodell eller "nested logit".

Vi tar fallet med färdmedels- och destinationsval som exempel. Vi kan då teckna nyttan av alternativet att välja destinationen d och färdmedlet m :

$$U_{dm} = \tilde{U}_d + \tilde{U}_m + \tilde{U}_{dm}$$

där \tilde{U}_d svarar mot den del av nyttan som beror av destinationen, \tilde{U}_m är nyttan (uppföringen) som enbart beror av färdmedlet och \tilde{U}_{dm} är den del av nyttan som beror av både färdmedlet och destinationen. Med tidigare notation erhålls:

$$U_{dm} = V_d + V_m + V_{dm} + \epsilon_d + \epsilon_m + \epsilon_{dm}$$

Till skillnad mot den vanliga multinomiala modellen har vi således även komponenter ε_d och ε_m som enbart berör vissa av valen i modellen. Detta innebär att nyttorna för alternativen inte längre är oberoende. I den strukturerade logitmodellen förutsätter vi att någon av de två komponenterna ε_d och ε_m har variansen noll (dvs. i praktiken att den är så liten att den kan försummas). Vi tillåter således en ytterligare en stokastisk komponent när modellen innehåller två nivåer, jämfört med den multinomiala logitmodellen.

Vi studerar som exempel samma valstruktur som i figur 4.3, dvs. en struktur med destinationsvalet överst och färdmedelsvalet underst. Denna struktur innebär av skäl som framgår nedan att vi förutsätter att $\text{var}(\varepsilon_m)=0$, dvs. att den komponent som enbart avser färdmedlen har försumbar varians. Vi betecknar den totala mängden destinationer med D och de möjliga destinationerna om man väljer färdmedel m med D_m . Vi betecknar vidare den totala mängden färdmedel med M och de möjliga färdmedlen om man väljer destination d med M_d . För enkelhetens skull anges inte något index för person n utan detta är underförstått.

Vi antar dessutom att:

- ε_d och ε_{dm} är oberoende
- ε_{dm} är identiskt Gumbelfördelade med skalparameter μ_m
- ε_d är fördelad så att $\max_{m \in M_d} U_{dm}$ är Gumbelfördelad med skalparameter μ_d

Med dessa förutsättningar kan den strukturerade modellen härledas. Vi tecknar först sannolikheten att välja alternativet d :

$$P(d) = \Pr \left(\max_{m \in M_d} U_{dm} \geq \max_{m \in M_{d'} } U_{d'm}, \forall d' \in D, d' \neq d \right) =$$

$$\Pr \left(V_d + \varepsilon_d + \max_{m \in M_d} [V_m + V_{dm} + \varepsilon_{dm}] \geq V_{d'} + \varepsilon_{d'} + \max_{m \in M_{d'} } [V_m + V_{d'm} + \varepsilon_{d'm}], \forall d' \in D, d' \neq d \right)$$

Eftersom vi förutsatt att ε_{dm} är Gumbelfördelad med skalparameter μ_m är även termen

$\max_{m \in M_d} [V_m + V_{dm} + \varepsilon_{dm}]$ Gumbelfördelad, men med parametrarna $\mu = \mu_m$ och

$$\eta = \frac{1}{\mu_m} \ln \sum_{m \in M_d} e^{(V_m + V_{dm})\mu_m}$$

Värdet på η ges av egenskap 7 hos Gumbelfördelningen (se avsnitt 4.1). Därmed kan vi skriva om sannolikheten ovan som:

$$P(d) = \Pr(V_d + V'_d + \varepsilon_d + \varepsilon'_d \geq V_{d'} + V'_{d'} + \varepsilon_{d'} + \varepsilon'_{d'}, \forall d' \in D, d' \neq d)$$

$$\text{där } V'_d = \frac{1}{\mu_m} \ln \sum_{m \in M_d} e^{(V_m + V_{dm})\mu_m}$$

$$\text{och där } \varepsilon'_d = \max_{m \in M_d} (V_m + V_{dm} + \varepsilon_{dm}) - V'_d$$

ε'_d är Gumbelfördelad med skalparameter μ_m (eftersom ε_{dm} enligt de valda förutsättningarna har denna fördelning). $\varepsilon_d + \varepsilon'_d$ är enligt de valda förutsättningarna Gumbelförde-

lad med skalparameter μ_d (eftersom $\max_{m \in M_d} U_{dm}$ förutsatts Gumbelfördelad med skal-

parameter μ_d och $\varepsilon_d + \varepsilon'_d$ är den stokastiska komponenten i detta uttryck). I analogi med härledningen av den multinomiala logitmodellen i avsnitt 4.2 blir därmed sannolikheten att välja alternativet d:

$$P(d) = \frac{e^{(V_d + V'_d)\mu_d}}{\sum_{d' \in D} e^{(V_{d'} + V'_{d'})\mu_d}}$$

Om vi sätter in uttrycket för V'_d och multiplicerar in μ_d samt definierar $\omega = \frac{\mu_d}{\mu_m}$ erhålls:

$$P(d) = \frac{e^{\mu V_d + \omega \ln \sum_m \exp(\mu V_{dm})}}{\sum_{d' \in D} e^{\mu V_{d'} + \omega \ln \sum_m \exp(\mu V_{d'm})}}$$

Av formeln framgår att skillnaden jämfört med den multinomiala modellen för motsvarande fall utgörs av parametern ω i den strukturerade modellen, samt att olika skalfaktorer tillåts på de två nivåerna. Eftersom variansen för en Gumbelfördelad variabel är omvänt proportionell mot kvadraten på skalfaktorn och $\text{var}(\varepsilon'_d)$ genom valet av förutsättningar är lika med $\text{var}(\varepsilon_{dm})$ och ε_d och ε_{dm} förutsatts oberoende erhålls:

$$\omega = \frac{\mu_d}{\mu_m} = \frac{\sqrt{\text{var}(\epsilon_{dm})}}{\sqrt{\text{var}(\epsilon_d + \epsilon'_d)}} = \frac{\sqrt{\text{var}(\epsilon_{dm})}}{\sqrt{\text{var}(\epsilon_d) + \text{var}(\epsilon_{dm})}}$$

Om varianserna för de stokastiska elementen i nyttofunktionen är lika på de två nivåerna, för vilket krävs att $\text{var}(\epsilon_d) = 0$, blir ω lika med ett och modellen blir identisk med den vanliga simultana logitmodellen.

Vi har redan tidigare noterat att $P(dm) = P(m|d) * P(d)$. Ovan härleds uttrycket för $P(d)$. Den betingade sannolikheten för att välja färdmedlet m givet att destinationen d väljs kan skrivas:

$$\begin{aligned} P(m|d) &= \Pr[U_{dm} \geq U_{dm'}, \forall m' \in M_d, m' \neq m | d \text{ vald}] \\ &= \Pr[V_{dm} + V_m + \epsilon_{dm} \geq V_{dm'} + V_{m'} + \epsilon_{dm'}, \forall m' \in M_d, m' \neq m | d \text{ vald}] \end{aligned}$$

De delar av den totala nyttan som består av V_d och ϵ_d kan utelämnas eftersom de är konstanta för alla alternativ i M_d . Eftersom ϵ_{dm} fyller kraven för den vanliga multinomiala logitmodellen erhålls i enlighet med härledningen i avsnitt 4.2:

$$P(m|d) = \frac{e^{\mu(V_m + V_{dm})}}{\sum_{m' \in M_d} e^{\mu(V_{m'} + V_{dm'})}}$$

Med vår härledning kan parametern ω ej vara över ett. Det går visserligen att formulera en strukturerad logitmodell konsistent med stokastisk nyttomaximering även för $\omega > 1$ (Anderstig och Mattsson 1991), men det fundamentala problemet är att modellen i det fallet kan ge korselasticiteter med fel storlek och/eller tecken (Williams och Senior 1977). Detta innebär till exempel att förbättringar som enbart gäller ett enda alternativ kan leda till att även andra alternativ får ökad valsannolikhet.

Om ω blir större än 1 vid estimeringen innebär detta att antagandet om $\text{var}(\epsilon_m) = 0$ bör ifrågasättas. Genom att vända på modellstrukturen och i stället studera $P(d|m)$ respektive $P(m)$ får parametern i allmänhet rätt storlek.

Ett skäl till att variansen är stor kan vara att valet är svårt att beskriva modellmässigt. Sådana val måste i konsekvens med resonemanget ovan placeras högt i modellstrukturen. Genom att föra in nya variabler eller genom att på andra sätt förbättra modellspecifikationen kan variansen i ϵ minskas. I stället för att vända på modellstrukturen kan man således försöka förbättra modellspecifikationen på den lägre nivån, om parametern ω blir större än ett.

Resonemanget visar även att den antagna modellstrukturen inte behöver innebära något antagande om vilken *ordning* valen sker i. Genom logsumvariabeln beaktas effekterna av valet på lägre nivåer när ett val på högre nivå sker. Vilken ordning i tiden valen följer är

därmed likgiltigt eftersom alla valen tänks påverka varandra. Valens tidsmässiga ordning kan dock i och för sig tänkas avspegla sig i strukturen. Långsiktiga beslut som fattas sällan kan ha varit påverkade av variabler som förändrat sig en hel del till undersökningstillfället (om data till modellerna samlas in genom en tvärsnittsundersökning). Vanliga exempel på sådana variabler är inkomst och hushållssammansättning - variabler som påverkar flera val inom trafikområdet. Det är därför sannolikt att beskrivningen av långsiktiga val blir osäker i nyttofunktionen och att sådana val måste placeras högt upp i modellstrukturen.

Hittills har vi endast diskuterat en strukturerad modell med två nivåer. Modellen kan dock generaliseras till ett godtyckligt antal nivåer. Generellt måste följande villkor vara uppfylla för de olika nivåerna i modellen:

- alla stokastiska komponenter som berör en viss nivå, men som inte också berör *alla* högre nivåer ska ha varians = 0
- alla stokastiska komponenter ska vara oberoende
- summan av de stokastiska komponenterna på en viss nivå och de stokastiska komponenterna på underliggande nivå ska vara identiskt Gumbelfördelad

På varje nivå består logsumvariabeln av logaritmen för summan av nämnaren för modellen på närmast lägre nivå (vilken i sin tur innehåller en logsumvariabel från ytterligare lägre nivåer).

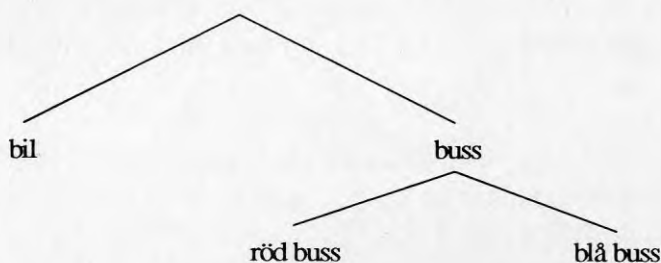
Som framgått är den strukturerade modellen mer generell än den multinomiala modellen. Det finns som diskuterats tidigare alltid anledning att tro att variansen för den stokastiska delen av nyttofunktionen kan skilja sig för olika valsituationer. Det är därför alltid motiverat att använda den strukturerade logitmodellen när flera val ska behandlas. De modeller som utvecklats i detta projekt är med få undantag av denna typ (undantagen gäller modeller som naturligen bara innehåller en enda nivå).

Oberoende av irrelevanta alternativ

Den strukturerade logitmodellen kännetecknas inte av oberoende av irrelevanta alternativ (IIA-egenskapen) *mellan* olika nivåer i modellen, däremot gäller egenskapen *inom* varje nivå.

Vi återvänder till exemplet med valet mellan bil och röd respektive blå buss. Eftersom vi har anledning att misstänka att alternativen inte är oberoende skulle vi kunna formulera följande modellstruktur:

Figur 4.4 Exempel på modellstruktur för beroende alternativ



Valet mellan de två bussarna sker på en lägre nivå i strukturen. På den högre nivån behandlas valet mellan bil och buss och där återfinns de variabler som förklarar valet mellan bil och buss (tider och kostnader). På den undre nivån återfinns variabler som förklarar valet mellan de olika bussarna. Om de röda och blå bussarna verkligen upplevs som helt lika av trafikanterna innebär detta att effekten av variabler som saknas och effekten av mätfel för variablerna är lika för båda bussalternativen. Variansen i slump termen ϵ på bussnivån blir då nära noll och parametern ω går mot noll. På den högre nivån försvinner därmed logsumtermens inverkan och fördelningen mellan bil och buss blir fortfarande 50 - 50 efter det att den nya bussen introducerats.

Vid en strukturering enligt ovan kommer ω således att spegla graden av likhet mellan alternativen. $\omega = 1$ ger den övre gränsen för den möjliga effekten av ett nytt alternativ (om det upplevs som helt fristående) och $\omega = 0$ den undre gränsen (om alternativet upplevs som identiskt med ett tidigare alternativ). Det kan noteras att den multinomiala modellen innebär att $\omega = 1$, vilket innebär att denna modell tenderar att överskatta effekten av att likartade alternativ tillkommer.

4.4 Alternativa ansatser

Probitmodellen

Om slump termerna ϵ i nyttofunktionen antas vara multivariat normalfördelade, med en godtycklig kovariansmatris, erhålls probitmodellen (Daganzo 1979). I probitmodellen kan således slump termerna i princip vara godtyckligt korrelerade. Detta innebär att modellen klarar av att hantera valsituationer där det finns mer komplicerade beroenden mellan alternativen.

Dessvärre har inte probitmodellen någon sluten analytisk form. För att skatta parametrarna i modellen har därför framför allt olika approximativa metoder utnyttjats. Dessa metoder har bara varit möjliga att använda vid valsituationer med ett starkt begränsat antal alternativ. Redan vid tre alternativ har det visats att felen i de approximativa metoderna kan vara stora (Horowitz et al 1982). Alternativa metoder som bygger på numerisk integration eller simulering har hittills krävt alltför omfattande beräkningar för att vara praktiskt användbara. En ytterligare komplikation med de algoritmer som hittills använts för att skatta probitmodellen är att det är svårt att veta om den erhållna lösningen är korrekt (om ett globalt eller lokalt maximum erhållits).

Om en godtycklig varians-kovariansmatris för alternativens slump termer tillåts måste ett mycket stort antal parametrar för dessa kovarianser skattas. Om restriktioner läggs på problemet, till exempel att ϵ förutsätts oberoende och lika normalfördelade, blir skillnaden mot den vanliga logitmodellen liten.

Forskning pågår för att utveckla estimeringsmetoder som gör probitmodellen praktiskt användbar. Framför allt är det den snabba datorutvecklingen som gjort att metoder som bygger på simulering nu börjar bli intressanta. Lam och Mahmassani (1991) utnyttjar beräkning med parallella processorer i superdatormiljö (Cray-datorer) och har hittills prövat metoden på fiktiva problem. Ben-Akiva och Bolduc (1991) använder också simulering men

även ansatser som begränsar antalet skattade parametrar i kovariansmatrisen. Praktiska exempel har visat på skillnader mot den vanliga logitmodellen i fall där beroenden mellan alternativen förekommer. Dessa praktiska exempel har dock ännu varit mycket begränsade (endast alternativspecifika konstanter och en enda ytterligare variabel) och endast omfattat 5 alternativ.

Mycket forskning återstår således innan probitmodellen är ett realistiskt alternativ för mer komplicerade modeller med många alternativ.

Alternativa funktionsformer

I våra analyser förutsätter vi att modellens funktionsform är given a priori. En intressant alternativ ansats är att samtidigt skatta parametrar för förklaringsvariabler och parametrar som bestämmer funktionsformen. Sådana analyser har framför allt utförts på aggregerade data och har då visats kunna ha betydelse för de resultat som erhålls (Gaudry och Wills 1978). Idag är denna typ av metoder oprövade för så komplicerade system av modeller som utvecklas i detta projekt. Området är dock intressant för fortsatt forskning.

5. Estimeringsteknik

I föregående kapitel valdes logitmodellen som matematisk funktionsform för trafikmodell-systemet. Denna kan - för en modell som är linjär i parametrarna (β) - skrivas

$$P_n(i) = \frac{e^{\beta' x_{in}}}{\sum_{j \in C_n} e^{\beta' x_{jn}}}$$

där $P_n(i)$ således beskriver sannolikheten för individ n att ur alternativmängden C_n välja alternativet i , som en funktion av (den transponerade) parametervektorn β' och attribut- eller variabelvärdesvektorn x_{jn}

För att skatta de modellparametrar - β - som behövs för att testa våra hypoteser krävs en skattningsmetod. Denna måste vara tillämpbar när det gäller att behandla ett antal praktiska problem som man ställs inför i samband med modellskattningen. Skattningsmetoden och de praktiska problemen diskuteras nedan. Framställningen är, där inget annat nämns, baserad på Ben-Akiva och Lerman (1985).

5.1 Val av skattningsmetod

I princip kan man tänka sig flera olika sätt att skatta en logitmodell. De viktigaste exemplen på möjliga skattningsmetoder är maximum likelihood-metoden och minsta kvadratmetoden. Även andra skattningsmetoder finns, exempelvis entropimaximeringsmetoden (Lundgren 1989). Anas (1983) har visat att logitmodellens parametervärden beräknas med identiska formler oavsett om man gör en individbaserad maximum likelihoodskattning eller om man använder samma material i aggregerad form i en entropimaximeringsberäkning. Entropimaximeringsmetoden ger dock inga möjligheter att beräkna variansen i enskilda parametervärden. När det gäller modeller med stora alternativmängder och ett individbaserat datamaterial finns det egentligen inget alternativ till maximum likelihood-metoden.

Maximum likelihood-metoden

Uppgiften är således att skatta parametervektorn β i logitmodellen. Utgångspunkten för skattningen är ett slumpmässigt urval hushåll, med vars hjälp vi ska skatta populationens β . För att definiera likelihoodfunktionen använder vi följande notation:

$$\begin{aligned} y_{in} &= 1 \text{ om beslutsenhet } n \text{ (ett hushåll eller en individ) väljer alternativ } i \\ y_{in} &= 0 \text{ om beslutsenhet } n \text{ (ett hushåll eller en individ) väljer alternativ } j \\ \hat{\beta}_k &= \text{det skattade parametervärdet för attribut } k \end{aligned}$$

x_{in} är en vektor av K olika attribut för alternativ i
 x_{jn} är en vektor av K olika attribut för alternativ j

För ett slumpmässigt urval innehållande N observationer gäller att urvalets likelihood-funktion är lika med produkten av observationernas likelihood. Om vi betecknar likelihoodfunktionen med L^* , kan den skrivas

$$L^* = \prod_{n=1}^N \prod_{i \in C_n} P_n(i)^{y_{in}}$$

Skattningen av β innebär att vi söker de parametervärden som maximerar likelihoodfunktionen. De parametervärden som ger maximivärdet på en funktion, ger också maximivärdet på en monoton transformation av funktionen. Detta gäller därför även logaritmerade funktioner (givetvis under förutsättning att funktionens värden kan logaritmeras).

Eftersom beräkningarna förenklas om likelihoodfunktionen logaritmeras, maximerar man den logaritmerade likelihoodfunktionen - L - i stället för den ursprungliga likelihoodfunktionen, dvs.

$$L = \sum_{n=1}^N \sum_{i \in C_n} y_{in} \left[\beta' x_{in} - \ln \sum_{j \in C_n} e^{\beta' x_{jn}} \right]$$

Första ordningens villkor för maximum är att förstaderivatorna av den logaritmerade likelihoodfunktionen med avseende på respektive parameter är lika med noll, vilket kan skrivas

$$\frac{\delta L}{\delta \beta_k} = \sum_{n=1}^N \sum_{i \in C_n} y_{in} \left[x_{ink} - \frac{\sum_{j \in C_n} e^{\beta' x_{jn}} x_{jnk}}{\sum_{j \in C_n} e^{\beta' x_{jn}}} \right] = 0$$

Eftersom

$$\sum_{i \in C_n} y_{in} \left[x_{ink} - \frac{\sum_{j \in C_n} e^{\beta' x_{jn}} x_{jnk}}{\sum_{j \in C_n} e^{\beta' x_{jn}}} \right] = \sum_{i \in C_n} y_{in} x_{ink} - \sum_{i \in C_n} y_{in} \frac{\sum_{j \in C_n} e^{\beta' x_{jn}} x_{jnk}}{\sum_{j \in C_n} e^{\beta' x_{jn}}}$$

$$\text{och } \sum_{i \in C_n} y_{in} \frac{\sum_{j \in C_n} e^{\beta' x_{jn}} x_{jnk}}{\sum_{j \in C_n} e^{\beta' x_{jn}}} = \sum_{i \in C_n} P_n(i) x_{ink}$$

(y_{in} är noll för alla utom ett alternativ, likgiltigt vilket, och summationsindex kan vara vilket som helst som går över alla alternativ) kan man också skriva

$$\frac{\delta L}{\delta \hat{\beta}_k} = \sum_{n=1}^N \sum_{i \in C_n} (y_{in} - P_n(i)) x_{ink} = 0$$

$$\text{eller } \frac{1}{N} \sum_{n=1}^N \sum_{i \in C_n} y_{in} x_{ink} = \frac{1}{N} \sum_{n=1}^N \sum_{i \in C_n} P_n(i) x_{ink}$$

Genom denna omskrivning blir det tydligt, att medelvärdet för en variabel beräknat för de valda alternativen är lika med medelvärdet beräknat över de modellberäknade sannolikheterna för respektive alternativ. Modellen kommer således att återskapa medelvärdena på alla medtagna variabler. I specialfallet att variabeln är en alternativspecifik konstant ($x_{ink} = 1$ för visst i och 0 annars) erhåller vi att

$$\frac{1}{N} \sum_{n=1}^N y_{in} = \frac{1}{N} \sum_{n=1}^N P_n(i)$$

eller med andra ord att den andel i urvalet som faktiskt valt alternativ i är lika med den modellberäknade andelen. Om man har med alternativspecifika konstanter för alla alternativ utom ett kommer således andelarna i urvalet att helt återskapas av den vanliga multinomiala modellen.

Andraderivatorna, vilka utnyttjas i den numeriska beräkningen av parametervärdena och deras varianser, ges av

$$\frac{\delta^2 L}{\delta \hat{\beta}_k \delta \hat{\beta}_l} = - \sum_{n=1}^N \sum_{i \in C_n} P_n(i) \left[x_{ink} - \sum_{j \in C_n} x_{jnk} P_n(j) \right] \left[x_{inl} - \sum_{j \in C_n} x_{jnl} P_n(j) \right]$$

McFadden (1974) visar att likelihoodfunktionen är globalt konkav redan under relativt svaga villkor, vilket innebär att om en lösning till maximeringsproblemet ovan existerar, så är den unik.

Några viktiga egenskaper hos maximum likelihoodestimatorn

Normalt skattas parametervärdena med hjälp av ett stickprov ur den population som vi vill studera. Detta innebär att resultatet kan avvika från populationsmedelvärdena, och därmed också att vi kan finna samband som inte gäller för populationen. För att kunna ta ställning till hur representativa de erhållna parametervärdena är, behöver vi veta något om estimatorns egenskaper i några olika avseenden. En önskvärd sådan egenskap skulle vara att väntevärdet är lika med populationsmedelvärdet vid varje urvalsstorlek, och en annan egenskap att estimatorns varians var känd vid varje urvalsstorlek. Maximum likelihood-

estimatoren saknar dessa egenskaper, vilka således avser givna urvalsstorlekar (s.k. små urval). Det kan däremot visas, att dessa egenskaper gäller asymptotiskt. För maximum likelihood-estimatoren gäller, att estimaten är

- konsistenta,
- asymptotiskt normalfördelade
- asymptotiskt effektiva

Konsistens

Denna egenskap innebär att estimaten konvergerar mot det sanna värdet i populationen när urvalet ökar. I små urval är estimaten således behäftade med en viss bias. Ortuzar och Willumsen anger - utan närmare motivering eller hänvisning - att urval på 500 - 1 000 observationer normalt är mer än tillräckliga (Ortuzar och Willumsen 1990).

Asymptotiskt normalfördelade

Detta innebär att estimatens avvikelser från de sanna värdena i populationen konvergerar mot normalfördelningen, när urvalets storlek ökar. Denna egenskap gör det möjligt att säga något om hur estimaten förhåller sig till det sanna värdet, om man kan beräkna variansens storlek för de olika parametrarna.

Asymptotiskt effektiva

För att kunna göra en så säker beräkning som möjligt av parametervärdena är det angeläget att estimatoren har en så liten varians som möjligt. Den minsta möjliga variansen ges av den s.k. Cramér-Rao-gränsen. Om β är en vektor med K element, så är $\text{var}(\hat{\beta}_N)$ en matris av dimensionen $K \times K$. Cramer-Rao-teoremet säger, att

$\text{var}(\hat{\beta}_N) \geq B^{-1}$, där

$$B = \begin{pmatrix} -E\left(\frac{\delta^2 L}{\delta \beta_1^2}\right) & -E\left(\frac{\delta^2 L}{\delta \beta_1 \delta \beta_2}\right) & \dots & -E\left(\frac{\delta^2 L}{\delta \beta_1 \delta \beta_K}\right) \\ -E\left(\frac{\delta^2 L}{\delta \beta_2 \delta \beta_1}\right) & -E\left(\frac{\delta^2 L}{\delta \beta_2^2}\right) & \dots & -E\left(\frac{\delta^2 L}{\delta \beta_2 \delta \beta_K}\right) \\ \dots & \dots & \dots & \dots \\ -E\left(\frac{\delta^2 L}{\delta \beta_K \delta \beta_1}\right) & -E\left(\frac{\delta^2 L}{\delta \beta_K \delta \beta_2}\right) & \dots & -E\left(\frac{\delta^2 L}{\delta \beta_K^2}\right) \end{pmatrix}$$

(E betecknar väntevärde.)

Denna gräns uppnås asymptotiskt (dvs. för stora urval). Variansen kan då också beräknas genom att beräkna matrisen av andraderivator.

5.2 Hypotestestning

För att kunna testa våra hypoteser har vi valt att formulera dem på ett sätt som innebär att de antingen kan testas genom att testa enstaka parametrar, eller genom att testa grupper av

parametrar. Maximum likelihood-estimatoren ger möjlighet till detta, dels genom det s.k. t-testet, och dels genom det s.k. likelihood ratio-testet.

t-testet

Maximum likelihood-estimatoren ger som nämnts ovan möjligheter att beräkna variansen i parametrarna $\hat{\beta}$, samtidigt som estimaten är (asymptotiskt) normalfördelade. Detta innebär, att vi kan genomföra ett test avseende storleken på de skattade parametrarna. Ett vanligt test är det s.k. t-testet, vilket innebär att man testat hypotesen att estimatet är lika med ett specificerat värde. En sådan test kan formuleras på följande sätt:

$$H_0: \beta_k = \beta^*$$

$$H_1: \beta_k \neq \beta^*$$

där H_0 betecknar nollhypotesen, och H_1 betecknar den alternativa hypotesen.

Om $\hat{\beta}_N$ är en vektor med normalfördelade element $\hat{\beta}_{Nk}$, så är statistikan $\frac{\hat{\beta}_{Nk}}{\hat{\sigma}_{Nk}}$

t-fördelad med $N-K$ frihetsgrader. N betecknar urvalets storlek, och K antalet skattade parametrar. Sannolikheten för att denna statistika ska ligga inom det intervall i t-fördelningen som ges av en given signifikansnivå α kan uttryckas på följande sätt:

$$\Pr \left[t_{N-K, \alpha/2} \leq \frac{\hat{\beta}_{Nk} - \beta^*}{\hat{\sigma}_{Nk}} \leq t_{N-K, 1-\alpha/2} \right] = 1 - \alpha$$

där $t_{N-K, \alpha/2}$ respektive $t_{N-K, 1-\alpha/2}$ anger de kumulerade frekvenserna för t-fördelningen med $N-K$ frihetsgrader. Vi kan då bilda en kritisk region med avseende på $\hat{\beta}_{Nk} - \beta^*$, för vilken nollhypotesen måste förkastas:

$$|\hat{\beta}_{Nk} - \beta^*| \geq t_{N-K, \alpha/2} \hat{\sigma}_{Nk}$$

Om t-värdet ligger innanför detta intervall (dvs. att olikheten inte gäller), så kan hypotesen H_0 inte förkastas. Det kan ibland finnas anledning att testa hypoteser av typen

$$H_1: \beta_k \geq \beta^*$$

$$H_0: \beta_k < \beta^*$$

Vi får då beräkna sannolikheten för en annat intervall, nämligen

$$\Pr \left[t_{N-K, \alpha} \leq \frac{\hat{\beta}_{Nk} - \beta^*}{\hat{\sigma}_{Nk}} \right] = 1 - \alpha$$

Den kritiska regionen blir då

$$\hat{\beta}_{Nk} - \beta^* \geq t_{N-K, \alpha} \hat{\sigma}_{Nk}$$

Detta test kallas ofta för ett ensidigt test, till skillnad från det föregående testet, som kallas dubbelsidigt.

Det som sagts ovan gäller när β_k är normalfördelad. I fallet med maximum likelihood-estimatoren är β_k endast asymptotiskt normalfördelad, och Cramér-Rao-gränsen används för att skatta variansen. I praktiken kommer vi att använda något som kan kallas "kvasi-t test", vilket innebär att vi använder t-testet med N-K frihetsgrader men bortser från att parameterestimatoren endast är asymptotiskt normalfördelad. När N-K är större än 30, blir det i praktiken fråga om en test mot normalfördelningen, eftersom t-fördelningen snabbt konvergerar mot normalfördelningen.

Om man vill testa skillnader mellan två koefficientvärden ($\hat{\beta}_1$ respektive $\hat{\beta}_2$) i samma modell kan man utnyttja statistikan

$$\frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\text{var}(\hat{\beta}_1) + \text{var}(\hat{\beta}_2) - 2\text{cov}(\hat{\beta}_1, \hat{\beta}_2)}}$$

för en asymptotisk t-test. Om man vill testa skillnader mellan koefficientvärden $\hat{\beta}_k$ i olika modeller a och b kan man utnyttja statistikan

$$\frac{\hat{\beta}_{ak} - \hat{\beta}_{bk}}{\sqrt{\text{var}(\hat{\beta}_{ak}) + \text{var}(\hat{\beta}_{bk})}}$$

för en asymptotisk t-test.

Likelihood ratio-testet

Givet en modell med ett antal parametrar $\hat{\beta}$, kan vi vilja testa om mer eller mindre restriktiva (i termer av antal parametrar) modeller är att föredra. Genom att först estimeras den icke begränsade modellen, erhåller vi värdet L^u på den logaritmerade likelihoodfunktionen. Genom att sedan estimeras en modell innehållande de restriktioner som vi vill testa,

erhåller vi värdet L^b på den logaritmerade likelihoodfunktionen för den begränsade modellen.

Genom att ha ett större antal parametrar i modellen, kommer vi åtminstone inte att få en sämre modell (i termer av värdet på likelihoodfunktionen), utan praktiskt taget alltid en viss förbättring. Frågan är naturligtvis om denna förbättring är så stor, att den inte bara avspeglar det ökade antalet frihetsgrader. Skillnaden mellan värdena för de logaritmerade likelihoodfunktionerna $L^b - L^u$ är detsamma som logaritmen för kvoten mellan värdena på de icke logaritmerade likelihoodfunktionerna - därav namnet likelihood ratio.

Theil (1971) har visat, att statistikan $-2(L^b - L^u)$ är fördelad som χ^2 med antalet frihetsgrader lika med antalet restriktioner i den begränsade modellen. Med hjälp av χ^2 -fördelningen kan man således testa nollhypotesen att den begränsade modellen gäller.

Det finns flera fall där det kan vara intressant att tillämpa detta test. Ett sådant fall är när man vill undersöka hela modellens giltighet. Därvid testas hypotesen att alla parametervärden är lika med noll, genom att först beräkna log likelihood-värdet för detta fall, och därefter log likelihood-värdet med de skattade modellparametrarna. Ur dessa värden kan sedan enkelt -2 *likelihoodratio-värdet beräknas och jämföras med det kritiska värdet i en χ^2 -tabell.

Oftast är modellerna mycket påtagligt signifikanta. Ett intressantare test kan vara att testa om de parametrar som inte utgörs av konstanter ger ett signifikant bidrag till en modell. Detta kräver på motsvarande sätt uppgifter om log likelihood-värdet för modellen med enbart konstanttermerna respektive den fullständiga modellen.

Det använda estimeringsprogrammet ALOGIT ger automatiskt log likelihoodvärdena för fallet med alla parametrar lika med noll samt för den fullständiga modellen. Vidare beräknas log likelihood-värdet motsvarande en modell med en full uppsättning alternativspecifika konstanter.

Ett ytterligare intressant testfall är när man vill undersöka om olika delar av populationen - marknadssegment - har samma parametervärden på alla eller på grupper av parametrar. I det förra fallet estimeras samma modell dels på hela urvalet, och dels på delpopulationerna var för sig. Delmodellernas log likelihood utgörs då av summan av log likelihood-värdena från respektive delmodell. Likelihood ratio-statistikan bildas på vanligt sätt, och antalet frihetsgrader blir lika med antalet delpopulationer minus ett gånger antalet parametrar i modellen.

5.3 Tester av modellstrukturen

Oberoende av irrelevanta alternativ

Logitmodellen förutsätter, som nämnts i kapitel 4, oberoende av irrelevanta alternativ. När detta antagande inte är uppfyllt, bör en strukturerad modell ansättas, vilken uppfyller denna förutsättning inom respektive nivå. Mot bakgrund av olikheterna mellan olika valdimensioner kan man förvänta sig vissa systematiska drag i modellstrukturen. Detta gäller exempelvis grupperingen i olika nivåer för färdmedelsalternativ och destinationsalternativ.

Tester av ansatta strukturer kan göras genom att testa om parametern för logsumvariabeln från lägre nivåer i en från början ansatt struktur är signifikant skild från ett. Man kan också testa om alternativen uppfyller förutsättningen om oberoende av irrelevanta alternativ genom att ansätta en strukturerad modell, och på samma sätt testa om logsumparametern är skild från ett. Våra tester av modellstrukturen har skett i form av sådana tester av logsumparametrar.

Smakvariation

Logitmodellen innebär att vi antar att såväl modellstruktur som modellparametrar gäller för hela den population som samplet representerar. Detta är naturligtvis ett mycket starkt antagande. Det finns olika sätt att hantera detta problem. Ett sätt är att ta med socioekonomiska variabler i modellen, vilka kommer att ta upp skillnaderna mellan olika kategorier i form av dummyvariabler. Ett annat sätt är att ansätta modeller som tillåter smakvariation, och ett tredje sätt är att dela upp urvalet i marknadssegment.

Om problemet exempelvis är att en del av populationen har en lägre känslighet för kostnader och man begränsar sig till att inkludera socio-ekonomiska dummyvariabler för att beakta detta, så kommer en förändring av reskostnaden inte att beskriva populationens faktiska beteende. Det är då mera relevant att ansätta olika kostnadsparametrar för de olika socio-ekonomiska segmenten.

Att ansätta en modell som explicit tillåter smakvariationer - som probitmodellen - är enligt tidigare inte beräkningstekniskt möjligt i vårt fall, eftersom modellerna är så omfattande. Detta sätt att hantera problemet innebär också att man antar att smakfördelningen är konstant, vilket kan vara fel om fördelningen mellan grupper med olika smak förändras.

Det sätt som normalt rekommenderas för att hantera problemet med smakvariationer utgörs av marknadssegmentering, vilket innebär att urvalet delas in i olika kategorier. Tester kan sedan genomföras avseende hela modellen eller delar av modellen i form av enskilda parametrar eller grupper av parametrar.

Tester avseende hela modellen genomförs med hjälp av likelihood ratio-testet som beskrivits ovan. Detta gäller också när man testar grupper av parametrar. Skillnader mellan enskilda parametrar kan testas med asymptotiskt t-test (se ovan). Det är dock inte säkert att skillnaderna mellan några parametrar blir signifikanta även om likelihood ratio-testet är det, och det är heller inte säkert att likelihood ratio-testet är signifikant även om skillnaderna mellan enskilda parametrar är det.

Konstant varians

Skalan i en logitmodell (se kapitel 4) är omvänt proportionell mot den stokastiska nytto-komponentens standardavvikelse. Eftersom skalan antas vara gemensam för alla observationer (på en viss nivå i modellen) förutsätts nyttofunktionen ha en konstant varians, vilket innebär att variansen i nyttofunktionens slumpterm är oberoende av nyttofunktionens storlek. Även detta förhållande kan testas, och också här kan man utnyttja en strukturerad ansats.

Om man exempelvis skattar en färdmedelsvalsmodell, och befarar att det på längre avstånd finns fler icke observerade faktorer av betydelse, kan man dela in observationerna i olika klasser och låta alla klasserna (utom en) ha olika logsumparametrar. Om logsumparametrarna inte blir signifikant skilda från ett kan inte misstankarna om heteroskedasticitet anses bekräftade.

5.4 Risker med hypotesprövning i svårupprepade urval

De tester som redovisas ovan innebär att hypoteser avseende hur olika parametervärden förhåller sig till populationens sanna värden förkastas under förutsättning att sannolikheten för olika händelser överstiger vissa kritiska värden. De skattade värdena är specifika för ett särskilt urval, och vissa parametrar kan råka ha en låg - eller hög - varians i just det aktuella urvalet. Detta innebär att man i vissa fall felaktigt kommer att förkasta eller inte förkasta dessa hypoteser.

Om man testar ett stort antal variabler, finns det risk för att man exploaterar urvalets mer extrema egenskaper. Även i ett helt slumpmässigt genererat datamaterial kan några variabelkombinationer visa sig korrelerade. Om det är förenat med stora svårigheter eller höga kostnader att åstadkomma ett nytt urval ur samma population, är det svårt att veta om man råkat ut för detta problem i något avseende.

Samtidigt är det heller inte försvarligt att inte fullt ut utnyttja datamaterialets möjligheter att testa olika specifikationer. Risken att felaktigt ha inkluderat eller uteslutit olika parametrar får därför belysas genom att motsvarande hypoteser testas på andra material.

5.5 Modellutveckling och statistiska tester i praktiken

Ovan har ett antal olika statistiska tester redovisats. Med hjälp av dessa kan olika hypoteser förkastas eller accepteras. Problemet att finna den "bästa" modellspecifikationen är inte enbart en fråga om att genomföra en rad statistiska tester på ett datamaterial. För att dessa tester ska vara helt konklusiva krävs såväl att den sanna funktionsformen är känd, som att data för alla relevanta variabler finns tillgängliga med tillräcklig precision och variationsvidd.

I verkligheten har vi förhållandevis små kunskaper om hur funktionsformen egentligen ska se ut, och står inför stora praktiska problem när det gäller att skatta mer komplicerade funktionsformer. Potentiellt viktiga data är svåra att mäta, eller uppvisar endast små variationer i praktiken.

Samtidigt har vi viss a priori information om ett antal parametrar, såväl när det gäller vilket tecken de ska ha, som i vilket intervall de bör ligga. Kostnads- och restidsparametrar bör normalt vara negativa, och förhållandet mellan dem får inte resultera i tidsvärden som är för extrema.

Det är därför knappast meningsfullt att alltför strikt låta modellspecifikationen bero på de olika statistiska testerna. Om en av de mer centrala parametrarna som exempelvis en res-

tidsparameter har rätt tecken, men inte riktigt når upp till de krav på precision som vi normalt ställer kan vi välja att ändå ha kvar variabeln i modellen. Detta rättfärdigas av att vi a priori vet att parametern ska vara negativ, av att vi vet att storleksordningen är rimlig och av att modellen annars skulle bli obrukbar i sin helhet.

Naturligtvis måste man i tillämpningen av en sådan modell vara medveten om att en central parameter är skattad med dålig precision, och att det är en angelägen uppgift dels att försöka utveckla modellens specifikation så att problemet försvinner eller minskar, och dels att försöka genomföra nya studier.

Ett annat fall, som kanske är vanligare, är frågan huruvida man bör slå samman parametrar som inte är signifikant skilda från varandra. Detta gäller dels uppdelning på restidskomponenter, och dels frågan om generiska respektive alternativspecifika variabler. Även här gäller, att skillnaderna kanske inte är signifikanta. Ändå kan vi välja att behålla uppdelningen om vi är övertygade att skillnaderna kan underbyggas med a priori information, som exempelvis resultat från andra studier.

5.6 Beräkningstekniska problem vid estimering av trädmodeller

Uttrycken för parameterestimaten är icke-linjära. Normalt beräknas parameterestimaten genom en iterativ procedur, exempelvis Newton-Raphsons metod. Denna innebär att man först finner en riktning i vilken parametervärdena ska ändras, och sedan gör en förändring i denna riktning. Om förändringen är tillräckligt liten, avbryter man proceduren, och de senast erhållna estimaten utgör slutresultatet. Proceduren utnyttjar den logaritmerade likelihoodfunktionens första- och andraderivator, och ger också, som biprodukt, kovariansmatrisen för estimaten.

Den vanliga Newton-Raphson-proceduren kan inte användas utan problem när det gäller att estimerade strukturerade modeller. Detta beror på att dessa inte är linjära i sina parametrar. När det gäller att estimerade strukturerade logitmodeller har det därför varit vanligt att estimerade modellerna sekventiellt. Detta har inneburit att man först estimerar modellen på den lägre nivån, och sedan beräknar värdet på logsumman för respektive observation och använder denna som en vanlig variabel på nästa nivå.

Det finns flera nackdelar med detta sätt att estimerade strukturerade modeller. En nackdel är att man på den lägsta nivån endast använder en del av den information som finns i datamaterialet, vilket kan innebära att man inte lyckas erhålla acceptabla resultat på denna nivå, och att det således inte går att testa hypoteser om modellstruktur. Exempel på sådana fall är modeller för inköpsresor i Västerås och Jönköping (Algers, Colliander och Widlert 1987).

En annan nackdel är att kovariansmatrisen för estimaten på den högre nivån inte blir korrekt beräknad i standardprogram för logitmodeller, utan i stället underskattad. Standardavvikelseerna blir därför för små för alla parameterestimat på den högre nivån. Det finns visserligen möjligheter att beräkna de korrekta standardavvikelseerna (McFadden 1981), men dessa används inte i praktiken eftersom de inte ingår i vanliga estimeringsprogram.

En ytterligare nackdel är att processen att genomföra sekventiell estimering är såväl tidsödande som utsatt för risker för fel vid beräkningen av logsummorna.

Dessa nackdelar har delvis eliminerats genom att den programvara som vi använt oss av - ALOGIT - utvecklats i en rad avseenden, delvis på vårt initiativ. De viktigaste förändringarna utgörs av möjligheterna att direkt estimeras hela trädstrukturer, och möjligheten att utnyttja programmet för att beräkna logsummer. Även om möjligheterna att estimeras hela strukturerade modeller har ökat betydligt är de modeller som skattats inom vårt avhandlingsarbete så omfattande, att vi ändå måste använda oss av sekventiell estimering i vissa fall. Därför är möjligheten att genomföra beräkningen av logsummer inom ramen för estimeringsprogrammet ändå mycket värdefull.

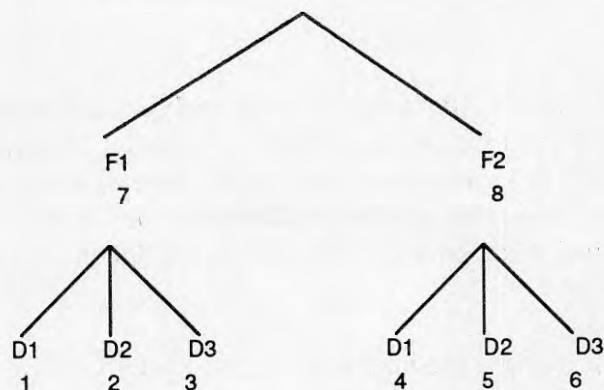
Skattningsmetoden för strukturerade modeller redovisas utförligt av Daly (1986 och 1982), och återges här i förkortad form.

En redovisning av estimeringsförfarandet vid trädmodeller kräver att man inför ett sätt att beskriva en trädstruktur. Det man gör när man strukturerar en valmängd C är att man definierar olika grupper av de ursprungliga alternativen C. Dessa grupper, eller "valelement", bildar således tillsammans med de ursprungliga alternativen en större mängd valelement S, i vilken den ursprungliga valmängden C är en delmängd. Trädet definieras med avseende på dessa valelement. Valelementen består således antingen av ett enda alternativ eller av övergripande alternativ högre upp i hierarkin. Dessa valelement får inte överlappa - för två element gäller att de antingen måste skilja sig helt eller att det ena helt innesluter det andra.

För att beskriva en bestämd hierarki definieras en funktion t , som för varje valelement s pekar på det alternativ som ingår på närmast högre nivå. Denna funktion är tillräcklig för att för varje valelement definiera två delmängder av S - dels de alternativ som denna mängd ingår i på högre nivåer (A), och dels de alternativ som ingår på den aktuella nivån (B).

För att illustrera hur en trädstruktur kan beskrivas på detta sätt, kan vi tänka oss ett exempel med färsätts- och destinationsalternativ. I detta exempel finns tre destinationer D1 - D3 och två färsätt F1 - F2. Trädstrukturen framgår av figuren nedan:

Figur 5.1 Valelement i trädstruktur



I detta fall finns 8 valelement. Varje alternativ i de två mängderna D1-D3 under F1 respektive F2 utgör ett valelement, vilket ger 6 valelement. På den högre nivån utgör F1 och

F2 varsitt valelement. Därutöver finns ett "dummyelement", som anger trädets rot. Strukturen kan nu anges genom att definiera funktionen $t(s)$. Förutsatt att strukturen inte innehåller några upprepningar kan en sådan funktion kallas ett träd, och utgör en entydig definition av trädstrukturen. I modellen ovan är alltså

$$\begin{aligned} t(1) &= 7 \\ t(2) &= 7 \\ t(3) &= 7 \\ t(4) &= 8 \\ t(5) &= 8 \\ t(6) &= 8 \\ t(7) &= 9 \\ t(8) &= 9 \end{aligned}$$

där 9 är rotelementet. Denna funktion är också tillräcklig för att för varje valelement definiera dels alternativen på högre nivåer (A), och dels de alternativ som ingår på samma nivå (B). Med informationen ovan kan A respektive B bestämmas (A och B har definierats så att det aktuella valelementet ingår i båda delmängderna):

$$\begin{array}{ll} A(1) = 1,7 & B(1) = 1,2,3 \\ A(2) = 2,7 & B(2) = 1,2,3 \\ A(3) = 3,7 & B(3) = 1,2,3 \\ A(4) = 4,8 & B(4) = 4,5,6 \\ A(5) = 5,8 & B(5) = 4,5,6 \\ A(6) = 6,8 & B(6) = 4,5,6 \\ A(7) = 7 & B(7) = 7,8 \\ A(8) = 8 & B(8) = 7,8 \end{array}$$

Med utgångspunkt från funktionen $t(s)$ och definitionen av A och B kan den strukturerade logitmodellen skrivas

$$P(i, t) = \prod_{a \in A(i, t)} l(a, B(a, t))$$

där

- $P(i, t)$ = sannolikheten att välja alternativ i från den ursprungliga valmängden C
- $A(i, t)$ = de valelement i strukturen t som ligger ovanför i (inklusive i 's nivå)
- $B(a, t)$ = de valelement som tillhör samma grupp valelement som a i strukturen t
- $l(a, B(a, t))$ = sannolikheten att välja alternativ a bland de alternativ som ingår i samma grupp beräknad med den multinomiala logitmodellen
- t = given struktur

Sannolikheten att välja alternativ i är således produkten av sannolikheterna på respektive nivå i i det av t -funktionen definierade träd. I exemplet ovan kan alltså sannolikheten att välja exempelvis alternativ nr 5 uttryckas som

$$P(5) = P(5, \{4, 5, 6\}) * P(8, \{7, 8\})$$

där klammerparentesen betecknar mängden valelement.

Med hjälp av notationen ovan kan vi (med samma utgångspunkt när det gäller sättet att göra urvalet som tidigare, dvs. ett slumpmässigt urval) ställa upp ett uttryck för likelihoodfunktionen (observera att β nu också innefattar strukturparametrarna, dvs. logsumparametrarna):

$$L^*(\beta, t) = \prod_{n=1}^N P_n(i_n, t)$$

där i_n är det valda alternativet för observation n och P definierats som ovan. Likelihoodfunktionen är således en funktion av såväl modellens parametervektor som trädstrukturen, och man skulle därför kunna tänka sig att maximera över båda dessa. Någon tillämpbar metod för att maximera även över t , som är en diskret till sin natur, har ännu inte utvecklats. Vi får därför begränsa oss till att maximera över parametervektorn, för en given struktur.

Som när det gäller den multinomiala modellen är det tillräckligt att maximera den logaritmerade likelihoodfunktionen. Att skriva likelihoodfunktionen som funktion av de strukturella och "normala" parametrarna skulle dock leda till en komplicerad notation. För en nivå s som ligger ovanför den lägsta kan man skriva

$$V_s = X_s + h_s \ln \sum_{t(u)=s} e^{V_u}$$

där X_s är de "normala" parametrarna multiplicerade med respektive variabelvärde, och h_s är den strukturella parametern för nivå s . Logsumman innehåller motsvarande uttryck för de alternativ som finns på nivån under (och vars trädfunktion t alltså pekar på s). Med utgångspunkt från nyttofunktionerna V kan den logaritmerade likelihoodfunktionen skrivas

$$L(\beta, t) = \sum_{n=1}^N \sum_{a \in A(i_n, t)} \left(V_{an} - \ln \sum_{b \in B(a, t)} e^{V_{bn}} \right)$$

där index n avser observationerna i urvalet. Vi utnyttjar då att

$$\ln(P(i, t)) = \sum_{a \in A(i, t)} \ln(l(a, B(a, t))) = \sum_{a \in A(i, t)} \left(V_a - \ln \sum_{b \in B(a, t)} e^{V_b} \right)$$

där V_a är nyttofunktionen för det valda valelementet på nivå a och V_b är nyttofunktionen för de valelement b som tillhör gruppen a .

På samma sätt som i det multinomiala fallet kan vi ställa upp första ordningens villkor för den logaritmerade likelihoodfunktionens maximum (summering över individer skrivs inte ut):

$$f^*_{*i} = \sum_{a \in A(r)} \left(V_{a,i} - \sum_{b \in B(a)} l(b, B(a)) V_{b,i} \right) = 0$$

där f^*_{*i} betyder derivatan av den logaritmerade likelihoodfunktionen med avseende på den i :te parametern i parametervektorn β , och $V_{a,i}$ betyder derivatan av V_a med avseende på parameter i . l betecknar som tidigare den multinomiala logitmodellen.

Det förefaller oklart om man här - som i fallet med den multinomiala logitmodellen - säkert kan säga att medelvärdena för de ingående variablerna återskapas av modellen. Enligt Daly 1987 är det möjligt att detta endast gäller asymptotiskt.

För att beräkna andraderivatan måste vi differentiera en gång till, och vi betecknar andra derivatan med index ij :

$$f^*_{*ij} = \sum_{a \in A(r)} \left(V_{a,ij} - V_{B(a),ij} + V_{B(a),i} V_{B(a),j} - \sum_{b \in B(a)} l(b, B(a)) V_{b,i} V_{b,j} \right)$$

där r betecknar det valda alternativet, och

$$V_{B(a),i} = \sum_{b \in B(a)} l(b, B(a)) V_{b,i}$$

De två sista termerna i uttrycket för andraderivatan är analogs med andraderivatan i det multinomiala fallet och utgör en negativt definit matris. De första två termerna består av andraderivator med avseende på V , vilka här inte är lika med noll eftersom V inte är linjär i parametrarna. Detta ser man om man differentierar uttrycket för V (för någon nivå s)

$$V_a = X_a + h_a \ln \sum_{S(a)} e^{V_b}$$

två gånger, varvid man erhåller

$$V_{a,i} = X_{a,i} + h_a V_{S(a),i}$$

respektive

$$V_{a,ij} = X_{a,ij} + h_a \left(V_{S(a),ij} - V_{S(a),i} V_{S(a),j} + \sum_{b \in S(a)} l(b, B(a)) V_{b,i} V_{b,j} \right)$$

där $S(a)$ är den mängd valelement b som uppfyller $t(b) = a$.

Förekomsten av dessa ytterligare båda termer i andraderivatan innebär att man inte längre kan vara säker på att beräkningsalgoritmen konvergerar. Modifieringen av algoritmen innebär, att man ibland (när den sanna matrisen av andraderivator är indefinit) utnyttjar en

annan matris som beräknats på annat sätt och som har de egenskaper som krävs för konvergens (att matrisen är definit). Denna andra matris utgörs av matrisprodukten av matrisen av förstaderivator och den transponerade matrisen av förstaderivator. Utnyttjandet av denna matris rättfärdigas dels av att denna matris säkert har de nödvändiga egenskaperna (är definit), och dels av att minus väntevärdet för denna matris är lika med väntevärdet för matrisen med andraderivator.

Algoritmen anges söka sig fram till det område där matrisen av andraderivator är definit inom 3-4 iterationer. Existensen av ett sådant område runt det "sanna" värdet på den logaritmerade likelihoodfunktionen - som åtminstone är ett lokalt optimum - anges också förekomma givet vissa villkor om regularitet. Vi har inte noterat att flera optima förekommit i våra modeller, trots att vi har skattat ett stort antal modeller med olika startvärden.

Estimatens kovariansmatris erhålls också vid likelihoodfunktionens optimum. Detta innebär att denna direkt kan utnyttjas för att exempelvis genomföra (approximativa) t-tester. Möjligheterna att utföra likelihood ratio-tester finns fortfarande kvar.

5.7 Några speciella estimeringsproblem

Förutom problemen med att estimerade strukturerade modeller finns det ytterligare två estimeringsproblem som kräver särskilda åtgärder. Det ena problemet utgörs av önskemålet att kunna skatta sammansatta storleksvariabler, och det andra problemet utgörs av kravet att kunna hantera stora mängder alternativ.

Storleksvariabler

Speciellt när det gäller destinationsalternativ uppkommer behov av att på något sätt beakta ett områdes storlek. För att illustrera detta behov kan vi ta en inköpsresa. Målpunkten för en inköpsresa är ju egentligen inte ett område som innehåller ett antal butiker, utan en viss butik, i vilken inköpet kan göras. Det finns en mycket stor mängd enskilda butiker, och det skulle föra mycket långt att ta med alla butiker som egna alternativ i modellen.

Det finns därför ett behov att ibland aggregera sådana s.k. elementaralternativ, vilka således är de alternativ en individ egentligen väljer mellan. Detta kan ske på följande sätt. Låt L beteckna den totala mängden elementaralternativ. Låt $P_n(\ell)$ beteckna sannolikheten att individ n väljer elementaralternativ $\ell \in L$. Dela upp mängden L i icke överlappande delmängder:

$$L_i \subseteq L, i = 1, \dots, J$$

där J är antalet aggregerade alternativ i valmängden C . Sannolikheten för individ n att välja det aggregerade alternativet $i \in C$ ges av

$$P_n(i) = \sum_{\ell \in L_i} P_n(\ell), i = 1, \dots, J$$

Detta innebär således, att sannolikheten att välja ett aggregerat alternativ är summan av sannolikheterna för elementaralternativen i det aggregerade alternativet. Sannolikheten att välja ett aggregerat alternativ antas bero på elementaralternativens nytta. Nyttan för ett elementaralternativ kan tecknas:

$$U_{\ell n} = V_{\ell n} + \varepsilon_{\ell n}$$

där $V_{\ell n}$ och $\varepsilon_{\ell n}$ är den systematiska respektive den stokastiska delen av elementaralternativet ℓ 's nytta för individ n . Eftersom elementaralternativen är varandra ömsesidigt uteslutande, kan nyttan av ett aggregerat alternativ definieras av

$$U_{in} = \max_{\ell \in L_i} (V_{\ell n} + \varepsilon_{\ell n})$$

Den kan också uttryckas som summan av väntevärdet, betecknat med V_{in} , och en stokastisk komponent, betecknad med ε_{in} :

$$U_{in} = V_{in} + \varepsilon_{in}, \quad i = 1, \dots, J,$$

$$\text{där } V_{in} = \mathcal{E} \left(\max_{\ell \in L_i} (V_{\ell n} + \varepsilon_{\ell n}) \right)$$

Elementaralternativens genomsnittliga systematiska nytta kan definieras på följande sätt:

$$\bar{V}_{in} = \frac{1}{M_i} \sum_{\ell \in L_i} V_{\ell n}$$

där M_i är antalet elementaralternativ i mängden L_i . Sambandet mellan det aggregerade alternativets systematiska nytta och elementaralternativens genomsnittliga nytta beror på de senares fördelningsfunktion.

Om det aggregerade alternativet i innehåller ett stort antal elementaralternativ, och om nyttorna är oberoende och identiskt lika fördelade (IID), kan det visas att det aggregerade alternativets nytta närmar sig Gumbelfördelningen, vilket innebär att vi kan ansätta en logitmodell för val mellan olika aggregerade alternativ. IID-antagandet innebär, att vi antar att elementaralternativen har samma medelvärde, medan den stokastiska komponenten är IID.

Om vi i stället antar, att elementaralternativen har olika medelvärden och att den stokastiska komponenten är IID Gumbel med skalparameter μ^* , så är nyttan för det aggregerade alternativet också extremvärdesfördelad med typvärde (mode, se egenskap 7 i kap 4.1)

$$\eta = \frac{1}{\mu^*} \sum_{\ell \in L_i} e^{V_{\ell n}}$$

vilket innebär att nyttan för det aggregerade alternativet - efter omskrivning - kan uttryckas som

$$U_{in} = \bar{V}_{in} + \frac{1}{\mu^*} \ln M_i + \frac{1}{\mu^*} \ln B_{in} + \varepsilon_{in}$$

där \bar{V}_{in} är elementaralternativens genomsnittsnytta, $\frac{1}{\mu^*} \ln M_i$ är en term som avspeglar alternativets storlek (i termer av antal elementaralternativ), och

$$B_{in} = \frac{1}{M_i} \sum_{\ell \in L_i} e^{\mu^*(V_{\ell n} - \bar{V}_{in})}$$

är ett mått på hur heterogena elementaralternativen är för alternativ i. Om alla elementaralternativ är homogena, är B_{in} lika med ett och har ingen betydelse. Det kan visas, att även om elementaralternativen är heterogena kan man bortse från detta om variansen för elementaralternativens nyttor är densamma för alla aggregerade alternativ.

En logitmodell för val mellan aggregerade alternativ kan således skrivas på följande sätt:

$$P_n(i) = \frac{e^{\mu V_{in}}}{\sum_{j=1}^J e^{\mu V_{jn}}}$$

$$= \frac{e^{\mu \bar{V}_{in} + \mu' \ln M_i + \mu' \ln B_{in}}}{\sum_{j=1}^J e^{\mu \bar{V}_{jn} + \mu' \ln M_j + \mu' \ln B_{jn}}}, \quad i = 1, \dots, J,$$

och där μ' är en positiv skalparameter och $\mu' = \mu/\mu^*$.

Antagandet om att elementaralternativens nyttor är IID Gumbel innebär att det här är frågan om en trädmodell, där valet av elementaralternativ ligger på en undre nivå. μ' återspeglar förhållandet mellan skalparametrarna på de två nivåerna, och därför också förhållandet mellan fördelningarna av de stokastiska termerna ε_{in} och $\varepsilon_{\ell n}$. Korrelationen mellan nyttan för elementaralternativ tillhörande olika aggregerade alternativ antas vara noll, medan korrelationen mellan nyttan för elementaralternativ i samma aggregerade alternativ är $(1-\mu'^2)$. Om korrelationen är låg, är alltså μ' nära ett. Om μ' sätts till 1 blir modellen okänslig för aggregeringsnivån.

För att skatta modeller med destinationsalternativ kan man således använda logaritmen av antalet s.k. elementaralternativ (exempelvis butiker i fallet med inköpsresor) som mått på ett områdes relativa attraktion. Parametern för denna variabel kan skattas som vilken annan parameter som helst, men låses vanligen till värdet ett, eftersom den endast för detta värde

är oberoende av områdesindelningen. Avvikelser från värdet ett kan vara tecken såväl på att homogenitetskravet inte är uppfyllt, som på att det valda måttet på antalet elementaralternativ på ett strukturellt sätt samvarierar med områdets på detta sätt definierade storlek.

Man kan emellertid vilja beskriva ett områdes storlek med fler än en variabel. Om vi återgår till exemplet med inköpsresorna kan man tänka sig att både inköpsställen för dagligvaror och inköpsställen för sällanköpsvaror utgör elementaralternativ - men i olika grad. Denna gradskillnad skulle på något sätt kunna ansättas a priori, men det är naturligtvis mest effektivt att låta även denna vara ett estimeringsresultat, i likhet med andra variabler.

En sammansatt storleksvariabel kan definieras på följande sätt:

$$M_i = \sum_{k=K'+1}^K \beta_k x_{ink} \quad , i = 1, \dots, J$$

där
 $(K-K')$ = antal storleksvariabler (K' är antalet övriga variabler, vilka beskriver området för övrigt)
 $x_{ink}, k=K'+1, \dots, K$ = variabelvärden för det aggregerade alternativet i och individ n

Uttrycket måste vara linjärt, eftersom storleksbeskrivningen annars blir beroende av aggregeringen (exempelvis en områdesindelning). Vidare måste koefficienterna och variablerna vara positiva, eftersom de ska avspegla storlek. Det är emellertid inte möjligt att identifiera en separat parameter för varje variabel, vilket kan visas genom följande omskrivning (k avser enbart storleksvariabler) :

$$\frac{e^{\ln \sum_k \beta_k x_{ik}}}{\sum_j e^{\ln \sum_k \beta_k x_{jk}}} = \frac{\sum_k \beta_k x_{ik}}{\sum_j \sum_k \beta_k x_{jk}}$$

Eftersom β -vektorn kan multipliceras med en godtycklig positiv konstant utan att kvoten ovan förändras, kan parametrarna inte bestämmas entydigt. Genom att sätta parametern för en av variablerna till ett, vilket innebär att denna variabel bestämmer skalan på storleksvariabeln, löses problemet.

Förekomsten av sammansatta storleksvariabler innebär att nyttofunktionen blir icke-linjär i sina parametrar, i likhet med den strukturerade modellen. För att tillåta estimering av sådana sammansatta storleksvariabler krävs således - liksom för strukturerade modeller - att beräkningsalgoritmen anpassas till logitmodeller som är icke-linjära i sina parametrar.

De modifieringar av den normala Newton-Raphson-metoden som krävs för detta är desamma som de som krävs för en simultan estimering av strukturerade modeller (se kap. 5.6). I själva verket genomfördes dessa modifieringar i syfte att kunna skatta sammansatta storleksvariabler, och har sedan utnyttjats även för den simultana estimeringen av strukturerade modeller.

Sampling av alternativ

När alternativen är homogena kan man således hantera problemet med stora alternativmängder med ett relativt enkelt aggregeringsförfarande. När alternativen inte kan betraktas som homogena, eller när man av andra skäl inte vill aggregera alternativen - som är fallet när det rör sig om områdesalternativ - måste andra metoder tillgripas för att antalet alternativ ska bli hanterligt i beräkningssammanhang.

Ett sätt att lösa problemet är att utnyttja en av logitmodellens egenskaper - oberoende av irrelevanta alternativ (IIA) - och göra ett urval ur mängden alternativ. Logitmodellen kan estimeras konsistent även om man bara utnyttjar en delmängd av de olika alternativen (inklusive det valda alternativet) (Ben-Akiva och Lerman 1985). Man utnyttjar då det faktum, att man kan uttrycka en betingad sannolikhet för att välja ett alternativ - betingad med avseende på de ingående alternativen.

Sannolikheten för att ur den totala alternativmängden C_n dra ett urval alternativ D samt det valda alternativet i kan skrivas

$$\pi_n(i, D) = \pi_n(D|i) \cdot P_n(i), \quad i = 1, \dots, J.$$

Genom att tillämpa Bayes teorem kan man skriva sannolikheten för att välja alternativ i , givet ett urval alternativ D

$$\pi_n(i|D) = \frac{\pi_n(D|i) \cdot P_n(i)}{\sum_{j \in D} \pi_n(D|j) \cdot P_n(j)}$$

Denna betingade sannolikhet $\pi_n(i|D)$ existerar om

$$\pi_n(D|j) > 0 \quad \forall j \in D$$

Denna egenskap - "positive conditioning property" - etablerades av McFadden (1978) som ett villkor för en konsistent estimator för logitmodellen med ett urval alternativ. Om man för in uttrycket för logitmodellen kan man skriva den betingade sannolikheten på följande sätt:

$$\pi_n(i|D) = \frac{e^{(\mu V_{in} + \ln \pi_n(D|i))}}{\sum_{j \in D} e^{(\mu V_{jn} + \ln \pi_n(D|j))}}$$

McFadden (1978) visar att maximering av den betingade logaritmerade likelihoodfunktionen

$$\sum_{n=1}^N \ln \pi_n(i|D)$$

ger konsistenta parameterestimater. Det bör noteras, att urvalet inte behöver göras likformigt för de olika observationerna i datamaterialet.

Eftersom vi inte har med all information när vi endast tar ett sample ur alternativmängden blir parameterestimaten inte variansminimala, och alltså inte effektiva. Det är naturligtvis en intressant fråga, hur många alternativ man bör ta med för att få tillräcklig precision i skattningen.

Daly (1982) redovisar några tester på hur känslig parameterskattningen är för antalet alternativ. Testerna visar att variansen minskar något när man går från 8 till 20 alternativ, medan man inte kan säga att parameterestimaten är olika. I dessa test genomfördes en helt slumpmässig sampling. Även Algers, Colliander och Widlert (1987) redovisar ett test av olika antal alternativ, med liknande resultat. I fallet med ett stratifierat urval bör skillnaderna bli mindre, eftersom man från början dragit "effektivare" alternativ.

Under förutsättning att man kan beräkna uttrycket $\ln \pi_n(D|j)$ kan man alltså begränsa alternativmängden vid skattningen till att avse ett urval alternativ. Uppenbarligen hänger sättet att beräkna denna korrektionsterm samman med urvalsstrategin. Två fall är av särskilt intresse. Det ena fallet utgörs av ett rent slumpmässigt urval, där alla alternativ har samma urvalssannolikhet, medan det andra fallet består av ett stratifierat urval.

I det första fallet, när det således rör sig om ett urval där alla alternativ har samma urvalssannolikhet, kan det lätt visas att man inte behöver göra någon korrektion alls. Detta beror på att alla delmängder av alternativ har samma sannolikhet att bli valda, vilket innebär att alla alternativ får samma korrektionsfaktor. Denna kan då förkortas bort, och man erhåller det vanliga uttrycket för logitmodellen.

För att få en så effektiv skattning som möjligt kan man önska en annan sorts urval. Det kan exempelvis vara så, att man vill försäkra sig om att variationsvidden i de olika variablerna i modellen blir tillräckligt stor, eller att ett visst alternativ av något annat skäl bör finnas med i samplet, trots att det har en låg urvalssannolikhet.

Intuitivt kan man tycka, att urvalssannolikheten bör stå i proportion till valsannolikheten. En urvalsstrategi som ökar sannolikheten för alternativ med hög valsannolikhet - "viktiga" alternativ - kan då vara intressant. En sådan urvalsstrategi utgörs av ett stratifierat urval, där urvalssannolikheterna är lika inom respektive stratum, men där urvalssannolikheterna är större för de strata som innehåller "viktiga" alternativ.

Urvalskorrektionerna - som således ska adderas till nyttofunktionen vid estimeringen - kan i detta fall beräknas som minus logaritmen för urvalssannolikheten för ett alternativ i det stratum som alternativet tillhör. Detta kan visas på följande sätt (Ben-Akiva och Lerman 1985):

Låt oss anta, att alternativmängden J stratifieras på R icke överlappande strata, så att

$$\sum_{r=1}^R J_{r n} = J$$

där $J_{r n}$ är antalet alternativ i stratum r för observation n . Sättet att få med fler "viktiga" alternativ består i att genomföra stratifieringen på ett sådant sätt att vissa strata innehåller en högre andel "viktiga" alternativ, och att ge alternativen i dessa strata en högre urvals-sannolikhet. Inom varje stratum har alla alternativ samma urvalssannolikhet.

Bestäm sedan hur många alternativ som ska tas med från varje stratum. Låt $\tilde{J}_{r n}$ beteckna det av oss bestämda antalet alternativ i stratum $r = 1, \dots, R$, och låt $r(i)$ beteckna det stratum som alternativ i tillhör. Dra i varje stratum - utom det som det valda alternativet tillhör - ett slumpmässigt urval (utan återläggning) innehållande $\tilde{J}_{r n}$ alternativ. Dra i det stratum som det valda alternativet tillhör ett urval innehållande $\tilde{J}_{r(i) n} - 1$ alternativ, och lägg till det valda alternativet.

På detta sätt kan antalet alternativ bestämmas på förhand, även om definitionen av de olika strata varierar mellan olika observationer. Detta är en fördel när det gäller att hantera datamaterialet i det praktiska estimeringsarbetet. I praktiken kan antalet utvalda alternativ komma att variera ändå, eftersom $\tilde{J}_{r n}$ kan vara större än $J_{r n}$ för vissa strata r och observationer n . I dessa fall är naturligtvis urvalssannolikheten ett för alternativen $J_{r n}$. Ett exempel på ett sådant fall kan vara när ett stratum definierats som de områden som ligger inom ett visst avstånd från bostaden. Detta antal kan skilja sig mellan olika observationer, och för vissa observationer kan det visa sig att antalet till och med är lägre än det antal som vi önskat ha med.

Följande uttryck, vilket återspeglar antalet möjliga kombinationer av alternativ i de olika strata, kan ställas upp för korrektionstermen $\pi_n(D|i)$:

$$\pi_n(D|i) = \left(\frac{J_{r(i)n} - 1}{\tilde{J}_{r(i)n} - 1} \right)^{-1} \prod_{\substack{r=1 \\ r \neq r(i)}}^R \left(\frac{J_{r n}}{\tilde{J}_{r n}} \right)^{-1}, \quad i \in D$$

Korrektionstermen $\pi_n(D|i)$ beskriver sannolikheten att få ett visst urval alternativ D givet att alternativet i valts. Sannolikheten att få detta urval bör vara lika med produkten över alla strata av att få med de alternativ som tillhör D inom varje stratum. Sannolikheten att få med dessa alternativ inom varje stratum är omvänt proportionell mot antalet möjliga kombinationer av alternativ i stratumet. Antalet kombinationer ges av binomialkoefficienten:

$$\binom{J_{r n}}{\tilde{J}_{r n}} = \frac{J_{r n}!}{\tilde{J}_{r n}!(J_{r n} - \tilde{J}_{r n})!}$$

Följande uttryck, vilket återspeglar antalet möjliga kombinationer av alternativ i stratumet med det valda alternativet samt i övriga strata, kan således ställas upp för korrektionstermen $\pi_n(D|i)$:

$$\pi_n(D|i) = \left(\frac{J_{r(i)n} - 1}{\tilde{J}_{r(i)n} - 1} \right)^{-1} \prod_{\substack{r=1 \\ r \neq r(i)}}^R \left(\frac{J_{rn}}{\tilde{J}_{rn}} \right)^{-1}, \quad i \in D$$

Observera att

$$\begin{aligned} \left(\frac{J_{rn} - 1}{\tilde{J}_{rn} - 1} \right)^{-1} \left(\frac{J_{rn}}{\tilde{J}_{rn}} \right)^{-1} &= \left(\frac{(J_{rn} - 1)!}{(\tilde{J}_{rn} - 1)!(J_{rn} - 1 - \tilde{J}_{rn} + 1)!} \frac{J_{rn}}{\tilde{J}_{rn}} \right)^{-1} \\ &= \left(\frac{J_{rn}!}{\tilde{J}_{rn}!(J_{rn} - \tilde{J}_{rn})!} \right)^{-1} = \left(\frac{J_{rn}}{\tilde{J}_{rn}} \right)^{-1} \end{aligned}$$

Uttrycket för korrektionstermen kan därför skrivas om på följande sätt:

$$\pi_n(D|i) = \frac{J_{r(i)n}}{\tilde{J}_{r(i)n}} Q_n(D), \quad i \in D$$

där

$$Q_n(D) = \prod_{r=1}^R \left(\frac{J_{rn}}{\tilde{J}_{rn}} \right)^{-1}$$

Denna del av korrektionstermen är densamma för alla alternativ, och kan därför förkortas bort i uttrycket för den betingade likelihoodfunktionen ovan. Det räcker därför att för respektive alternativ beakta den alternativberoende delen, genom att addera termen

$$\ln \frac{J_{r(i)n}}{\tilde{J}_{r(i)n}}, \quad i = 1, \dots, J,$$

vilket är detsamma som att subtrahera den logaritmerade urvalssannolikheten q_{in} för respektive alternativ, där således:

$$q_{in} = \frac{\tilde{J}_{r(i)n}}{J_{r(i)n}}, i = 1, \dots, J,$$

Parametrarna estimeras alltså genom att maximera likelihoodfunktionen med de betingade sannolikheterna

$$\pi_n(i|D) = \frac{e^{(\mu V_{in} - \ln q_{in})}}{\sum_{j \in D} e^{(\mu V_{jn} - \ln q_{jn})}}$$

När man estimerar trädmodeller simultant, och destinationsvalet ligger på en lägre nivå, så kommer logsumman från den lägre nivån att innehålla en bias, genom att den valda destinationen alltid är med. Vi känner inte till någon referens, som strikt visar att urvalsförfarandet ger konsistenta estimat i detta fall. Intuitivt förefaller det dock som om det inte borde vara något problem - biasen bör vara lika stor för alla ovanförliggande alternativ, eftersom det valda området ingår i logsumman för alla dessa.

Vi har inom avhandlingsarbetet tillämpat tekniken med urval av alternativ på tre olika sätt. Det första sättet innebär att man definierar strata med utgångspunkt från det extra avståndet för att komma till en destination, givet att man reser mellan två fixa områden. Denna strategi, som utnyttjas i modeller för sekundära destinationer, innebär att sannolikheten för närliggande områdesalternativ ökar, vilket intuitivt är rimligt. Korrektionstermen blir i detta fall minus logaritmen för kvoten mellan antalet medtagna alternativ i stratomet och det totala antalet alternativ i samma stratum.

Det andra sättet att tillämpa stratifierat alternativurval består i en stratifiering som är en kombination av avståndsberoende och centralitet. I detta fall utgörs strata av områden i olika avståndsklasser samt områden i innerstaden respektive city. Strata får dock inte vara överlappande. Denna strategi tillämpas vid de flesta av våra modeller som innehåller destinationsalternativ. Även här blir korrektionstermen minus logaritmen för kvoten mellan antalet medtagna alternativ i stratomet och det totala antalet alternativ i samma stratum.

Det tredje sättet har att göra med kombinerade destinationsalternativ. Om en modell avser ett kombinerat destinationsval måste också alternativmängden utgöras av kombinerade alternativ. Stratifieringen görs här i två steg - det första steget utgörs av en stratifiering med samma princip som den föregående, och det andra steget utgörs av en kombination av dessa strata. Korrektionstermen består här av minus logaritmen av kvoten mellan antalet medtagna alternativ i den aktuella stratumkombinationen och produkten av antalet alternativ i de två strata som den aktuella stratumkombinationen utgörs av.

5.8 Kausalitet och korrelation

Det kan finnas anledning att också något beröra frågan om kausalitet och korrelation. Estimeringstekniken ger möjlighet att testa om det föreligger signifikanta samband mellan olika variabler. Förekomsten av ett sådant samband säger egentligen ingenting om kausaliteten i sambandet - denna fråga bör ju egentligen vara en del i det teoretiska underlaget för modellspecifikationen där beroende och oberoende variabler definieras. Det finns dock fall, när dessa frågor inte är särskilt entydiga.

Sådana fall uppkommer i samband med tjänste- och skolresemmodellerna, där det kan vara tveksamt vad det är för beslut som modelleras. Ett exempel är frekvensmodellen för tjänsteresor, som kan visa sig ha en signifikant parameter för tillgänglighet. Det kan då vara oklart om antalet tjänsteresor beror på en god tillgänglighet, eller om en god tillgänglighet har inneburit att företag med många tjänsteresor har lokaliserat sig där. En utförligare diskussion av detta problem förs i anslutning till redovisningen av dessa modeller.

5.9 Tvärsnittsdata

Det kan diskuteras om tvärsnittsdata eller tidsseriedata ska ligga till grund för att testa våra hypoteser. Tvärsnittsdata är det helt dominerande underlaget för denna typ av modeller. Samtidigt finns vissa problem förknippade med ett sådant dataunderlag.

Modeller som skattas på tvärsnittsdata baseras på antagandet att det resmönster som modellen beskriver beror av de variabler som registreras vid samma tidpunkt. Huruvida detta antagande är korrekt eller ej beror på i vilken utsträckning det förekommer någon fördröjning mellan förändringar i dessa variabler och eventuella förändringar i resmönstret (under förutsättning att den aktuella variabeln påverkar resmönstret).

Även om det förekommer påtagliga fördröjningseffekter kan ett tvärsnittsmaterial vara relevant som underlag för modellstudier. För att detta ska vara fallet krävs dock att de variabler som påverkar resmönstret inte förändrats i nämnvärd utsträckning under en tillräckligt lång tid.

Utvecklingen i Stockholm hade i några avseenden en svacka under den tidigare hälften av 1980-talet. Detta innebar att inkomstutvecklingen var ungefär noll procent mellan 1980 och 1985 (Fridén et al 1985), och att trafiken över innerstadssnittet (Stockholms Gatukontor 1985) var ungefär lika stor 1985 som 1976. Även om det kan ha funnits skillnader i förväntningar under denna period, så har levnadsvillkoren varit förhållandevis lika. Det finns därför anledning att anta att den typ av problem som nämnts ovan bör ha varit måttliga.

Ett alternativ till tvärsnittsdata är data som beskriver just förändringar. Sådana material kan erhållas i s.k. panelundersökningar, där samma urval intervjuas vid olika tillfällen. Även sådana studier är förknippade med problem av olika slag. Ett av de viktigaste problemen är att vidmakthålla panelen över tiden.

Ett annat sätt att erhålla sådana data är genom retrospektiva frågor i tvärsnittsstudier. Detta är dock svårt när det gäller information om alla dimensioner i resmönstret. En förutsättning när det gäller att kunna utvinna information ur paneldata är naturligtvis också att det förekommer nämnvärda förändringar i de oberoende variabler som påverkar resmönstret.

Ett ytterligare problem när det gäller tidsseriedata är att det tar längre tid att genomföra undersökningen. Detta är särskilt problematiskt när det gäller tillämpningar utanför rena forskningsmiljöer, där resultaten ska komma till praktisk användning så snart som möjligt. En tvärsnittsstudie valdes därför som grund för modellarbetet.

6. Datamaterial

6.1 Inledning

För att skatta ett modellsystem med den omfattning som beskrivs i kapitel 3 krävs betydande mängder indata. I detta projekt används som diskuterats i avsnitt 3.2 utslutande uppgifter om hur individer och hushåll faktiskt betett sig när de stått inför en valsituation (det vill säga uppgifter om "Revealed Preferences"). Den viktigaste delen utgörs därför av en *resvaneundersökning* som ger information om hur ett urval hushåll valt att resa under en viss mättag. Utöver denna uppgift krävs även uppgifter om vad de kunde gjort istället, det vill säga vilka alternativ de stod inför. Delvis får vi denna information från resvaneundersökningen (där vi kan se om hushållet har tillgång till bil, om personen ifråga har körkort, etc.). En ytterligare viktig källa till information om de olika alternativen utgörs av *attraktivitetsregister* som ger information om de möjligheter som finns att utföra olika aktiviteter i olika målområden. Tillgängligheten till olika målområden ges i sin tur av *trafiksystemdata* som huvudsakligen hämtas från kodade trafiknät.

6.2 Resvaneundersökning

Inom ramen för projektet genomfördes två olika undersökningar. Dels genomfördes en hushållsundersökning vars syfte var att ge ett underlag för att utveckla projektets modeller, dels genomfördes även en antalsmässigt omfattande individundersökning (netto ca. 30 000 intervjuer) där i gengäld mängden insamlad information per intervju begränsades starkt. Individundersökningen genomfördes främst för att ge möjlighet att redovisa geografiskt mer nedbrutna deskriptiva resultat och för att ge resmatriser som bland annat används som en informationskälla bland flera andra för att skapa de nulägesmatriser som prognossystemet utnyttjar.

I detta avsnitt behandlas uppläggning och genomförande av hushållsundersökningen. En mer utförlig redovisning återfinns i en särskild teknisk rapport (Rosenqvist 1991). Individundersökningen berörs inte vidare här. Även genomförandet av denna dokumenteras i en teknisk rapport (Mouwitz 1992).

Planering och uppläggning

Planeringen av resvaneundersökningen kom tidigt att präglas av behovet av uppgifter om hela hushållets resande och inte minst av att alla i hushållet måste beskriva *samma* dags resande. Kravet på att samma dag beskrivs orsakas av att modellerna behandlar substitutionsmöjligheter och samband mellan olika hushållsmedlemmar, vilka naturligtvis påverkas av de resor som utförs.

Tidigare resvaneundersökningar i landet har oftast varit individorienterade (Lundberg 1982a, SCB 1984, Mouwitz 1983). I de fall de gällt hela hushåll (TU 71 1974) fanns inte det absoluta kravet på att alla medlemmar skulle beskriva samma dag. Mycket av arbetet

kom därför att kretsa kring möjligheterna att utforma undersökningen så att partiella bortfall av enskilda hushållsmedlemmar kunde undvikas.

Eftersom den planerade datainsamlingen var dyrbar och projektet stöddes med forskningsmedel gavs både anledning och möjlighet att systematiskt prova olika utformningar och att genomföra en datainsamling av hög kvalitet. Kontakter togs därför med svenska och internationella organ med erfarenhet inom området. Stellan Lundberg hade nyligen med stöd av Bygghörsningsrådet genomfört en resvaneundersökning i Jönköping med avsikten att utveckla en förebild för kommande undersökningar (Lundberg 1982b och c). Uppläggningsen av Stockholmsundersökningen diskuterades därför med honom. Samarbete inleddes även med företaget Socialdata i München som utvecklat en tysk nationell resvaneundersökning som genomförts ett flertal gånger med mycket stora urval (Schwertner et al 1983, Brög et al 1983).

Den tyska undersökningen kännetecknas av en långt gången omsorg om utformningen av formulären med en särskild resdagbok och en genomtänkt grafisk utformning som gör följdfrågor och hopp i formulären begripliga utan omfattande instruktioner. Undersökningen genomfördes helt per post. I samband med utformningen av den tyska undersökningen utfördes omfattande tester av olika utformningsaspekter. Resultaten av testerna kunde direkt komma till nytta vid planeringen av Stockholmsundersökningen. Socialdata medverkade därför under planeringsperioden och under provundersökningarna, ända fram till dess att huvudundersökningen startade.

Provundersökning 1

Den första provundersökningen genomfördes hösten 1983. Totalt 300 hushåll i tre områden i Stockholm intervjuades. Varje hushåll tilldelades en viss mätvecka. Samtliga förflyttningar under denna vecka registrerades. Hushållen fick först ett informationsbrev per post. Därefter kontaktades de per telefon och ett antal inledande frågor ställdes. Resdagböcker och frågeformulär sändes därefter ut i lagom tid före första mätdagen. Mitt i mätveckan togs en ytterligare telefonkontakt för att motivera hushållet. Hushållet skickade slutligen in materialet per post.

Huvudsyftet med det första provet var dels att pröva möjligheterna att intervjua hela hushåll och att få dem att beskriva en hel veckas resande, dels att pröva formulärutformning och att få underlag till kostnadsberäkningar.

Provet visade att det gick mycket bra att få alla hushållsmedlemmar att beskriva samma period. Det visade sig även att den använda dagboksmetoden registrerade fler förflyttningar än tidigare använda telefonintervjuer. En rad synpunkter på formulärens utformning och på uppläggningsen av undersökningen inkom under provet.

Den första provundersökningen redovisas i Planeringsrapport 3 (1985).

Provundersökning 2

Under hösten 1984 genomfördes en andra provundersökning. Totalt intervjuades 580 hushåll. Intervjumaterialet hade inför detta prov omarbetats fullständigt. Den använda resdagboken hade nu utformats utifrån de tyska erfarenheterna. Tre olika varianter med en, tre eller sju dagar i resdagboken prövades.

Centrala frågeställningar i det andra provet var hur den nya utformningen av materialet fungerade, hur antalet dagar i dagboken påverkade resultaten och hur väl hushållen förmodade fylla i dagböckerna. Den sista aspekten behandlades genom att intervjuare oanmälda besökte vissa hushåll under mätperioden. För övrigt användes samma principiella uppläggning som i det första provet.

Resultaten visade att intervjuarnas besök höjde rapporteringsgraden i resdagboken. Antalet dagar i resdagboken hade visserligen inte någon avgörande betydelse för svarsbenägenheten, men det fanns tecken på att det framför allt var de regelbundna resorna som verkligen rapporterades alla dagar.

Provundersökning 3

Den tredje provundersökningen genomfördes under försommaren 1985. Huvudsyftet var att pröva möjligheterna att utföra intervjuerna helt per post med liknande uppläggning som den tyska. Provet kom framför allt att fungera som ett prov inför individundersökningen. Resultaten visade att det var möjligt att genomföra en datainsamling per post, under förutsättning att formulärets omfattning begränsades.

Resultaten från provundersökning 2 och 3 redovisas i Planeringsrapport 4 (1985).

Huvudundersökningen

Uppläggning och genomförande

Efter erfarenheterna från de tre provundersökningarna beslöts att använda en kombination av i förväg utskickat material och personliga kontakter med en intervjuare för att få så hög kvalitet som möjligt. Eftersom data för en hel vecka ledde fram till komplexa modeller som var svåra att passa in i den modellstruktur som valts, och eftersom sjudagarsdagboken både var svårare och mer tidskrävande att samla in, beslöts att varje intervjuperson enbart skulle beskriva en mätdag.

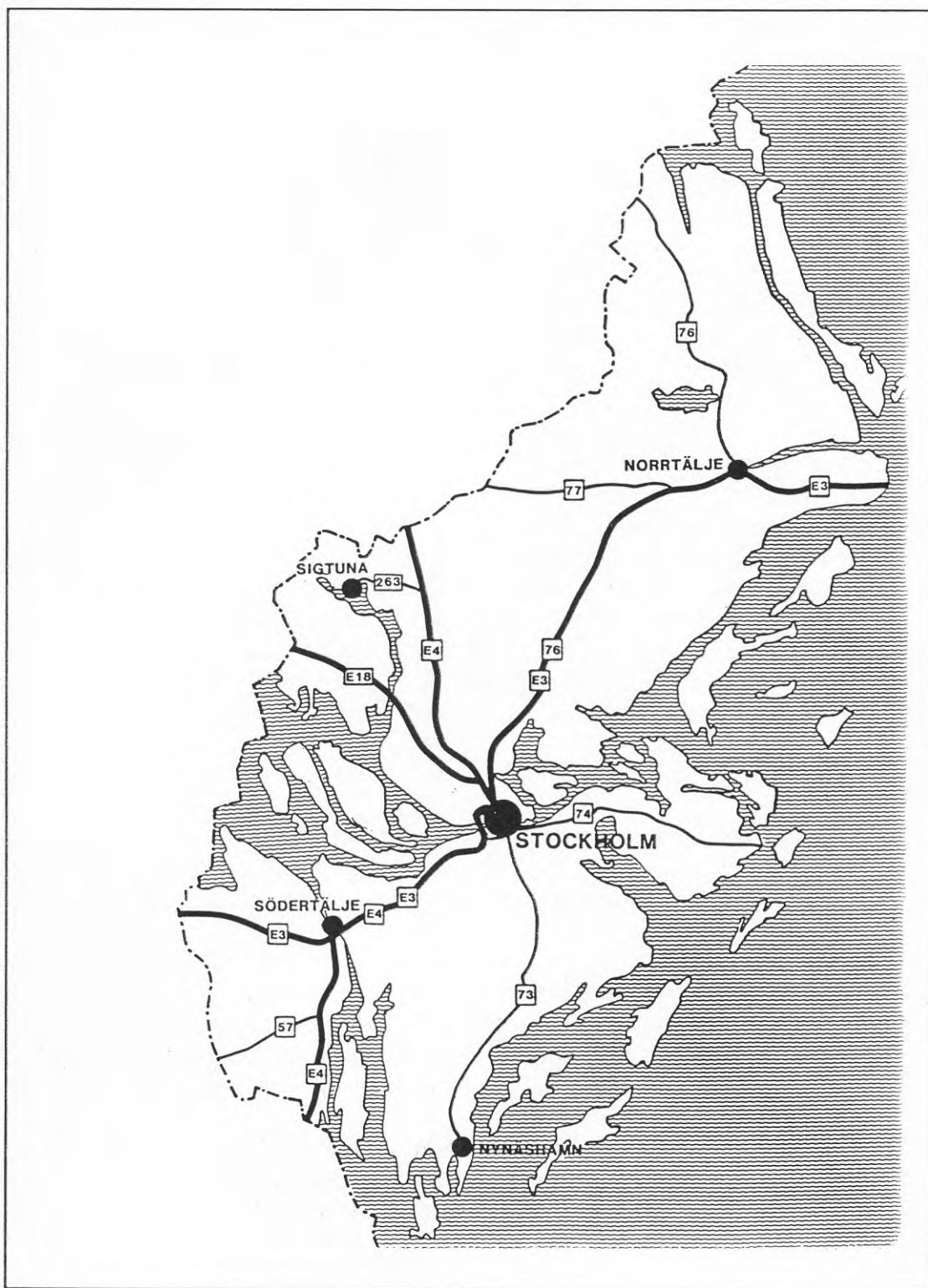
Undersökningen fick därför följande uppläggning:

- Två veckor före mätdagen fick hushållet ett introduktionsbrev
- En vecka före mätdagen tog intervjuaren kontakt per telefon. Ett antal frågor om hushållets sammansättning ställdes. Intervjuaren bokade tid för hemintervju.
- Efter telefonkontakten skickade intervjuaren en mapp till hushållet med resdagböcker till samtliga personer som var 12 år eller äldre.
- På den avtalade tiden (normalt dagen efter mätdagen) besökte intervjuaren hushållet, ställde frågorna i ett hushållsformulär, gick igenom alla dagböcker och lät alla hushållsmedlemmar fylla i ett särskilt individformulär.

Fältarbetet genomfördes av 27 intervjuare från Statistiska Centralbyrån, SCB. Intervjuerna pågick under ett helt år för att få representativa resultat. Intervjuerna började i mars 1986 och slutade i mars 1987. Uppehåll gjordes under sommaren och julen. Bortfallsuppföljningar utfördes under tre olika perioder.

Urval

Målpopulationen för undersökningen var personer bosatta i Stockholms län i åldern 12-74 år. Undersökningsområdet framgår av figur 6.1. Urvalet drogs ur SCBs register över totalbefolkningen.



Figur 6.1 Undersökningsområde

Totalt drogs 3 750 personer. Urvalet stratifierades efter kommun. Urvalet var proportionellt mot befolkningen i den aktuella åldersgruppen i kommunen, men med en undre gräns vid 90 personer per kommun. Hushållen bildades genom att de utvalda personerna under intervjun fick uppge vilka personer som tillhörde det egna hushållet. Inom varje kommun spreds mätdagarna jämnt över hela undersökningsperioden.

Svarsprocent

En målsättning för intervjuarbetet var att nå upp till 80% genomförda intervjuer. Samtidigt med den planerade undersökningsstarten utbröt en massmediadebatt om integritet i samband med undersökningar. Debatten orsakades av att förekomsten av en tidigare ej offentliggjord undersökning, där information insamlades genom samkörning av olika register, avslöjades (den så kallade "Metropolitundersökningen"). Det krävdes därför stora extra insatser för att nå upp till det satta målet. Det slutliga resultatet framgår av tabell 6.1. Den slutliga svarsprocenten blev således 79,5.

Tabell 6.1 Resultat av fältarbete

	Antal	Procent
Bruttourval	3 750	
Övertäckning	98	
Nettourval	3 652	100,0
Intervjuade hushåll	2 904	79,5
Förhindrad medverkan	35	1,0
Ej anträffad	168	4,6
Vägran	545	14,9

Gruppen "övertäckning" utgörs av personer som avlidit eller flyttat ut ur länet. "Förhindrad medverkan" innebär att intervjupersonen haft något fysiskt eller psykiskt hinder för att delta i undersökningen.

Kvalitet

För att få representativa resultat är det viktigt att undvika att mätdagar flyttas. Det finns en risk för att hushåll som är bortresta, på sjukhus, eller som av annan anledning inte är hemma under den ursprungliga mätdagen, blir kontaktade senare och då beskriver ett resande som inte alls är representativt för den ursprungliga mätdagen. Denna tendens leder till att det inomregionala resandet överskattas. Totalt 8,5% av hushållen fick sin mätdag flyttad för att intervjuaren inte kunde få kontakt med hushållet i tid. När mätdagar flyttats kontrollerade intervjuaren om hushållet varit bortrest utanför länet eller på sjukhus under den ursprungliga mätdagen. Om så var fallet behölls den ursprungliga mätdagen.

Vi har tidigare diskuterat att problemen med att få samtliga hushållsmedlemmar att beskriva samma dag var en av de frågor som ägnades störst uppmärksamhet under planeringen av datainsamlingen. Tabell 6.2 visar hur väl detta lyckades under huvudundersökningen.

Tabell 6.2 Partiella bortfall i resvaneundersökningen

	Antal	Procent
Hushåll med partiellt bortfall av individformulär	7	0,2
Hushåll med partiellt bortfall av resdagbok	7	0,2
Hushåll med partiellt bortfall av samtliga formulär för någon individ	20	0,7
<hr/>		
Hushåll med någon form av partiellt bortfall av individ	34	1,1

Totalt är det enbart 1,1% av de intervjuade hushållen som saknar ett eller flera formulär för någon person i hushållet. Med den valda uppläggningsen har det således praktiskt taget alltid varit möjligt att intervjua samtliga personer trots kravet att samma dag skulle beskrivas.

Partiella bortfall på enstaka frågor i undersökningen förekommer främst för inkomst där ca 10% vägrat lämna uppgift (eller angett att de inte vet inkomstens storlek). Näst högst bortfall erhöles för utbildningsfrågan som 2% vägrade besvara.

Granskning, kodning och logiska kontroller

Omedelbart när formulären inkommit till undersökningsledningen granskades de. Saknade uppgifter samlades in per telefon. 14% av hushållen kontaktades på detta sätt i efterhand för att lämna ytterligare uppgifter.

Undersökningen kodades i ett särskilt system för datorstödd kodning som delvis utvecklats just för detta projekt. Systemet innebär att kodaren sitter vid en terminal och skriver in intervjuformulärets uppgifter. Systemet gör direkt vissa logiska kontroller och tillåter dessutom enbart svar inom tillåtna intervall. I systemet finns även ett adressregister som gör att de adresser som intervjupersonerna uppgett för förflyttningarnas start- och målpunkter direkt kan omvandlas till koordinater och områdeskoder. Adressregistret hämtades ursprungligen från Centrala Fastighetsdataregistret (CFD-registret) och från länsstyrelsens adressregister. Registret kompletterades därefter med "kända platser". Successivt lades dessutom adresser som ej kunde matchas in i registret. Totalt kunde mer än 95% av de uppgivna adresserna ges både koordinater och områdeskoder vilket är ett anmärkningsvärt gott resultat.

Under kodningsperioden kontrollkodades 10-15% av alla intervjuer. Kontrollerna var särskilt intensiva när nya kodare började sitt arbete.

Efter att materialet kodats har även ett antal logiska kontroller utförts. Kontrollerna har gällt åldern hos personer som uppgett att de kört bil som förare, åldern på dem som lämnat individformulär (minst 12 år), att tiden för ankomst till målet är senare än avresetidpunkten, ankomsttider för inköpsresor, etc.

Sammanfattning

Den genomförda undersökningen föregicks av osedvanligt omfattande provundersökningar. Ambitionsnivån under insamlingsfasen var också hög, både vad gäller svarsprocent och kontroll av materialets fullständighet. Ett omfattande arbete lades ner av undersökningsledningen under hela undersökningsåret. Resultatet har blivit ett datamaterial av mycket hög kvalitet. Det interna bortfallet är litet, både när det gäller individer och svar på de olika frågorna. Jämförelser med andra undersökningar tyder också på att en hög andel av förflyttningarna verkligen rapporterats (Rosenqvist 1991).

6.3 Bearbetning av resvanedata

Huvudresa och huvudärende

Olika resvaneundersökningar arbetar med olika definitioner av vad som utgör en resa. De flesta använder ett eller flera av följande resbegrepp:

- reselement
- delresa
- huvudresa

För att förklara dessa begrepp måste vi även införa begreppet "besöksställe". Med besöksställe menar vi varje plats där en person gjort ett uppehåll för att uträtta ett ärende (inklusive bostaden).

Med ett *reselement* menar vi den del av en förflyttning mellan två besöksställen som skett med samma färdmedel. En förflyttning där man gått till bussen, åkt buss och därefter gått från hållplatsen till besöksstället består således av tre reselement.

En *delresa* är en förflyttning mellan två besöksställen och en *huvudresa* är en resa som består av en eller flera delresor och där hänsyn tas till sambandet mellan de olika delresorna i reskedjan.

I resvaneundersökningen samlas uppgifterna om intervjupersonernas resande in med särskilda resdagböcker. Dagböckerna är utformade så att förflyttningar mellan platser där man gjort uppehåll för att uträtta ett ärende registreras som en enhet (som en separat kolumn i dagboken). Resdagboken samlar således in *delresor*.

Undersökningen gör det inte generellt möjligt att identifiera reselement. Det är visserligen tillåtet att ange flera färd sätt för en förflyttning, men det finns ingen garanti för att alla element uppgivits.

När materialet ska användas för att skatta trafikmodeller är det nödvändigt att ta hänsyn till att delresor i en reskedja har ett samband (t.ex. att det finns ett beroende mellan valet av färdmedel till och från arbetet). Om en reskedja innehåller mer än ett besöksställe identifieras därför ett enda huvudbesöksställe och ett enda huvudärende för resan.

Vi definierar en huvudresa som ett antal sammanhängande delresor som både börjar och slutar i någon av följande brytpunkter:

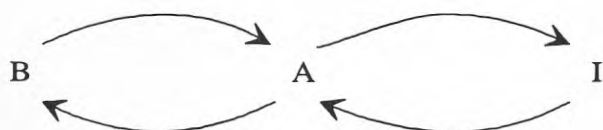
- den egna bostaden
- den egna fritidsbostaden
- den egna arbetsplatsen/skolan

Delresorna i en huvudresa bildar således en sammanhängande rundtur som börjar och slutar i samma punkt. Om flera av brytpunkterna ingår i samma huvudresa har egen bostad högst prioritet när startpunkten ska bestämmas. Näst högst prioritet har fritidsbostaden.

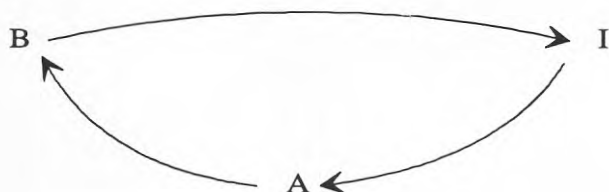
Några exempel illustrerar principen. Vi använder följande beteckningar:

- B = bostaden
- A = arbetsplatsen
- I = ett inköpsställe
- S = en servicepunkt
- R = en rekreationspunkt

Vi betraktar först ett exempel där en person rest från bostaden till arbetet, därefter besökt en butik under lunchen och återvänt till arbetsplatsen, samt slutligen åkt hem igen efter arbetets slut:



Dessa fyra delresor bildar två huvudresor. Den första huvudresan sker från bostad till arbete och åter till bostad. Den andra sker från arbete till inköp och åter till arbete.



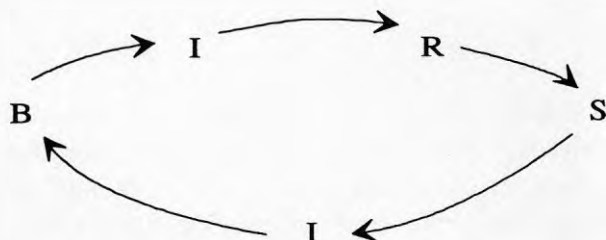
De tre delresorna i det andra exemplet utgör en enda huvudresa. Huvudresan har startpunkt i bostaden.

När en huvudresa innehåller mer än ett besöksställe måste ett huvudärende för resan bestämmas. Huvudregeln är att alla ärenden i delresorna som bildar huvudresan rangordnas med avseende på vistelsetiden i besöksstället. Vistelsetiden är skillnaden mellan ankomsttiden till och avresetiden från en målpunkt. Det har visats (Antoinisse, Daly och Gunn 1986) att denna klassificering väl överensstämmer med trafikanternas uppfattning om viktigaste ärende. Från denna regel finns tre undantag. Det första är att om någon målpunkt är arbete eller skola är huvudärendet arbete respektive skola. Det andra undantaget är tjänstresor, som alltid är högst i rang om huvudresan inte innehåller någon arbets/skolresa. Det

tredje undantaget är ärendet hämta och lämna andra personer, som alltid har lägsta rang. Rangordningen är alltså följande:

1. Arbete/skola
2. Tjänste
3. Övriga ärenden (utom hämta och lämna) efter vistelsetid
4. Hämta och lämna

Det tredje exemplet visar en resa där en rad olika ärenden utförts:



Reskedjan består av två delresor med inköpsärende, en delresa med rekreationsärende och en delresa med serviceärende. De fem delresorna bildar en enda huvudresa. Den målpunkt där vistelsetiden är längst utgör huvudbesöksställe och ger huvudresans ärende.

En alternativ beteckning för huvudresor som används i rapporten är "rundtur" eller enbart "tur".

Omvandlingen av resdagbokens delresor till analysernas turbegrepp innebär en förenkling av verkligheten som medför att vissa resor försvinner. Modellerna beskriver enbart vissa aspekter av kedjeresorna, övriga besök som utförs i kedjorna elimineras. Detta innebär att resultaten från modellerna kommer att underskatta trafikarbetet något. Detta korrigeras i prognosprogrammen med korrektionsfaktorer.

Hushållsturer

I resdagböckerna samlas information om *individens* resande. Flertalet av de modeller som utvecklas behandlar dock *hushållets* resande. Vid intervjuerna användes även ett separat formulär som visade vilka övriga hushållsmedlemmar som rest tillsammans med intervjupersonen under varje delresa. Med hjälp av denna information kan *hushållsturer* bildas. Vår definition av en hushållstur är att de olika personerna som rest tillsammans ska ha utträttat reskedjans huvudärende i samma målområde och att minst en av delresorna både till och från huvudärendet ska ha utförts gemensamt. Dessutom användes ett krav att starttidpunkterna för resorna inte fick skilja sig mer än 15 minuter. Individerna behöver således inte ha färdats gemensamt på alla delar av resan. Individernas delresor kan ha rapporterats olika i dagboken, kravet är bara att de uppgivit att de åkt tillsammans och att resorna startat vid ungefär samma tidpunkt. Definitionen motiveras av att ett antal problemfall ska kunna hanteras, t.ex.:

- en resa där man samåker en del av vägen på väg till olika resmål är inte en gemensam hushållsresa på huvudresenivån
- en resa där man åker tillsammans till ett centrum och där uträttar olika ärenden och sedan åker hem är en gemensam hushållsresa (trots de olika ärendena)
- en resa där man åker gemensamt, men där den ena personen uträttar ett ärende på vägen (t.ex. handlar i samband med ett färdmedelsbyte) är en gemensam hushållsresa
- helt identiska resor kan rapporteras olika i resdagböckerna (t.ex. skillnader i tider för avresa eller ankomst)

Huvudfärdsätt

Uppläggningsen av resdagboken innebär att samma person kan ha uppgivit flera olika färdmedel för samma delresa. Under en reskedja kan naturligtvis även olika färdmedel ha använts för de olika delresorna (t.ex. kollektivtrafik till centrum, sedan till fots mellan de enskilda målpunkterna i centrum). För varje tur har därför ett huvudfärdmedel bestämts enligt följande regler:

Tabell 6.3 Klassning till huvudfärdmedel

Huvudfärd- medel	Registrerat färdmedel i undersökningen:							taxi	övrigt
	bil som förare	bil som pass.	koll.	cykel	gång	färd- tjänst			
1 förare	x			(x)	(x)				
2 passagerare		x		(x)	(x)				
3 infartsparkering	x		x	(x)	(x)				
4 kiss-and-ride		x	x	(x)	(x)				
5 kollektivt			x	(x)	(x)				
6 cykel				x	(x)				
7 färdtjänst					(x)	x			
8 taxi					(x)		x		
9 gång					x				
10 övrigt								x	
11 övriga kombinationer									

x = färdmedlet ska ha använts, (x) = färdmedlet kan ha använts

Vid klassningen till huvudfärdmedel för hushållsturen används samma hierarki som i tabellen.

6.4 Trafiksystemdata

Områdesindelning

Samtliga modeller i systemet använder sig av en områdesindelning där Stockholms län delas in i ca 850 delområden. Detta innebär att de nät som kodats mäter tider och avstånd mellan dessa områden och att alla övriga data gäller denna områdesindelning.

Nätdata

För påkodningen av restider till resvaneundersökningen krävs tillgång till nätuppgifter för respektive färdstätt och tidsperiod. Indelningen för biltrafiken beaktar att det råder trängsel under rusningstrafiken, att perioden mellan morgon- och eftermiddagsrusningen innebär en viss trängsel, samt att trafiken under den övriga delen av dygnet knappast har några framkomlighetsproblem. Efter en genomgång av trafikräknedata beslöts följande indelning för biltrafiknäten:

Tabell 6.4 Tidsperioder för bilnät

Tidsperiod	Nätbelastning
Vardagar	
1. Morgonrusning resan påbörjad 06:45-08:45	Maxtimmatris enligt regionplan 1985 kalibrerad mot infartsräkningen 1985
2. Eftermiddagsrusning resan påbörjad 16:00-18:00	Transponerad morgonmatris
3. Mellantidsmatris resan påbörjad 08:45-16:00 samt 18:00-19:00	Summan av de båda föregående matriserna, multiplicerad med 0,35
4. Lågtrafikmatris resan påbörjad 00:00-06:45 samt 19:00-00:00	Fritt flöde
Lördagar	
5. Mellantidsmatris resan påbörjad 10:00-16:00	Samma som 3.
6. Lågtrafikmatris resan påbörjad 00:00-10:00 samt 16:00-00:00	Fritt flöde
Söndagar	
7. Lågtrafikmatris	Fritt flöde

För varje tidsperiod finns trafiknätsuppgifter för bilrestid och bilavstånd mellan varje områdesrelation.

För kollektivtrafiknäten används följande indelning i tidsperioder:

Tabell 6.5 Tidsperioder för kollektivtrafiknät

Tidsperiod	Gäller from, - tom.
Vardagar	
1. Morgonrusning	06-08
2. Eftermiddagsrusning	15-17
3. Mellantrafik	09-14
4. Tidig kvällstrafik	18-19
5. Mellankvällstrafik	20-21
6. Övrig trafik	22-05
Lördagar	
7. Högtrafik	10-15
8. Övrig trafik	16-09
Söndagar	
9. Högtrafik	11-14
10. Övrig trafik	15-10

För varje tidsperiod finns separata matriser för restid i fordonet, gångtid till och från fordonet, väntetid, bytestid samt antalet byten.

Den viktigaste orsaken till skillnader mellan tidsperioder för kollektivtrafiken är olika tur-tätheter under olika tider. Det förekommer dock även linjer som trafikeras enbart vissa tider.

Restids- och reslängdsmatriserna genererades med hjälp av tillgänglig programvara för bil- respektive kollektivtrafik (EMME/2, för en beskrivning av algoritmer se EMME/2 User Manual). Kodningen genomfördes på ett sådant sätt, att de förhållanden som rådde under undersökningsåret efterliknades så långt möjligt.

Genereringen av kollektivrestiderna förutsätter att man kan ange vilka vikter som ska åter-spegla värderingen av de olika restidskomponenterna. Detta innebär emellertid en viss risk för cirkelbevis, vilket kan belysas med ett exempel. Låt oss anta, att vi studerar en popula-tion, som värderar alla komponenter precis lika. Om vi tror oss veta, att värderingen av

väntetider är betydligt högre och därför genererar restidsmatriser med en mycket hög väntetidsvärdering, så kommer resvägar med lång väntetid att utnyttjas mindre ofta, vilket resulterar i restidskomponenter där väntetiden är relativt kort.

Effekten av detta blir högre parameterestimater som kompenserar de korta väntetiderna, och lägre parametrar för övriga variabler som kompenserar de ökade värdena för dessa. På detta sätt har vi - falskeligen - bevisat att väntetider har högre vikt.

Vi har dock inte gjort någon känslighetsanalys för dessa antaganden, vilket egentligen vore önskvärt. Några dramatiska effekter kan man knappast förvänta sig, eftersom ruttvalet inte är särskilt känsligt för förändringar i dessa antaganden i de relationer där det inte finns så många intressanta alternativ. Genom den stora mängden nätdata är känslighetsanalysen emellertid en ganska resurskrävande uppgift.

Vi har i vår studie använt vikterna 1 för åktid, och 2 för övriga komponenter. Detta motsvarar de värden dessa parametrar normalt får i denna typ av studier, vilket också bekräftats genom andra studier, där nätdata inte genererats på detta sätt (exempelvis Stated Preference-studier, se Widlert 1990). De resulterande vikterna (som sammanfattas i avsnitt 8.7) ligger också i rimlig närhet av de antagna - dock med en tendens till lägre vikter för väntetid och bytestid och högre för gångtid till hållplats.

Parkeringsdata

I vissa delar av länet har parkeringskostnaderna stor betydelse för färdmedelsvalet. Främst gäller detta Stockholms innerstad, men även i vissa andra områden har parkeringskostnaden betydelse. En komplikation utgörs av att kostnaderna inte är likformiga inom de olika områden som studeras, utan ofta varierar med resärendet. I vissa områden är t.ex. parkeringen vid centrumbildningen (som besöks vid inköpsresor) avgiftsbelagd medan parkeringen är gratis i bostadsområdena (dvs. gratis vid besöksresor). I andra områden är förhållandet det motsatta. Den avgiftsbelagda tiden (både över veckan och dygnet) varierar dessutom mellan olika områden.

Vi behöver således ett parkeringsregister där avgifter för parkering registreras på en ännu finare områdesnivå än den 850-områdesnivå som annars används. Registret måste dessutom innehålla uppgifter om avgiftsbelagd tid olika dagar i veckan. Eftersom sådana uppgifter saknades genomfördes en särskild enkät till trafikingenjörerna i länets kommuner. Enkäten följdes upp per telefon av trafikkontorets kommunbevakare för respektive kommun. Trots att ett omfattande arbete lades ner visade det sig omöjligt att få in uppgifter från alla kommuner. Dessutom kände trafikingenjörerna inte till avgiftsnivåer och avgiftsbelagda tider i alla bostadsområden. Vi tvingades därför att själva utföra en inventering i de områden där trafikkontorets kommunbevakare bedömde att avgifter kunde förekomma.

Med hjälp av de insamlade uppgifterna kunde därefter det önskade parkeringsregistret upprättas. Eftersom registret avser 850-områdesnivån registrerades separata uppgifter för besöksresor (till bostadsområden) och för övriga resärenden. Det är dock värt att lägga på minnet att det är en problematisk uppgift att beräkna relevanta parkeringskostnader. En stor del av dem som parkerar gör detta i strid mot gällande bestämmelser. Detta innebär både att otillåtna platser utnyttjas och att avgifter ej erläggs. För många bilister är det där-

för inte i första hand parkeringsavgiften som är ett mått på kostnaden, utan snarare risken för att få en parkeringsbot och dennas storlek.

I registret antas att alla platser inom samma 850-område har samma kostnad för ett visst resärende. Detta är en förenkling eftersom det i verkligheten inte alls är ovanligt att både avgiftsbelagda och ej avgiftsbelagda platser finns inom samma område. Slutsatsen blir att parkeringsregistrets uppgifter enbart kan utgöra en förhållandevis grov indikation på kostnaden för att parkera i olika områden.

För arbetsresor frågar vi i resvaneundersökningen om intervjupersonens kostnad för att parkera i det valda målområdet - oavsett om intervjupersonen åker bil eller inte. För de alternativa ej valda målområdena används genomsnittliga faktiska parkeringskostnader för de valda destinationerna. För samtliga övriga resärenden används parkeringsregistret. Hän-syn tas både till avgiftsnivån i målpunkten, parkeringens varaktighet och till om resärendet är besök eller ej.

6.5 Områdesdata

För att beskriva nyttan av ett område krävs egentligen information om såväl antalet elementaralternativ som den genomsnittliga nyttan för elementaralternativen (och gärna också variansen i denna). När det gäller antalet elementaralternativ - storleksdata - finns en hel del information tillgänglig, vilket redovisas nedan. Beträffande data som beskriver nyttan är läget sämre. Den typ av information som skulle vara intressant i sammanhanget utgörs av prisnivå, kvalitet, grad av service etc. Någon sådan information har vi inte haft tillgång till, utan vi har i stället fått ansätta dummyvariabler för några olika områden eller områdestyper där vi haft anledning att tro att den genomsnittliga nyttan i området kunnat avvika. Exempel på detta är dummyvariabler för områden med stormarknader och dummyvariabler för cityområdet.

I resvaneundersökningen får vi veta vilka målpunkter intervjupersonerna har besökt. Vid analyserna behöver vi också veta vilka andra målpunkter de kunde besökt istället, samt hur stora dessa olika målpunkter är (hur många elementaralternativ de innehåller).

Under projektets planeringsfas lades ett omfattande arbete ner på att undersöka hur ett sådant register bäst skulle kunna byggas upp. Samtliga befintliga register inventerades. Särskilt för butiker finns ett antal olika register med olika innehåll och syfte. Vårt önskemål var att få med så många olika storleksmått som möjligt i registret.

För inköpsresor undersöktes möjligheterna att få med följande mått:

- antal anställda
- försäljningsytor
- omsättning

De register som inventerades var SCBs basregister (Centrala Företagsregistret), Dagligvaruleverantörernas Förbunds register (DLF-registret), Statens Pris- och Konkurrensverks register (SPK-registret), KFs register över kooperativa butiker, Stockholms kommuns serviceregister, samt ett antal andra mindre kommunala register. Olika register visade sig under arbetets gång vara behäftade med olika brister. De flesta av registren innehåller bara

vissa av de önskade uppgifterna och de flesta täcker antingen bara vissa butikstyper eller bara viss del av länet.

Avsikten var från början att SCBs basregister skulle utgöra stommen i vårt register. Basregistret hade nyligen enligt SCBs uppgifter förbättrats avsevärt i och med att det skulle börja användas som en del i den registerbaserade Folk- och bostadsräkningen. Vi försökte därför aggregera Basregistret till områdesnivå. Detta arbete visade till att börja med att 30-40% av registrets företag utanför Stockholms kommun saknade korrekta områdeskoder. Ett försök att själva föra på områdeskoder för dessa företag med hjälp av det adressmatchningssystem som använts i resvaneundersökningen ledde till att nästan hälften av de saknade områdeskoderna kunde kompletteras.

En jämförelse mellan det aggregerade Basregistret och uppgifter över sysselsatt dagbefolkning enligt Folk- och bostadsräkningen visade på stora skillnader för ungefär vart femte område (i vår 850-indelning). En detaljgranskning av ett antal områden visade att Basregistret genomgående var mindre tillförlitligt. I ett antal fall redovisade Basregistret områden med mycket blygsam sysselsättning där Fob-materialet visade en sysselsättning på flera tusen personer. Slutsatsen av analysen blev att Basregistret i många fall fortfarande inte registrerar företagen på arbetsställenivå utan på företagsnivå, dvs. samtliga sysselsatta redovisas ofta fortfarande på huvudkontoret och inte på de enskilda arbetsplatserna, trots att detta varit en uttalad målsättning vid förbättringen av registret.

Bristerna och ofullständigheterna i de olika registren ledde till att storleksdata till slut nästan enbart kom att hämtas från Fob-data. Vid denna tidpunkt fanns Fob-85, som var den första registerbaserade folk- och bostadsräkningen, tillgänglig. Denna uppvisade dock också så stora brister att den var oanvändbar för våra syften. Vi var därför hänvisade till att använda Fob-80. De storleksuppgifter som slutligen användes kom därför huvudsakligen att utgöras av antal anställda i olika branscher i respektive målområde. Tidigare studier (Widlert, Tegnér och Berglund, 1977) har också visat att olika storleksmått är starkt korrelerade och därför ger likartade resultat i analysen.

Ärendeindelning

Den ursprungliga ärendekodningen i resvaneundersökningen omfattar 31 slag av resärenden (se tabell 6.6).

Tabell 6.6 Resärenden i resvaneundersökningen

1.	till arbetsplats
2.	tjänstebesök
3.	till skola
4.	inköp av livsmedel
5.	annat inköp
6.	till bostaden
7.	post
8.	tele
9.	sjuk- och socialvård, tandvård, sjukbesök, myndigheter
10.	bank
11.	frisör, solarium
12.	restaurang, lunch, etc
13.	hotell
14.	teater, bio
15.	övriga nöjen, träffa kompis på stan
16.	annan utbildning än skola
17.	facklig, politisk eller religiös verksamhet
18.	konsert
19.	bibliotek
20.	museum
21.	promenad eller motion i direkt anslutning till bostaden, gå ut med hunden
22.	övrig motion och idrott utomhus, gå ut med barnen
23.	motion och idrott inomhus
24.	promenad, ströva i natur
25.	egen fritidsbostad
26.	annans bostad och annans fritidsbostad
27.	rundtur utan egentligt mål eller ärende
28.	jaga, fiska
29.	hämta och lämna person, möta person
30.	yrkesmässig förflyttning
31.	annat

Undersökningen ger således möjligheter till en mycket detaljerad ärendeindelning. Frågan om separata modeller behövs för olika resärenden (respektive om särskilda variabler behövs för vissa ärenden) är empirisk. Vi kunde dock redan från början utesluta möjligheten att utveckla modeller med en så fin ärendeindelning som den ovanstående. För att underlätta arbetet och för att minska på informationsmängden i registren skapades därför en aggregerad ärendeindelning enligt tabell 6.7.

Tabell 6.7 Reviderad ärendeindelning

Aggregerat ärende	Ingående ärenden
Arbete	
Tjänste	
Skola	
Inköp av dagligvaror	
Sällanköp	
Service	post, tele, sjuk- och socialvård, tandvård, sjukbesök, myndighet, bank, frisör, solarium
Restaurang	
Kultur	hotell, teater, bio, övriga nöjen, träffa kompis, annan utbildning, facklig/politisk/religiös verksamhet, konsert, bibliotek, museum
Utomhusmotion	övrig motion utomhus, gå ut med barnen, promenad, ströva i natur, jaga, fiska
Inomhusmotion	
Besök	besök i annans bostad, egen och annans fritidsbostad
Övrigt	rundtur utan egentligt mål eller ärende, hämta/lämna, annat

Ärendet "promenad eller motion i direkt anslutning till bostaden, gå ut med hunden" samt ärendet "yrkesmässig förflyttning" tas ej med i modellanalyserna. Dessutom försvinner ärende 6 "åter till bostaden" när delresorna omvandlas till rundturer. De modeller som utvecklats har täckt in ett eller flera av dessa aggregerade ärenden.

Storleksuppgifter

Storleksuppgifter för arbets- och tjänsteresor

För arbets- och tjänsteresor används uppgifter från Fob 80 om sysselsatt dagbefolkning (minst en timme per vecka) fördelat efter näringsgren som storleksuppgift. Näringsgrensindelningen har gjorts efter Nordisk Yrkesklassificering på begränsad nivå (NYK). Denna klassning har aggregerats till följande nio grupper:

Tabell 6.8 Indelning i yrkesgrupper

Grupp	Ingående yrken enligt NYK	Kod i NYK
1	tekniskt, kemiskt, fysikaliskt, medicinskt, religiöst, juridiskt, litterärt, övrigt naturvetenskapligt, militärt	00-03, 06-09, 98
2	hälso- och sjukvård	04
3	pedagogiskt	05
4	administrativt	1
5	kameralt och kontorstekniskt	2
6	kommersiellt (exklusive övrigt kommersiellt)	3 exkl. 33
7	övrigt kommersiellt	33
8	lantbruk, skogsbruk, fiskeri, gruv och stembrytning, tillverkning	4, 5, 7-8
9	sjöbefäl, civil bevakning och skydd, hushåll, servering, fastighetsskötsel städning, övrigt service, ej identifierbara yrken	6, 90-94, 99

Registret innehåller således uppgifter om antalet sysselsatta i vardera av dessa nio grupper för vart och ett av de 850 områdena.

Storleksuppgifter för skola, inköp, service och övrigt

Storleksuppgifter för skola, inköp, service, restaurang och kultur har hämtats från en näringsgrensfördelad Fob 80. Näringsgrensindelningen har gjorts enligt SNI-kod på fyrsiffernivå för nyckelkodsområden. Dessa områden har därefter aggregerats till 850-områdesnivå.

Storleksuppgifter för besöksresor

Som storleksmått för besöksresor används antalet boende i varje område. Dessa uppgifter är de enda som hämtats från Fob 85. I registret redovisas även ytorna i respektive område. Dessa ytor har hämtats från regionplanekontorets områdesdatabas och avser markyta exklusive kommunikationsytor, ytor för försvarsändamål, samt vissa impediment och restytor. Totalt ingår 95% av länets markyta.

Storleksuppgifter för rekreationsresor

Som storleksmått för rekreationsresor finns förutom uppgifter om antal anställda i olika näringsgrenar även ett särskilt register över utbud av olika rekreationsanläggningar i varje område. Registret har upprättats genom en särskild inventering från kartmaterial. Följande utbud är registrerade:

- friluftsgård
- campingplats
- tältplats
- motionsgård
- elljusspår
- slalombacke
- slalombacke med lift
- friluftsbad

- bassängbad
- golfbana
- skjutbana

Enbart förekomsten av dessa utbud har registrerats, inte några uppgifter om deras kvalitet eller omfattning.

6.6 Mätproblem

Nätdata

De använda trafiknäten saknar helt trafikstandarduppgifter för resor inom ett enda område (eller innehåller enbart en enda skafftid för sådana resor). Även vid korta resor som inte sker helt inom samma område blir naturligtvis trafikstandarduppgifterna osäkra, eftersom den exakta start- och målpunkten inom områdena kan ha stor betydelse för de verkliga tiderna och avstånden.

Även för gångavstånd till och från kollektivtrafik innebär angreppssättet problem eftersom gångavståndets variation inom zonerna kan vara mycket stort.

Kostnadsuppgifter

I samtliga modeller används rörliga bilreskostnader som baseras på de avdrag i självdeklarationen som Riksskatteverket medger för körsträckor över 1 000 mil. Detta avdrag som ska motsvara den rörliga kostnaden för en medelstor bil var i genomsnitt under undersökningsperioden 6,95 kronor per mil. Bensinpriset var i genomsnitt 4,15 kronor per liter (96 oktan, snabbtank). I verkligheten kommer naturligtvis olika personer såväl att ha olika kostnader som olika uppfattning om hur stor reskostnaden är.

För arbetsresor förutsätter vi dessutom att de intervjupersoner som antingen gör bilavdrag, eller anser att han skulle kunna göra detta om han åkte bil, räknar helt rationellt på resavdragen. Detta innebär att vi förutsätter att han verkligen upplever att bilreskostnaden reduceras i proportion till avdragsregler och den egna marginalskatten. Kollektivreskostnaden förutsätts inte påverkas av resavdragen eftersom kostnaden för månadskort under hela året inte översteg schablonavdraget för kostnader för intäkternas förvärvande (detta uppgick till 3 000 kronor). Om man har andra avdrag för inkomsternas förvärvande så att summan av avdragen överstiger schablonavdraget kan i realiteten även den upplevda kollektivkostnaden påverkas av avdragen.

Intervjupersoner som har leasingbil antas i de flesta modeller sakna rörlig kostnad för bilresor, vilket svarar mot den vanligaste modellen för bilförmån och att arbetsgivaren täcker alla rörliga kostnader. Om hushållet har flera bilar och inte alla är leasingbilar antar vi att det är mannen som disponerar leasingbilen.

För samtliga resor beräknas reskostnaden för kollektivresor utifrån de färdbevis intervjupersonen innehade vid intervjutillfället.

Vid *arbetsresor* räknar vi ut vilket färdbevis som blir billigast med hänsyn till resans längd (antalet taxezoner som passeras) och antalet arbetsdagar per månad som intervjupersonen uppgett i resvaneundersökningen. Blir månadskortet billigast räknar vi med detta pris dividerat med antalet arbetsdagar per månad.

Vid övriga resor antar vi att de som har månadskort, oavsett varför de har skaffat det, räknar med att marginalkostnaden är noll vid varje enstaka resa. Om man innehar förköpshäfte används denna kostnad och om man inte har något färdbevis används kontanttaxan.

För gång- och cykelresor antas genomgående att reskostnaden är noll.

Sammanfattningsvis kan vi konstatera att beräkningen av reskostnader förutsätter en lång rad antaganden. Om vi systematiskt felskattar intervjupersonernas upplevda kostnad kommer detta direkt att avspeglar sig i modellernas kostnadsvariabler och därmed även i de tidsvärden som beräknas med hjälp av modellerna. Tidsvärdena är tämligen proportionella mot de rörliga kostnader som förutsätts. Att vi inte tar hänsyn till att personer med högre inkomster i genomsnitt har bilar med högre rörliga kostnader än personer med lägre inkomster gör att vi systematiskt kommer att underskatta skillnaden i tidsvärden mellan grupper med höga respektive låga inkomster.

Det är angeläget med ytterligare forskning som belyser hur individerna upplever kostnaderna för olika färdmedel i olika situationer. Så länge sådan information saknas är det dessutom angeläget att kostnaderna beräknas på ett likartat sätt i olika projekt, så att de tidsvärden och andra resultat som erhålls blir så jämförbara som möjligt.

7. Validering

Det är naturligtvis önskvärt att på något sätt undersöka giltigheten dels hos erhållna delmodeller och dels hos hela modellsystemet som sådant. Syftet med modellerna är att beskriva ett visst tillstånd med avseende på resmönstret som funktion av ett antal oberoende variabler. Två tester kan då tänkas för att beskriva modellernas validitet:

- modellerna ska kunna beskriva tillståndet i det datamaterial de estimerats på
- modellerna ska kunna återskapa ett annat känt relevant tillstånd

7.1 Validering baserad på estimeringsdata

Det första kravet innebär att anpassningen till estimeringsdata ska vara tillräckligt god. Detta bör - förutom de anpassningsmått som ges av estimeringsprogrammet - också avse olika delmängder av datamaterialet. Ett exempel på detta kan vara att modellerna bör vara giltiga för såväl bilinnehavare som icke bilinnehavare. Ett annat exempel kan vara att modellerna bör vara giltiga såväl för kategorier som gör långa resor, som för kategorier som gör korta resor. Om en modell beskriver den genomsnittliga reslängden väl, men uppvisar stora skillnader för olika undergrupper kan validiteten ifrågasättas.

Förekomsten av sådana problem kan analyseras med hjälp av s.k. valideringstabeller, vilka kan produceras med hjälp av den programvara för estimering - ALOGIT - som vi använt. Valideringen innebär att observationerna i datamaterialet redovisas i en tabell, uppdelad dels efter de olika enskilda eller grupperade alternativen i modellen, och dels efter en godtycklig variabel som klassificerar observationerna i varandra uteslutande kategorier.

I samma tabell redovisas också standardavvikelseerna för antalet observerade val, samt de beräknade valsannolikheterna för de olika (enskilda eller grupperade) alternativen. Detta gör det möjligt att särskilja de fall, där den modellberäknade valsannolikheten särskilt mycket avviker från det faktiska valet. Detta leder ofta till att olika fel kan korrigeras och till förbättrade modellspecifikationer.

Valideringstabellerna gör det således möjligt att belysa validiteten hos enskilda delmodeller. Sådana tabeller redovisas i anslutning till de olika delmodellerna i avhandlingens empiriskt inriktade andra och tredje delar.

7.2 Validering baserad på externa datakällor

Det är inte tillräckligt att validera modellerna enbart utifrån de data som de estimerats på. Det är nästan alltid möjligt att hitta värden på olika modellparametrar som beskriver estimeringsdata tillräckligt bra. Detta är dock inte samma sak som att modellparametrarna också beskriver förändringar i resmönstret på ett bra sätt.

Ett sätt att validera enskilda delmodeller eller hela trafikmodellssystemet kan vara att undersöka dess beskrivning av ett känt resmönster. Det är då frågan om att utnyttja oberoende datakällor - antingen avseende det resmönster som resvaneundersökningen avser att beskriva (samtida data), eller något annat känt resmönster (exempelvis med hjälp av förefter-studier). Därvid krävs data inte bara avseende detta resmönster, utan också avseende förutsättningarna för resmönstrets uppkomst i form av modellens förklaringsvariabler. En sådan validering erbjuder stora svårigheter av olika slag.

Samtida data

För att beskriva ett resmönster krävs egentligen en resvaneundersökning. Sådana är dyra att genomföra, och finns därför sällan att tillgå. Parallellt med den hushållsbaserade resvaneundersökning som ligger till grund för modellskattningarna genomfördes också som tidigare nämnts en individbaserad resvaneundersökning. Den senare genomfördes i större skala, men med en annan intervjuteknik (brevenkät kombinerad med telefonintervjuer).

Den jämförelse som gjorts mellan de båda undersökningarna visar (Rosenqvist 1991) att även om hushålls- och individvariablerna stämmer rätt väl överens, så innehåller den individbaserade undersökningen ett systematiskt bortfall av korta resor jämfört med hushållsundersökningen. Skillnaden mellan undersökningarna innebär att ett försök till validering av modellsystemet mot individundersökningens datamaterial också skulle innehålla en jämförelse av olika enkätmetoder, vilket skulle göra det svårt att dra bestämda slutsatser.

Ett resmönster avspeglar sig också i antal fordon på olika väglänkar, och i form av antal personer som stiger på och av olika kollektivtrafikfordon. Sådana uppgifter är förhållandevis billiga att samla in, och detta görs också regelbundet av andra skäl. För att validera modellsystemet med hjälp av sådana data krävs dels att man kan särskilja de fordon och trafikanter som modellsystemet avser att beskriva, och dels att man på något sätt överför resmönstret till dessa storheter.

När det gäller vägtrafik är det förhållandevis svårt att ur befintliga fordonsräkningar särskilja det antal fordon som modellsystemet avser. En stor andel utgörs av yrkestrafik av olika slag, vilken dessutom sannolikt har en annan fördelning än de resor som modellsystemet avser.

Vad avser kollektivtrafikresor torde den helt avgörande andelen resor vara sådana som modellsystemet avser. De räkningar som man numera gör inom stora delar av kollektivtrafiknätet är dessutom av såpass god kvalitet (Råbäck 1991) att dessa data bör vara väl lämpade att beskriva modellsystemets kollektivtrafikresor. Uppgifter finns dessutom för en mycket stor mängd enskilda delar i trafiknätet.

Gång- och cykelresor räknas endast undantagsvis, och kan därför inte utnyttjas för valideringsändamål. Det är dessutom svårt att särskilja gångresor i modellsystemets mening (rena gångresor) från de gångförflyttningar som är en del i en kollektivtrafikresa eller i en bilresa.

När det gäller att jämföra modellsystemets resmönster med länkdata av den typ som beskrivits ovan, uppkommer ytterligare ett problem - att uttrycka resmönstret i form av länkdata. Detta görs genom att utnyttja ett nätutlägningsprogram som fördelar resorna på

olika länkar. Detta innebär att valideringen delvis kommer att avse nätutläggningsprogrammets förmåga att fördela resandet på de olika länkarna.

Länkdata kan inte användas för att validera enskilda ärenden, eftersom dessa data inte finns uppdelade på resärenden. Valideringsmöjligheterna inskränker sig därför till att omfatta hela modellsystemet samtidigt.

Data för andra tidsperioder

För att undersöka modellsystemets förmåga att beskriva förändringar skulle man kunna tänka sig att vidta någon åtgärd, och sedan genomföra en undersökning i syfte att erhålla det resmönster som uppstått efter det att åtgärden genomförts. Även om detta vore genomförbart, så skulle också detta sätt att validera modellerna stöta på problem.

Som nämnts ovan (se 5.9) innebär användningen av tvärsnittsdata att modellerna avser ett läge, som innebär en långsiktig anpassning av resmönstret till förutsättningarna i form av modellsystemets förklaringsvariabler. Denna anpassning sker olika snabbt, beroende på vilket resbeslut det handlar om. Det är exempelvis rimligt att tänka sig, att anpassningar i fråga om färdmedelsval och destinationsval för inköpsresor sker snabbare än anpassningar i fråga om destinationsvalet för arbetsresor.

Det förekommer säkert också en viss tröghet i anpassningen - att ändra beteende kan innebära en slags transaktionskostnad, exempelvis i form av att lära känna butiker i ett annat område. Det kan också ta tid att lösgöra medel, exempelvis för att köpa bil.

Ett ytterligare sätt att undersöka modellsystemets prediktiva förmåga vore att utnyttja kunskaper om ett tidigare resmönster. Förutom svårigheterna att ta fram data för prognosvariablerna för detta tillfälle har man kvar problemen med anpassningsmekanismerna.

Det är inte säkert att alla de svårigheter som vi har pekat på här utesluter varje försök till validering. Det handlar naturligtvis om att göra en bedömning av i vilken utsträckning de olika svårigheterna påverkar resultaten, och om att låta denna vara vägledande för beslutet.

Innan modellsystemet är helt implementerat, saknar vi dock möjligheter att validera modellerna på annat sätt än genom att utnyttja estimeringsmaterialet. Redovisningen av sådana valideringstabeller har därför fått förhållandevis stort utrymme i de empiriska delarna.

8. Ett modellsystem för Stockholms län

I detta avsnitt beskrivs de modeller som utvecklats i form av modellstrukturer. De viktigaste parametrarna i varje modell kommenteras verbalt. Det primära resultatet från arbetet med modellerna är egentligen de skattade parametrarna som visar inflytandet från de olika variablerna. Utrymmet medger dock inte att dessa värden här redovisas och kommenteras mer i detalj. Den mer detaljerade redovisningen återfinns istället i avhandlingens andra och tredje delar. Vissa resultat från de olika modellerna redovisas och jämförs dock översiktligt i avsnitt 8.7.

Kapitlet avslutas med en kortfattad beskrivning av det prognossystem som byggts upp för modellerna. Prognossystemet innefattas ej i denna avhandling utan beskrivs enbart för fullständighetens skull.

8.1 Arbetsresor

Modellstrukturen för arbetsresor består av delmodeller för de mer långsiktiga valen av bilinnehav och destination (arbetsplats), modeller för de mer kortsiktiga valen av resfrekvens, allokering av bilen inom hushållet och valet av färdmedel, samt modeller för valet av sekundär destination.

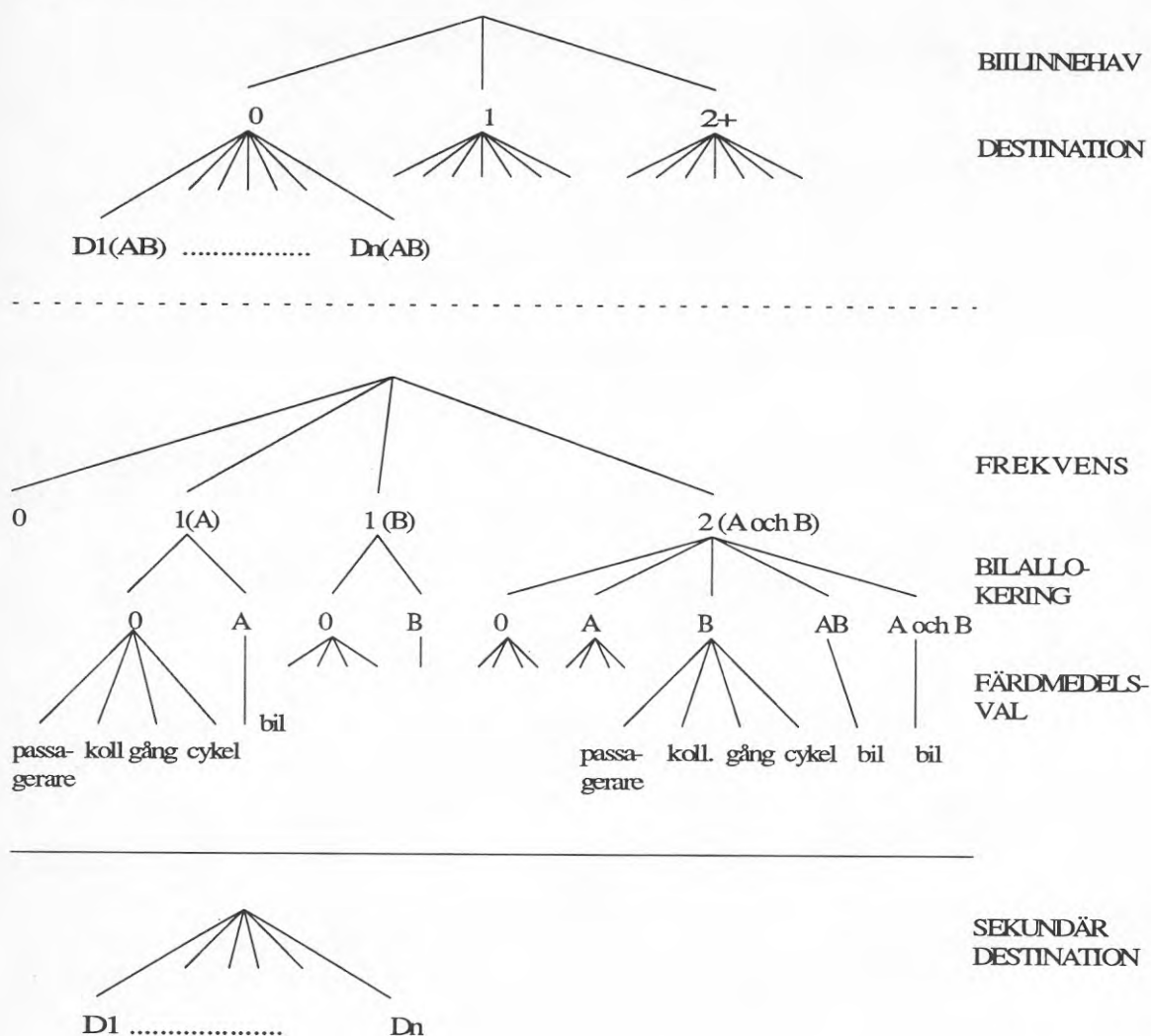
Bilnehavsmodellen har placerats i arbetsresestrukturen eftersom arbetsresorna utgör en långsiktigt styrande bas för det totala resandet och det därför är extra viktigt att fånga in hur bilnehavet påverkas av tillgängligheten för detta resärende. I bilnehavsmodellen ingår samtliga hushåll - oavsett om de förvärvsarbetar eller ej. I arbetsresestrukturens övriga modeller ingår enbart förvärvsarbetande hushåll.

Struktur

Figur 8.1 visar arbetsresemodellens struktur. I figuren är inte alla delar fullständigt utritade. Avkortade streck utan beteckning illustrerar sådana delvis utelämnade delar.

Det är praktiskt omöjligt att estimeras hela den visade strukturen simultant i ett enda steg. Istället delas den upp längs linjerna i figuren och de olika delarna estimeras sekventiellt. Delarna länkas vid estimeringsarbetet samman genom logsumvariabler. Detta innebär att modellerna för sekundära destinationer estimeras först. Dessa används sedan för att beräkna logsumvariabler som ingår i färdmedelsvalsmodellen när den mittersta strukturen skattas. Denna struktur används på samma sätt för att beräkna logsumvariabler till destinationsvalsmodellen när den översta strukturen skattas.

Figur 8.1 Arbetsresemodellen



Figur 8.1 visar valen för ett hushåll med två förvärvsarbetande medlemmar som vi betecknar A och B. Kodningen har utförts så att A normalt är mannen i hushållet och B normalt kvinnan. Hushåll med mer än två förvärvsarbetande delas upp så att de yngsta medlemmarna bildar separata enpersonshushåll. Modellen för hushåll med enbart en förvärvsarbetande blir väsentligt mindre komplex än den visade.

Bilnehavsmodellen behandlar valet mellan att inte ha bil, att ha en bil eller två (eller flera) bilar. Såväl ägda som leasade bilar behandlas i modellen. Destinationsvalsmodellen behandlar det samtidiga valet av arbetsplats för båda förvärvsarbetande i hushållet. Detta innebär att alternativen utgörs av kombinationer av arbetsplatser för person A och person B (ett alternativ kan t.ex. utgöras av att person A arbetar i innerstaden och person B i det egna bostadsområdet). Alternativen betecknas $D_1 - D_n$.

I frekvensmodellen finns alternativen att inte göra någon resa under den studerade dagen, att enbart person A reser, att enbart person B reser, samt att både person A och B gör var sin arbetsresa.

Bilallokeringsmodellen behandlar vem i hushållet som ska få tillgång till bilen. Om enbart person A reser finns alternativen att ingen använder bilen för arbetsresor (0), eller att person A använder den (A). Motsvarande alternativ återfinns om enbart person B gör en arbetsresa. Om båda gör arbetsresor finns följande alternativ:

- 0: ingen använder bilen
- A: mannen använder bilen
- B: kvinnan använder bilen
- AB: båda använder samma bil (samåkning i hushållet)
- A och B: mannen och kvinnan använder olika bilar (alternativet finns enbart om hushållet disponerar minst två bilar)

Om enbart person A gör en resa och han inte väljer att åka bil som förare (alternativ A på bilallokeringsnivån), kan han på färdmedelsnivån välja mellan bil som passagerare i ett annat hushålls bil eller med ej förvärvsarbete medlemmar av det egna hushållet, kollektivtrafik, gång och cykel. Samåkning med andra förvärvsarbete i hushållet behandlas enligt ovan som ett explicit alternativ på allokeringnivån och samåkning med andra hushåll som ett passageraralternativ på färdmedelsnivån. Om både A och B gör arbetsresor och B använder bilen enligt allokeringmodellen kan person A välja mellan bilpassagerare, kollektivtrafik, gång och cykel på färdmedelsnivån. Om ingen använder bilen (alternativ 0 på allokeringnivån) utgörs alternativen av en kombination av A:s och B:s övriga alternativ (dvs. 16 alternativ erhålls på färdmedelsnivån).

Underst i strukturen återfinns modellen för val av sekundära destinationer. En sekundär destination kan t.ex. utgöras av en butik som besöks på väg till eller från arbetet. Det är rimligt att anta att individen väljer målpunkt för denna typ av resor oberoende av övriga medlemmar i hushållet. Dessa modeller är därför de enda i arbetsresestrukturen som avser individens resande, inte hushålls.

I fortsättningen av kapitlet beskrivs arbetsresemodellerna nerifrån och upp, det vill säga i den ordning de estimerats.

Sekundära destinationer

Valet av sekundära destinationer modelleras som två explicita val: valet av att utföra ett sekundärt besök eller ej (vilket beskrivs i inköps-, service-, rekreations- och besöksresemodellerna), samt valet av destination om ett sådant besök utförs. I detta fall antas lokaliseringen av bostad och arbetsplats vara given, liksom färdmedlet för resan till den primära destinationen, det vill säga arbetsplatsen. Eftersom färdmedlet är givet skattas separata modeller för varje färdmedel.

De färdmedel som studeras är bil och kollektivtrafik. Gång och cykel utesluts eftersom de främst används vid korta resor inom en zon.

Variablerna i modellen är attraktionsvariabler, trafikstandardvariabler och alternativspecifika konstanter. Modellen innebär att vi kan modellera transportsystemets inverkan på valet av sekundära destinationer på ett realistiskt sätt.

Färdmedelsval, bilallokering och resfrekvens

På färdmedelsvalsnivån återfinns vi olika tids- och kostnadsvariabler med högre värdering av väntet-, bytes- och gångtid än åktid i fordon. Som väntat har personer som ibland använder sin bil i arbetet högre sannolikhet att åka bil en viss given dag. En bilkonkurrensvariabel mäter konkurrensen om bilen med ej förvärvsarbete med medlemmar av hushållet (konkurrens med de förvärvsarbete med medlemmarna modelleras explicit på allokeringsnivån). En dummyvariabel fångar in den högre sannolikheten för att resa som passagerare om man tillhör ett bilhushåll. Denna parameter speglar möjligheten att åka med ej förvärvsarbete hushållsmedlemmar. Att åka tillsammans med andra förvärvsarbete med medlemmar av hushållet ingår som ett särskilt alternativ i modellen.

Logsumvariabeln från modellen för sekundära destinationer är signifikant skild från noll. Detta innebär att tillgängligheten till olika målpunkter på väg till och från arbetet med ett visst färdmedel påverkar färdmedelsvalet vid arbetsresan - precis som antogs när modellstrukturen specificerades.

På bilallokeringsnivån visar könsvariabler att kvinnor - om allt annat är lika - har en lägre sannolikhet att få tillgång till hushållets bil än männen. Detta kan tolkas som en sämre förhandlingsposition på grund av traditionella könsroller. Dummyvariabler visar att yngre kvinnor och kvinnor med högre utbildning tycks vara mer jämlika med männen när de "förhandlar" om hushållets bil. En logsumvariabel mäter hur tillgängligheten påverkar bilallokeringen. Parametern visar att tillgängligheten är en viktig faktor när hushållet avgör vem som ska få tillgång till bilen.

I frekvensmodellen finns variabler som fångar in den lägre sannolikheten för att göra arbetsresor på lördagar och söndagar. Dummyvariabler visar att personer som arbetar deltid reser mer sällan än personer som arbetar heltid, det vill säga att deltidsarbetet inte bara innebär att man arbetar färre timmar per dag, utan även färre dagar per vecka. En variabel för hushåll med två förvärvsarbete och små barn visar att sannolikheten för att enbart kvinnan reser en given dag minskar. Detta är troligen en effekt av att det är vanligare att kvinnor stannar hemma och vårdar sjuka barn eller besöker barn på dagis än att män gör detta. Logsumvariabeln från allokeringsmodellen är liten och inte signifikant skild från noll. Tillgängligheten har därmed ej kunnat visas ha en signifikant inverkan på antalet arbetsresor i förvärvsarbete hushåll. Tillgängligheten kan däremot tänkas påverka förvärvsfrekvensen - ett val som inte behandlas i modellsystemet.

Destinationsval och val av bilnehav

Som nämnts tidigare behandlar destinationsvalsmodellen kombinationer av arbetsplatser för båda förvärvsarbete i hushållet (om hushållet har två förvärvsarbete). I destinationsvalsmodellen ingår en logsumvariabel från bilallokeringsnivån. Denna mäter hushållets tillgänglighet till olika målpunkter, givet de färdmedel som hushållet har tillgång till. På detta sätt är det möjligt att ta hänsyn till att hushållets förvärvsarbete kan välja sina

arbetsplatser så att möjligheterna till samåkning underlättas. Modellen innehåller även variabler som speglar att Södertälje och Nynäshamn utgör mer självständiga arbetsmarknader. Olika avståndsrelaterade variabler fångar in informationseffekter och tidsbudgetrestriktioner.

I bilinnehavsmodellen finns inkomstvariabler som visar en starkare inkomsteffekt för hushållets andra bil än för den första, samt olika variabler som mäter hushållets storlek och sammansättning - också med en starkare effekt på andrabilen. Parkeringskostnaderna i bostadsområdet visar sig påverka bilinnehavet signifikant. En logsumvariabel som mäter den ökade tillgängligheten till olika alternativa destinationer när hushållet har en eller två bilar visar sig ha ett starkt förklaringsvärde för sannolikheten att ha bil.

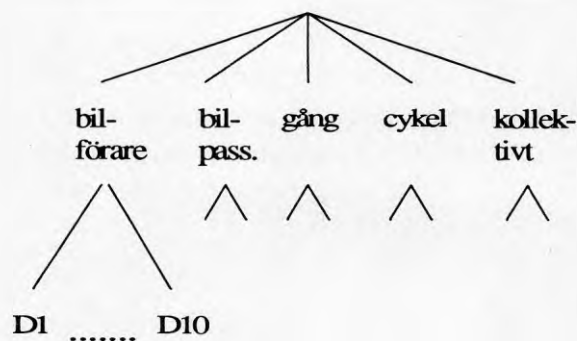
8.2 Skolresor

Skolmodellen avser resor med huvudärendet skola som utförts av personer som har "studier" som huvudsaklig sysselsättning. Skolmodellen avser således *individens* resor. Eftersom resvaneundersökningen enbart behandlar resor som utförts av personer som är 12 år eller äldre ingår i huvudsak resor utförda av elever på högstadiet och uppåt (även vissa mellanstadieelever som intervjuats på våren när de gick i sjätte klass ingår).

Struktur

Skolelever på mellanstadie-, högstadie- och gymnasienivå har normalt inte något val av om de ska resa till skolan eller ej. Modeller för resfrekvens är därför inte relevanta för dessa grupper. I undersökningen finns ej tillräckligt många observationer för universitetsstuderande för att tillåta att separata modeller estimeras för denna grupp. För skolresorna modelleras därför enbart destinations- och färdmedelsval. När modellen används utnyttjas genomsnittliga resfrekvenser för olika åldersgrupper för att bestämma antalet resor.

Figur 8.2 Skolresemodellen



Skolmodellens struktur är resultatet av empiriska tester som konsekvent visade att destinationsvalet måste placeras under färdmedelsvalet för att ge korrekta värden på logsumparametern.

Modellen innehåller fem färdmedelsvalsalternativ. För estimeringen gjordes ett stratifierat slumpmässigt urval av 10 destinationer.

Destinations- och färdmedelsval

För yngre elever används antalet elever i området som attraktivitetsvariabel. För studenter på universitetsnivå används antalet sysselsatta lärare. Eftersom eleverna på grundskole- och gymnasienivå inte har ett fritt val av skola innehåller modellen även en dummyvariabel som beskriver den höga sannolikheten att yngre elever går i skola i sitt eget bostadsområde (mätt på 850-områdesnivå).

På färdmedelsvalsnivån återfinns de traditionella trafikstandardvariablerna - alla med parametrar som skattats med god statistisk kvalitet. Resultaten visar att bilpassagerare framför allt är ett alternativ för gymnasieelever - inte för grundskoleelever som oftast har sin skola så nära bostaden att gång eller cykel är de dominerade färdmedlen och inte heller för universitetsstuderande. Resultaten illustrerar föräldrarnas benägenhet att skjutsa barnen till skolan. Studenter på universitetsnivå har ofta flyttat hemifrån och bor de med sina föräldrar är avståndet till skolan ofta så långt att det är mindre sannolikt att föräldrarna är beredda att skjutsa.

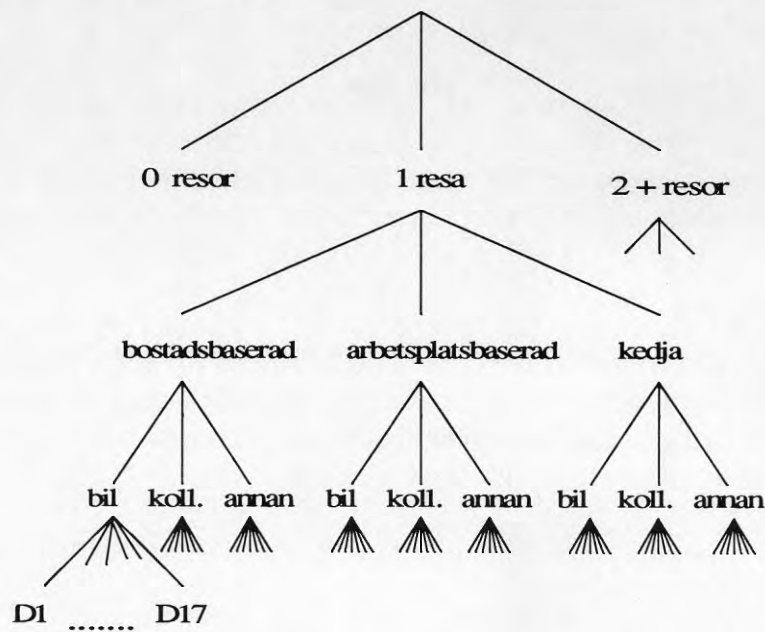
Cykelalternativet har en högre sannolikhet om eleven går på gymnasiet. Orsakerna är desamma som diskuterats ovan. Cykel har naturligt nog en lägre sannolikhet under vintern. Alternativet bil som förare är enbart möjligt om eleven har fyllt 18 år och har körkort. Sannolikheten att köra bil varierar starkt med den relativa tillgången på bilar i hushållet (antalet bilar i förhållande till antalet personer med körkort).

8.3 Tjänsteresor

Struktur

Systemet täcker alla tjänsteresor i Stockholms län utom dem som utförs av personer som kör bil yrkesmässigt (buss- och lastbilschaufförer, patrullerande poliser etc.). Tjänsteresomodellen har sysselsatta *individer* som bas.

Figur 8.3 Tjänsteresemodellen



Högst upp i strukturen finns en separat estimerad frekvensmodell med alternativen att inte resa, att göra en resa eller att göra två eller fler tjänsteresor. Restypmodellen behandlar alternativen att göra tjänsteresan bostadsbaserat, arbetsplatsbaserat, eller i samband med resan till eller från arbetet. För att förenkla strukturen antar vi att valet av färdmedel för tjänsteresan sker oberoende av valet av färdmedel till arbetet. Tre färdmedel särskiljs - bil (förare, passagerare och taxi), kollektivtrafik och annat (gång och cykel). På destinationsvalsnivån har 17 destinationer dragits slumpmässigt från fem olika strata som relaterats till avståndet från arbetsplatsen. Valet av restyp, färdmedel och destination modelleras samtidigt i en struktur med 153 alternativ på den understa nivån.

För kedjealternativen använder vi den *extra* restiden och kostnaden om alternativen ifråga väljs istället för att resan görs direkt mellan arbetet och bostaden (eller omvänt). Tiden beräknas - som i alla övriga modeller - för den tidsperiod då resan utfördes, det vill säga med hänsyn till det trafikutbud och den trängsel som finns vid olika tider på dygnet.

Val av destination, färdmedel och restyp

På destinationsvalsnivån mäts attraktiviteten med det totala antalet sysselsatta personer i varje målområde. Färdmedelsvalmodellen innehåller de traditionella trafikstandardvariablerna, en del socioekonomiska variabler och en logsumvariabel från destinationsvalsnivån.

Den viktigaste variabeln på restypsnivån är logsumvariabeln från de underliggande nivåerna. Logsumvariabeln visar att tillgängligheten har en signifikant effekt på valet av restyp. Modellen innehåller även variabler som beskriver skillnaden i sannolikheter för olika restyper om tjänstebesöket ska utföras tidigt på morgonen eller sent på eftermiddagen, och för om tjänstebesöket har lång varaktighet. Modellen innehåller även könsvariabler som speglar skillnader i tjänsteresemönster mellan män och kvinnor.

Val av resfrekvens

De viktigaste variablerna i frekvensmodellen är knutna till individens yrke eller till den typ av arbetsplats han arbetar på. Företag och yrken med många tjänsteresor är - som väntat - överrepresenterade i Stockholms innerstad. När variabler för yrke och typ av arbetsplats introduceras i modellen överskattas resfrekvensen i innerstaden. Dummyvariabler korrigerar för denna effekt.

I frekvensmodellen antar vi att destinationerna väljs oberoende av om en eller två tjänstebesök ska utföras under dagen. Logsumvariabeln för 2+ alternativet beräknas därför helt enkelt som 2,3 gånger logsumman för enresealternativet (i genomsnitt utförs 2,3 tjänsteresor per dag i detta alternativ). Logsumparametern är signifikant skild från noll och har ett värde på 0,8 vilket indikerar att tillgängligheten har en stark effekt på resfrekvensen.

8.4 Inköpsresor

Resmönstret utgörs av individers resor mellan olika målpunkter och med olika färdmedel. Dessa resor kan i varierande utsträckning vara resultat av individuella beslut, beslut utanför hushållet eller beslut fattade inom ramen för ett samspel mellan olika hushållsmedlemmar. När det gäller arbets-, tjänste- och skolresor så är beslutet att genomföra en resa i stor utsträckning ett resultat av antingen överenskommelser mellan enskilda hushållsmedlemmar och utomstående (t.ex. att förvärvsarbeta, även om det även här kan förekomma inslag av hushållsinteraktion) eller beslut helt utanför hushållet (skolplikt, beordrade tjänsteresor). Det blir härigenom också givet vem det är som utför resan. Förekomst och genomförande av inköpsresor är emellertid ett beslut som till alla delar fattas inom hushållet.

Hushållets frihet att själv organisera inköpsresandet ställer ett antal krav på modellutformningen. Trots att vårt syfte endast är att beskriva frekvens-, destinations- och färdvalsval så behöver egentligen även följande aspekter beaktas:

- hushållets efterfrågan på varor av olika slag
- hushållets val av inköpsstrategi
- fördelningen av resandet på individer
- inverkan av restriktioner avseende tillgängliga tidsluckor

Som namnet antyder, så kan hushållet i allmänhet antas ha en kontinuerlig efterfrågan på dagligvaror. Detta behöver dock inte innebära, att hushållet för den skull genomför dagliga inköpsresor. Varje inköpsresa innebär en kostnad för hushållet, och eftersom hushållet kan överblicka sin konsumtion för en tid framåt, finns det möjligheter att minska kostnaderna för inköpen genom att företa färre, men mer omfattande, inköpsresor. Med kostnader avses här generaliserade kostnader, där således även restiden ingår. Avvägningen mellan omfattning och antal inköp kallar vi här hushållets inköpsstrategi.

Förekomsten av sådana inköpsstrategier innebär att det finns ett beroende mellan olika restyper. Om exempelvis bensinkostnaden ökar, så blir det dyrare att genomföra ett stor-

marknadsinköp. Detta kan då resultera i att hushållet byter inköpsstrategi, och i stället för ett stormarknadsinköp per vecka genomför 5 dagliga inköpsresor.

Inköpsstrategierna är sannolikt oftast veckobaserade. Det innebär, att vi bör observera en hel veckas inköpsresande för att kunna definiera en vald inköpsstrategi. Svårigheterna att genomföra en sådan datainsamling ledde emellertid till att undersökningen begränsades till att omfatta en dags resande. Detta innebär, att vi inte har kunnat modellera inköpsresandet med det beroende som finns mellan olika inköpsresor. Vi har i stället tvingats välja en enklare ansats, där vi antar oberoende mellan de olika inköpsresorna.

Fördelningen på individer är inte given för inköpsresor, till skillnad från ärenden som avser exogent "kontrakterade" individer, som när det gäller arbets-, tjänste- och skolresor. Hushållet kan exempelvis minska kostnaderna för att genomföra inköp genom att låta en hushållsmedlem utföra inköpet under hemresan från arbetet i stället för att en annan hushållsmedlem genomför en bostadsbaserad inköpsresa. För att kunna beakta hushållets möjligheter att fördela inköpsresorna på olika hushållsmedlemmar kan en explicit modell för att allokera hushållets resor till olika individer ställas upp.

Denna modell måste också beakta individernas möjligheter att genomföra resorna i olika sammanhang. Inköpsresor kan genomföras på flera sätt - dels som en rundtur från ett mer permanent uppehållsställe som bostaden eller arbetsplatsen, och dels som en kombinationsresa, där inköpet görs i samband med andra ärenden. När resan inte är en renodlad inköpsresa uppkommer problem med beroende mellan olika ärendetyper när modellen ska tillämpas. Ska man först generera exempelvis serviceresor för att ha som utgångspunkt för möjligheten att göra en kombinerad inköps- och serviceresor, eller ska man göra tvärtom?

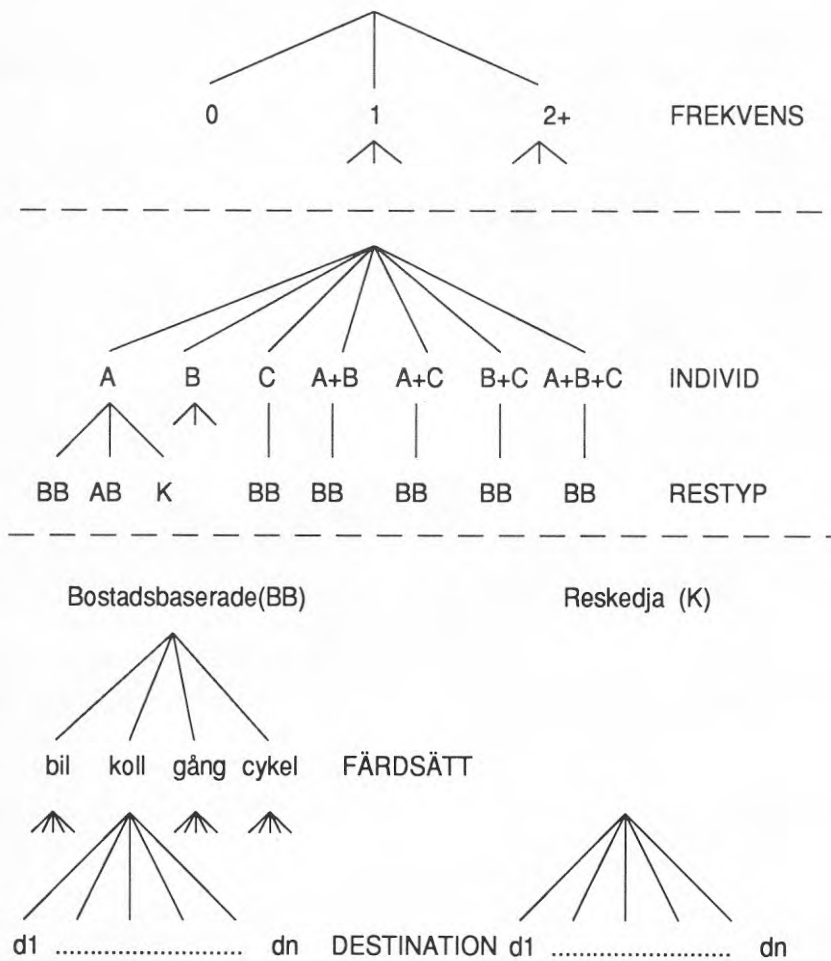
För att fullt ut beakta beroendet mellan olika ärenden krävs därför en modellansats som simultant beskriver olika resärenden. Detta för emellertid till betydligt mer komplicerade modeller, vilket inte bedömts genomförbart inom ramen för detta projekt. I stället ansätts ett förenklat beroende, vilket innebär att arbetsresan ses som överordnad övriga resor. Givet förekomsten av en arbetsresa innehåller modellen såväl alternativet att genomföra en arbetsplatsbaserad inköpsresa som alternativet att genomföra en inköpsresa i kombination med resan från arbetet till bostaden. Ett generaliserat beroende mellan övriga ärenden och arbetsresan existerar dock såtillvida att färdmedelsvalet för arbetsresan också beror av möjligheterna att genomföra ett ytterligare -"genomsnittligt" ärende mellan arbetsplatsen och bostaden.

Såväl fördelningen på individer som valet av inköpsställe och valet av färd sätt är direkt beroende av restriktioner avseende tillgängliga tidsluckor. Dessa tidsluckor bestäms dels av verksamheter som är låsta i tiden, som t.ex. arbete, och dels av inköpsställets öppettider. Tidsluckans storlek måste inte nödvändigtvis vara lika med den faktiska vistelsetiden för ett genomfört besök. Det kan tänkas, att det vore tillräckligt att hinna till butiken strax före stängningsdags. Definitionen av stängningsdags är inte heller helt given, eftersom olika butiker har olika stängningstider. Olika antaganden avseende kravet på tidslucka har prövats i modellarbetet.

Struktur

Vid tidigare studier av destinations- och färdmedelsval för inköpsresor har det visat sig, att destinationsvalet bör placeras längst ner i strukturen. Detta är därför utgångspunkt i detta arbete. Det är också rimligt att anta, att restyps- och individvalet bör placeras högre upp i strukturen. Att frekvenssteget ligger högst upp är självklart, eftersom de övriga valen förutsätter att en resa görs. Denna struktur ansattes under planeringen av projektet, och har också bekräftats av estimeringsarbetet. Strukturen redovisas i figur 8.4 nedan:

Figur 8.4 Inköpsresemodellen



På den översta nivån ligger således frekvensvalet. Det är formulerat som hushållets val mellan att inte resa, att göra en resa och att göra två eller flera resor per dag. Givet att någon resa görs, så följer individvalet. Person A i figuren är mannen i hushållet, medan person B är kvinnan. Person C är övriga personer i hushållet, vanligen barnen. Modellens alternativ är alla kombinationer av dessa personer, dvs. 7 st.

För personerna A och B, vilka kan ha gjort en arbetsresa, följer sedan val av restyp. Denna nivå innehåller förutom alternativet med en bostadsbaserad inköpsresa också alternativet med en arbetsplatsbaserad inköpsresa och alternativet att göra ett inköp på vägen mellan arbete och bostad. För de resor som innebär att person A eller person B reser tillsammans med någon annan person antas att endast en bostadsbaserad resa är ett rimligt alternativ.

För bostadsbaserade resor följer sedan färdmedels- och destinationsval. Färdsättsalternativen utgörs av bil, kollektivt färdstätt, gång och cykel. Destinationsalternativen utgörs av de områden som innehåller ett för inköpsresorna relevant utbud.

För de arbetsplatsbaserade inköpsresorna och för inköpen på vägen mellan arbetet och bostaden förekommer endast destinationsval, eftersom färdstättet bestäms av färdstättet till arbetet. I princip finns dock en viss återkoppling till valet av färdstätt för arbetsresan, eftersom färdmedelsvalsmodellen för arbetsresorna innehåller en logsumvariabel som beskriver den förväntade nyttan av att utträta ett ärende på vägen oavsett ärendets karaktär.

Modellen för val av sekundär destination har skattats separat för alla ärenden samtidigt, eftersom materialet inte är stort nog för att tillåta ärendewis estimering. Destinations- och färdmedelsvalet har skattats i ett annat separat steg. Individ- och restypsvalsdelen estimerades i ett ytterligare separat steg, medan frekvensdelen har estimerats som ett sista separat steg. De streckade linjerna i figuren markerar olika estimeringssteg.

Inköpsresorna är egentligen inte en särskilt homogen grupp. Det har visat sig att man bör göra en uppdelning på dagligvaruinköp med kort varaktighet å ena sidan, och övriga dagligvaruinköp samt sällanvaruinköp å den andra sidan. Modellstrukturen är dock densamma för dessa båda undergrupper.

Resultat

Sammanhängande modellsystem har skattats separat för de båda inköpskategorier som beskrivits ovan. Dessa modellsystem innebär att åtgärder i trafiksystemet kan beskrivas med avseende på

- antalet resor
- fördelningen på olika hushållsmedlemmar
- fördelningen på olika restyper
- fördelningen på destinationer för olika restyper
- fördelningen på färdstätt

Resultaten visar bl.a. att

- restider och reskostnader har statistiskt signifikant betydelse för destinations- och färdmedelsval, och därmed för tillgängligheten till olika inköpsställen
- tillgängligheten till inköpsställena för bostads- och arbetsplatsbaserade inköp har statistiskt signifikant betydelse när det gäller restypsväl
- tillgängligheten till inköpsställena för olika hushållsmedlemmar har statistiskt signifikant betydelse för vilka hushållsmedlemmar som genomför inköpet
- tillgängligheten till inköpsställena för hela hushållet har statistiskt signifikant betydelse för antalet inköpsresor

Modellsystemet bekräftar därmed i huvudsak de hypoteser som ställdes upp inledningsvis. Det enda undantaget utgörs av att den tillgänglighet som kedjeresorna medför inte tycks ha något inflytande på restypsvalet. Detta kan bero på att såväl tillgängligheten med avseende på inköp som restiden varierar med arbetsresans längd, och att nyttan av den ökade tillgängligheten därför kan uppvägas av den hårdare tidsrestriktion som är förknippad med den längre restiden.

Frekvensval

Antalet resor som ett hushåll genomför har visats bero av hushållets sammansättning och inkomst (förutom av tillgängligheten som nämnts ovan). Det har dock inte varit möjligt att behandla inköpsresorna som ett val av inköpsstrategi, med vilket avses en kombination av mindre och större inköp över en veckocykel. Detta innebär, att hänsyn inte kunnat tas till det beroende mellan större och mindre inköp som sannolikt finns i verkligheten.

Individ- och restypsval

Individ- och restypvalsmodellerna beaktar också i viss utsträckning hushållsinteraktioner, vilket innebär att resor kan omfördelas mellan olika individer. De variabler som påverkar en sådan omfördelning är dels tillgängligheten till inköpsställena - vilken påverkas av trafiksystemet - samt arbetstiderna för mannen och kvinnan i hushållet. Resultaten visar att dessa har större betydelse för inköpsresorna än exempelvis öppettider. De påverkar såväl restypsvalet som vem i hushållet som genomför inköpet.

Färdmedels- och destinationsval

I färdmedels- och destinationsvalsmodellerna är restids- och reskostnadsvariabler samt områdesbeskrivningsvariabler de centrala. Modellerna innehåller signifikanta estimat av alla dessa parametrar, vilka genom att modellsystemet är sammanhängande får effekter i alla valdimensioner.

8.5 Service- och rekreationsresor

Förutsättningarna för service- och rekreationsresorna är mer olikartade än när det gäller inköps- och besöksresorna. Vissa resor kan ha karaktär av försörjningsaktivitet (vissa serviceresor), medan andra resor har karaktär av aktivitet för enbart den resandes egen skull (exempelvis restaurangresor). Man kan också tänka sig mellanformer, som exempelvis en resa till en fotbollsplan, där ett barn åtföljs av någon förälder. Det är då en resa för barnets skull, medan vem som helst av föräldrarna (eller båda) kan följa barnet.

Vissa serviceresor (exempelvis hämta eller lämna paket på posten) kan således vara föremål för en fördelningsprocess mellan medlemmarna i hushållet, vissa rekreationsresor och kulturreSOR kan vara det när det gäller vissa individer (vem följer barnet till fotbollsmatchen eller till museet), medan det är svårare att tänka sig att restaurangresorna är det.

Man kan alltså i viss utsträckning tänka sig en fördelningsprocess för vissa resor inom dessa ärendetyper (möjligen med undantag för restaurangresorna).

Även om möjligheterna att omfördela resor mellan de olika hushållsmedlemmarna är mer begränsade här än när det gäller inköpsresor kan de således inte helt uteslutas. Det är därför rimligt att pröva hypotesen om en fördelning av resor på olika individer och kombinationer av individer som en funktion av såväl tillgänglighet som andra variabler. Det har också syntts rimligt att pröva hypotesen om fördelning på olika restyper som funktion av samma variabler.

Detta innebär att modellstrukturen kommer att ha samma form som för inköpsresorna (se figur 8.4).

Resultat

Även för service- och rekreationsresorna har vi skattat ett sammanhängande modellsystem, som innebär att åtgärder i trafiksystemet kan beskrivas med avseende på:

- antalet resor
- fördelningen på olika hushållsmedlemmar
- fördelningen på olika restyper
- fördelningen på destinationer för olika restyper
- fördelningen på färdsätt

Resultaten visar bl.a. att

- restider och reskostnader har statistiskt signifikant betydelse för destinations- och färdmedelsval, och därmed för tillgängligheten till olika målpunkter
- tillgängligheten till service- och rekreationsutbudet när det gäller bostadsbaserade och arbetsplatsbaserade resor har statistiskt signifikant betydelse när det gäller restypsväl
- tillgängligheten till service- och rekreationsutbudet för olika hushållsmedlemmar har statistiskt signifikant betydelse för vilka hushållsmedlemmar som genomför resan
- tillgängligheten till service- och rekreationsutbudet för hela hushållet har statistiskt signifikant betydelse för antalet resor av denna typ

Modellsystemet bekräftar därmed i huvudsak de hypoteser som ställdes upp inledningsvis. Även i detta fall tycks den tillgänglighet som kedjeresorna medför inte ha något inflytande på restypsvalet. Detta kan även här bero på att såväl tillgängligheten med avseende på exempelvis serviceärenden som restiden varierar med arbetsresans längd, och att nyttan av den ökade tillgängligheten kan uppvägas av den hårdare tidsrestriktion som är förknippad med den längre restiden.

Frekvensval

Antalet resor som ett hushåll genomför har visats bero av hushållets sammansättning och inkomst (förutom av tillgängligheten som nämnts ovan). Inkomsten har en betydligt större betydelse för antalet service- och rekreationsresor än för inköpsresorna. Detta gäller också tillgänglighetens inverkan, vilket indikerar att substitutionseffekterna mellan hushållsmedlemmarna är mindre än när det gäller inköpsresorna.

Individ- och restypval

Dessa modeller beaktar också i viss utsträckning hushållsinteraktioner, vilket innebär att resor kan omfördelas mellan olika individer. De variabler som påverkar en sådan omfördelning är dels tillgängligheten till rekreations- och serviceutbudet - vilken påverkas av trafiksystemet - och dels arbetstiderna för mannen och kvinnan i hushållet. Dessa variabler påverkar såväl restypvalet som vem i hushållet som genomför resan.

Färdmedels- och destinationsval

I dessa modeller är restids- och reskostnadsvariabler samt områdesbeskrivningsvariabler de centrala. Modellerna innehåller signifikanta estimat av alla dessa parametrar, vilka genom att modellsystemet är sammanhängande får effekter i alla valdimensioner.

8.6 Besöksresor

Besöksresorna skiljer sig från de övriga resorna på flera sätt - såväl när det gäller beteendet, som när det gäller vår förmåga att beskriva dem. Medan inköp är en försörjningsaktivitet som kan utföras av olika hushållsmedlemmar, så är besök en aktivitet som oftast genomförs för den besökandes (och den besöktes) egen skull. Det är därför svårare att tänka sig att besöksresor är föremål för en direkt fördelningsprocess inom hushållet (även om sådana fall är tänkbara, exempelvis om besökets syfte är att lämna något). Ofta är det den besökte, som genom inbjudan avgör vilken/vilka hushållsmedlemmar som är aktuella.

En ytterligare skillnad är att en besöksresa - för det mesta - förutsätter en överenskommelse med den/de besökta om tidpunkt. Därmed beror besöksresorna inte bara på tidsbudgetrestriktioner i det egna hushållet, utan också på restriktioner i det besökta hushållet. Detta förhållande utgör naturligtvis en betydligt mer begränsande faktor på vår förmåga att beskriva destinationsvalet än vad exempelvis öppettiderna utgör för inköpsresorna.

Besök har ofta en längre varaktighet än inköp, vilket innebär att möjligheterna att genomföra dem i samband med arbetsresor också är begränsade. Valmöjligheterna när det gäller restyp blir därför färre än när det gäller inköpsresor.

Flera av de förhållanden som nämnts ovan innebär att möjligheterna att göra modeller för besöksresorna blir mer begränsade. Speciellt gäller detta destinationsvalet, där vi naturligtvis inte har information om var bekantskapskretsen är bosatt för de olika hushållen.

Man kan diskutera huruvida besöksresorna, med tanke på att de är förhållandevis individberoende, bör behandlas som individbeslut snarare än som hushållsbeslut. Eftersom sällskapsstorleken kan beaktas även i en individansats bör detta inte spela så stor roll för färdmedels- och destinationsvalsdelen. Ett restypsväl är tänkbart även i en individansats. I detta fall är det dock rimligt att tänka sig en mer begränsad utbytbarhet än när det gäller de resor som tidigare behandlats.

Individvalet blir inte aktuellt i en individansats. Frekvensvalet kommer att avse resor per person i stället för resor per hushåll, och kan då innehålla variabler som beskriver individen, som exempelvis kön, sysselsättning och ålder. Dessa variabler ingår mer eller mindre explicit i frekvensvalet i en hushållsansats genom logsumman från individvalet. Även en individorienterad ansats kan innehålla hushållsvariabler. Exempel på sådana är hushållsstorlek, hushållsinkomst och bilinnehav.

Det kan således göras sannolikt att det inte är särskilt meningsfullt att utgå från en hushållsansats när det gäller besöksresor. Eftersom detta egentligen bör visas empiriskt har denna hypotes ändå testats, men förkastats. Arbetet redovisas därför med utgångspunkt från en individansats, men testen av hushållsansatsen redovisas också.

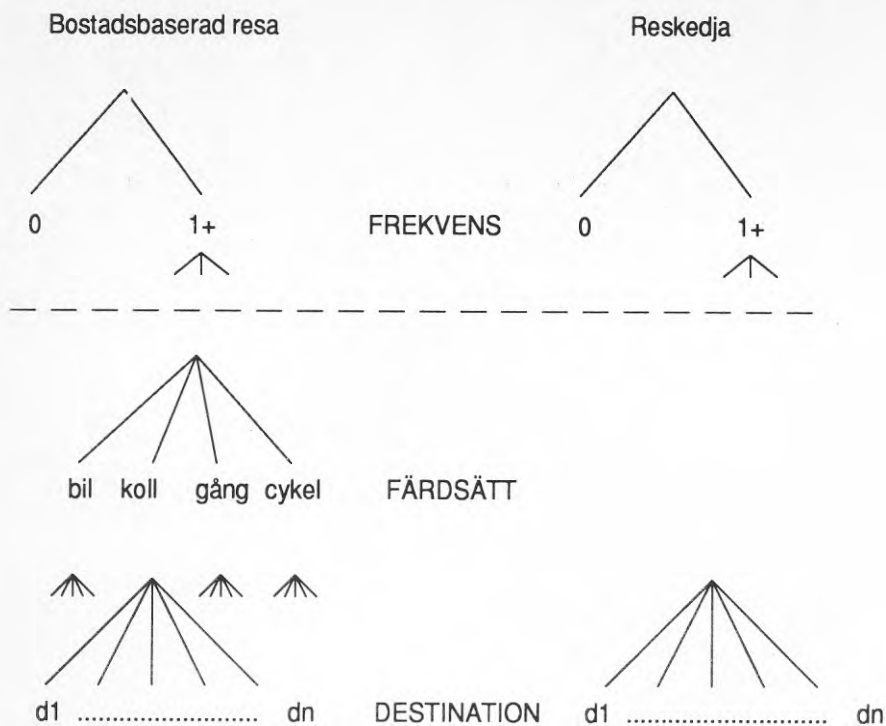
Struktur

Strukturen för besöksresorna blir således betydligt mer individorienterad än när det gäller inköpsresorna. Eftersom varken restypsväl- eller individvalsstegen visat sig tillämpbara, så återstår endast färdmedelsväl, destinationsväl och frekvensväl. Det förekommer både bostadsbaserade besöksresor och besöksresor i samband med resan mellan arbetet och bostaden, vilket innebär att det krävs två separata strukturer - en för vardera restypen.

Strukturschemat för besöksresorna (se figur 8.5) innehåller därför en struktur för bostadsbaserade resor, och en struktur för kedjeresor. Den bostadsbaserade strukturen är tillämpbar på alla individer i populationen i respektive område. Strukturen för kedjeresor är tillämpbar på dem som gjort en arbetsresa under dagen. Modellen för destinationsväl för denna grupp är dessutom betingad av färdmedelsvalet för arbetsresan. Detta innebär - naturligt nog - att kedjeresan är lika beroende av arbetsresan som kedjeresorna i inköpsresomodellerna.

I denna modellstruktur har frekvenssteget skattats separat för såväl bostadsbaserade resor som för kedjeresorna.

Figur 8.5 Besöksresemodellen



Resultat

Även för besöksresorna har ett sammanhängande modellsystem skattats, som innebär att åtgärder i trafiksystemet kan utvärderas med avseende på

- antalet bostadsbaserade resor
- fördelningen på destinationer för olika restyper
- fördelningen på färd sätt

Resultaten visar bl.a. att

- restider och reskostnader har statistiskt signifikant betydelse för destinations- och färdmedelsval, och därmed för tillgängligheten till olika potentiella besökspunkter (den övriga befolkningen)
- tillgängligheten till dessa besökspunkter har statistiskt signifikant betydelse för antalet besöksresor

Det har också visats att

- tillgängligheten till besökspunkterna saknar statistiskt signifikant betydelse när det gäller restypval
- tillgängligheten till besökspunkterna för olika hushållsmedlemmar saknar statistiskt signifikant betydelse för vilka hushållsmedlemmar som genomför besöket

Modellsystemet bekräftar därmed i huvudsak de hypoteser som ställdes upp inledningsvis. Det enda undantaget utgörs av att den tillgänglighet som kedjeresan medför inte heller här tycks ha något inflytande. Detta kan - som när det gäller de övriga resärendena - bero på att såväl tillgängligheten med avseende på besökspunkter som restiden varierar med arbetsresans längd, och att nyttan av den ökade tillgängligheten kan uppvägas av den hårdare tidsrestriktion som är förknippad med den längre restiden. Med tanke på det låga värdet på logsumparametern i frekvensmodellen för de bostadsbaserade besöksresorna är det också rimligt att tänka sig att tillgängligheten har avsevärt mindre betydelse när det gäller besöksresorna.

Frekvensval

Antalet resor som en individ genomför har visats bero av individtyp och hushållets sammansättning (förutom av tillgängligheten som nämnts ovan). Hushållsinkomsten har en viss positiv effekt på sannolikheten att göra ett besök (upp till en viss nivå).

Individ- och restypval

Försöken att skatta dessa modeller har visat att besöksresornas fördelning på restyp och individ inte beror på trafiksystemet. Detta är tämligen trivialt utifrån de förväntningar som kunde ställas härpå, och försöken torde ha sitt största värde som illustration till motsvarande modeller för de övriga resärendena eftersom de visar att en sådan modellformulering inte automatiskt resulterar i signifikanta modellparametrar.

Färdmedels- och destinationsval

I dessa modeller är restids- och reskostnadsvariabler samt områdesbeskrivningsvariabler de centrala. Modellerna innehåller signifikanta estimat av alla dessa parametrar, vilka genom att modellsystemet är sammanhängande får effekter i alla valdimensioner.

8.7 Jämförelse av resultaten för olika resärenden

I detta avsnitt diskuteras skillnader och likheter i de empiriska resultaten från modellstrukturerna för de olika resärendena.

Resfrekvens

Alla resfrekvensmodeller i systemet - utom modellerna för tjänsteresor - beskriver antalet resor som en funktion av hushållets sammansättning.

Modellerna för inköp, service, rekreation och besök innehåller även inkomstvariabler. Dessa inkomsteffekter skiljer sig kraftigt för de olika resärendena. När hushållets ekonomiska styrka ökar, ökar inköpsresorna endast till en viss gräns. Detta gäller också besöks-

resorna. Service- och rekreationsresorna ökar däremot påtagligt. Detta förhållande förefaller rimligt - med ökad inkomst ökar de resor som är förknippade med "positiva" och utgiftskrävande aktiviteter. Anledningen till att inköpsresorna inte ökar kan vara dels att inkomstskillnader mer avspeglar sig i vad som köps än hur ofta det köps, och dels att man för att få tid med de mer "positiva" aktiviteterna rationaliserar sina inköp. Att besöksresorna inte ökar kan avspegla en mer konstant efterfrågan på sociala aktiviteter. Det kan också vara så, att de sociala kontakterna i större utsträckning sker i samband med de andra, mer "positiva" och utgiftskrävande aktiviteterna.

I frekvensmodellerna för arbets- och tjänsteresor ingår inte några inkomstvariabler. För dessa resärenden är det också mindre rimligt att anta att inkomstskillnader ska ha ett orsakssamband med antalet resor. Däremot finns effekter som att deltidsarbetande - vilka i genomsnitt har lägre inkomster - reser mer sällan. Sådana effekter fångas istället upp med mer direkta förklaringsvariabler.

Frekvensmodellerna för tjänsteresor skiljer sig mer från modellerna för de övriga resärendena eftersom antalet resor framför allt förklaras av typen av arbetsplats och individens yrkestillhörighet.

I alla frekvensmodeller utom modellen för arbetsresor ingår logsumvariabler med signifikanta parametrar. För arbetsresor väntar vi oss inte att tillgängligheten ska ha en märkbar effekt på antalet resor som utförs. För samtliga övriga ärenden där frekvensmodeller skattats visar sig dock tillgängligheten ha en signifikant effekt på antalet resor som utförs.

Restyps- och individfördelning

Modeller för restyps- och individfördelning har skattats för inköps- samt service- och rekreationsresor. För övriga resärenden räknar vi inte med att möjligheterna till utbytbarhet mellan olika individer i hushållet är reella. Modeller för restypsfördelning (utan fördelning på individ) har även skattats för tjänsteresorna.

Vem i hushållet som genomför resorna beror på vem personen är och tillgängligheten till resmålen. Arbetstiden för förvärvsarbetande spelar också in. Ju längre arbetstid en person har, desto mindre är sannolikheten att resan görs av denna person. Om såväl mannen och hustrun har långa arbetstider, ökar sannolikheten att hustrun gör resan - och då som en arbetsplatsbaserad resa. Detta förhållande är mindre utpräglat när det gäller service- och rekreationsresor, vilket kan återspegla en mer begränsad utbytbarhet mellan personer i detta fall. Arbetstiden får här också en större effekt på restypsvalet.

För tjänsteresorna påverkas valet av restyp framför allt av när på dagen resan ska utföras och hur långt tjänstebesöket är.

I alla restypsvalsmodeller utom besöksmodellerna ingår signifikanta logsumparametrar som speglar tillgänglighetens betydelse vid valet mellan olika reskedjeformer. I allmänhet är parametrarna som avser logsumvariabler för tillgängligheten vid arbetsplatsbaserade resor osäkert bestämda, medan de som avser tillgängligheten vid bostadsbaserade resor skattats med högre precision.

Signifikanta logsumparametrar återfinns även i individvalsmodellerna för inköp samt service och rekreation.

Destinationsval

När det gäller destinationsval finns naturligtvis stora skillnader när det gäller att beskriva målpunkterna. Det visar sig också, att arbets- och besöksresorna tycks ha den storleksvariabel som bäst beskriver områdesstorleken, medan service- och rekreationsresorna har den variabel som minst tycks avspegla enbart storlek (antal elementaralternativ). Inköpsresorna ligger däremellan.

I inköps- samt service- och rekreationsresemmodellerna överskattar storleksmättet innerstadens attraktivitet, vilket i modellerna korrigeras genom särskilda variabler. En förklaring kan vara att antalet sysselsatta i den aktuella branschen kan innebära en överskattning av antalet elementaralternativ, eftersom vissa av de sysselsatta kan ha andra uppgifter (till exempel kontorsanställda som registreras med näringsgren handel). Förekomsten av sådana grupper kan antas vara särskilt stor i innerstaden, där många företag har sina huvudkontor.

Färdmedelsval

Färdmedelsvalsmodellerna innehåller i allmänhet det vidaste spektrumet av förklaringsvariabler.

När det gäller trafiksystemvariablerna, finner vi vissa skillnader mellan resärendena. En skillnad gäller de tidsvärden som kan beräknas ur modellerna. Följande åktidsvärden kan härledas:

Tabell 8.1 Tidsvärden

Ärende	Tidsvärde kr/timme
Arbete	20
Skola	16
Tjänste bil	103
Tjänste kollektivt	128
Inköp, kortvariga dagligvaruköp	24
Övriga inköp	20
Besök, service- och motion	15
Restaurang- och kulturreSOR	25

För arbetsresor redovisas i tabellen resultatet från en modell som definierats på motsvarande sätt som övriga modeller (således ej slutmodellen som har separata kostnadsparametrar för olika grupper).

Alla tidsvärden utom värdena för tjänsteresor ligger inom intervallet 15 - 25 kr/timme, och skillnaderna är inte signifikant skilda från noll (förutom tjänsteresorna jämfört med övriga resor). I detta avseende är modellerna tämligen lika, men trots dessa likheter kan parametrarna för de olika restids- och reskostnadsvariablerna vara mycket olika, och därmed ge olika effekter när modellerna tillämpas.

I nästa tabell finns därför reskostnads- och åktidsparametrar för de olika modellerna redovisade:

Tabell 8.2 Åktids- och reskostnadsparametervärden

Ärende	Åktidsparameter	Kostnadsparameter
Arbete	-0,0178	-0,0524
Skola	-0,0288	-0,1101
Tjänste bil	-0,0276	-0,0161
Tjänste kollektivt	-0,0344	-0,0161
Inköp, kortvariga dagligvaruköp	-0,0817	-0,2064
Övriga inköp	-0,0380	-0,1131
Service- och motion	-0,0246	-0,0985
Restaurang- och kulturreSOR	-0,0246	-0,0593
Besök	-0,0202	-0,0807

Spännvidden i parametervärden är betydligt större än spännvidden i tidsvärden. Detta är rimligt, eftersom modellens skala har betydelse för parametervärdena. Skalan kan dock inte skattas separat i logitmodellen (se kapitel 4), så vi kan inte utläsa i vilken utsträckning skillnaderna i parametervärden mellan de olika modellerna verkligen beror på olikheter i skala.

Det förefaller emellertid som om modellen för kortvariga dagligvaruinköp har den största skalan, följt av övriga inköp. Övriga ärenden förefaller ha en ytterligare något lägre skala. Den lägsta skalan tycks återfinnas för arbetsresorna. Samtidigt är det rimligt att anta, att vår förmåga att beskriva målpunkterna följer ungefär samma skala. Vi kan därför förmoda, men inte visa, att förmågan att beskriva målpunkterna har resulterat i den avtagande skalan för de olika resärendena ovan. Det är därför möjligt att man genom att förfina beskrivningen av olika områdens attraktivitet skulle kunna minska variansen och därmed erhålla mer likartade värden på restids- och reskostnadsparametrarna.

Att genomföra en sådan utvidgning skulle ställa stora krav på datainsamling och intervjuundersökning, och motsvarande krav vid tillämpning av modellerna. Detta får därför kanske snarast ses som en möjlig utveckling vid studier av enskilda ärendetyper (som t.ex. inköpsresor), och inte en praktisk möjlighet när syftet som här i första hand är att skapa ett trafikprognossystem.

Sammanfattningsvis kan sägas, att de olika resärendena skiljer sig åt i en rad olika avseenden, vilket återspeglas i de olika parametervärdena. Uppdelningen på de olika ärendetyperna kan därför motiveras på alla modellnivåer.

Tabell 8.3 Vikter på restidskomponenter (vikt för åktid i fordon = 1,0)

	Vänte- och bytestid	Gångtid till håll- plats
Arbete	1,1	2,9
Skola	-	-
Tjänste	1,4	1,4
Inköp, kortvariga dagligvaruinköp	-	-
Övriga inköp	2,0	1,7
Besök, service och motion	1,4	2,5
Restaurang- och kulturreSOR	1,4	2,5

Tabell 8.3 visar att vikten för vänte- och bytestid i allmänhet ligger förhållandevis lågt jämfört med många tidigare studier (se t.ex. Algers, Colliander och Widlert 1987) medan vikten för gångtid till hållplats i de flesta modellerna ligger något högre. Det bör observeras att osäkerheten i skattningarna gör det omöjligt att dra säkra slutsatser om skillnader mellan de olika resärendena.

8.8 Prognossystem

För att modellerna ska kunna användas för praktisk planering krävs att de byggs in i ett prognossystem. Detta arbete ligger vid sidan av ämnet för denna avhandling och berörs därför endast mycket översiktligt. Arbetet med prognossystemet har i huvudsak utförts av Hague Consulting Group i Holland (se Daly och Lindveldt 1992).

Principerna för prognossystemet innebär att modellerna körs med hjälp av befolkningsdata för de 850 områden som studeras. Befolkningen i varje område delas in i ett stort antal olika grupper som definieras på olika sätt för olika resärenden. Gruppindelningen kan vara efter faktorer som inkomst, hushållsstorlek, antal förvärvsarbete, etc. Resvaneundersökningen används för att bestämma de olika kategoriernas storlek. Modellerna körs därefter för varje sådan grupp i varje område.

När modellerna estimeras innehåller de ett ganska stort antal variabler av socioekonomisk natur. Alla dessa kan inte läggas till grund för särskilda kategorier när modellerna används eftersom körningstiderna då skulle bli alltför långa. Arbetet med prognossystemet har därför också inneburit att försöka identifiera de faktorer som är viktigast att använda vid kategoriindelningen. Även om vissa variabler inte alls utnyttjas (eller endast grova medelvärden används) när modellerna implementeras i prognossystemet kan de vara väsentliga att ha med vid estimeringsarbetet eftersom de kan bidra till att övriga parametrar i modellen kan skattas säkrare.

När modellerna estimeras används ett begränsat urval av alla destinationer. Även när prognossystemet körs är det lämpligt att sampla bland destinationerna. Att köra modellen för

alla kategorier i alla 850 startområdena och ta med alla 850 målområdena leder till långa körningstider. Istället körs prognoserna för alla kategorier i alla startområdena men med ett urval av destinationer. För närvarande används ca 200 destinationer som samplats så att de 100 närmaste från varje område, 50 i innerstaden samt 50 i övriga regionen tas med. Både det totala antalet alternativ och metoden att sampla kan komma att förändras vartefter mer erfarenhet av praktiska tillämpningar vinnns och vartefter datorernas hastighet ökar. Målsättningen är att hela systemet ska kunna köras på högst ett dygn (inklusive utläggning på nät).

Modellerna ger resmatriser på 850-områdesnivå. Dessa modellmatriser används för att förändra de nulägesmatriser som byggts upp. Matriserna läggs därefter ut på nät med befintliga nätutläggningsprogram.

Resultaten från systemet erhålls dels i form av tabeller och grafik från prognossystemet, dels i form av resultat från nätutläggningsprogrammen. Genom att resultaten från modellerna kan knytas till de olika befolkningskategorier som systemet arbetar med kan prognosresultaten särredovisas efter en rad olika socioekonomiska kriterier. Detta innebär att konsekvensanalyser kan utföras för en lång rad olika dimensioner.

Det är även möjligt att använda modellerna för att förutsäga hur individerna i det urval som modellerna skattas för skulle reagera på olika förändringar (så kallad Sample Enumeration). Fördelarna med detta är att fullständigt disaggregerade data används vilket eliminerar aggregeringsfelen och gör det möjligt att utföra detaljerade fördelningspolitiska analyser. Sådana körningar har utförts under projektets gång. Det praktiska problemet utgörs främst av beräkningen av de logsumvariabler som krävs för att länka samman de modelldelar som estimerats separat.

9. Överföring

Att utveckla modellsystem med den omfattning och komplexitet som systemet i detta projekt är både tidsödande och dyrbart. Modellernas användbarhet ökar naturligtvis påtagligt om modeller från en ort även kan användas på andra orter. För att detta ska vara möjligt krävs att modellen är geografiskt stabil, eller med andra ord att den är överförbar. Diskussionen i detta avsnitt tar sin utgångspunkt i arbete som vi tidigare redovisat i rapporten "Logitmodellen - Användbarhet och generaliserbarhet" (Algers, Colliander och Widlert, 1987).

9.1 Begreppet överförbarhet

Med överförbarhet avses att en modell som estimerats för en viss plats vid en given tidpunkt också kan anses vara giltig för framtida tidpunkter och för andra platser. För att en modell ska vara överförbar krävs därför att modellens parametrar är stabila över tid och rum.

Det stora intresset för modellernas överförbarhet orsakas till stor del av att tidsmässig stabilitet är en förutsättning för att modellen ska kunna användas för prognoser om framtiden.

Om modellerna är geografiskt stabila är det dessutom möjligt att spara tid och pengar genom att använda samma modeller på flera platser.

Givet att en tillräcklig grad av stabilitet finns är det också möjligt att använda erfarenheterna från tidigare modeller när nya modeller ska kalibreras. Inte minst gör detta det möjligt att förenkla datainsamlingen (genom att begränsa denna till de variabler som tidigare visat sig mest avgörande för det val som ska studeras).

En första fråga är vad som egentligen avses med att en modell är överförbar. Flera författare (Ben-Akiva 1981, Hansen 1981) menar att man kan identifiera olika nivåer på överförbarhet. Hansen identifierar fyra nivåer:

1. Teorin bakom modellerna (t.ex. nyttomaximering) är överförbar
2. Modellformuleringen är överförbar
3. Modellspecifikationen är överförbar
4. De erhållna parametrarna är överförbara

Här berör vi främst de två understa nivåerna för vilka de högsta kraven på modellerna gäller.

Det finns två principiellt olika synsätt på överförbarhet. Det första representeras av ett antal författare som har studerat överförbarhet genom att jämföra en tidigare modell med en helt nyestimerad modell och med olika statistiska test undersökt om man kan förkasta hypotesen att de kan vara identiska (givet den statistiska precision som modellerna skattats

med). Testerna anger antingen att en modell är, eller inte är överförbar. Ben-Akiva (1981) ger en översikt över dessa studier.

Bakom de statistiska testerna ligger synsättet att det skulle kunna finnas en bakomliggande "sann" modell som är giltig överallt. De genomförda studierna har givit växlande resultat. Oftast har hypotesen att modeller från olika platser skulle vara identiska kunnat förkastas.

Eftersom alla modeller i verkligheten är mer eller mindre ofullständigt specificerade kan det vara olämpligt att kategoriskt klassa en modell som överförbar eller inte.

Författare som Lerman (1981), McCarthy (1981) och Gunn och Pol (1985) argumenterar för en mer pragmatisk ansats. Eftersom det inte är möjligt att utveckla en "perfekt" modell står valet mellan modeller med större eller mindre grad av osäkerhet och approximation.

Givet att det inte finns någon "sann" modell har de statistiska testerna betydande svagheter. Om datamaterialen är tillräckligt små (och precisionen i estimaten därmed är tillräckligt liten) kommer testerna att acceptera hypotesen om överförbarhet. Om å andra sidan datamaterialen är tillräckligt stora och av tillräckligt god kvalitet kommer hypotesen alltid att kunna förkastas.

Louviere (1981) har påpekat risken med att dra generella slutsatser om modellers överförbarhet utifrån erfarenheter med modeller som är behäftade med ofullkomligheter både vad avser specifikation och använda indata. Louviere menar att det egentligen ännu är alltför tidigt att dra slutsatser om modellernas överförbarhet. Innan detta kan ske krävs en större enhetlighet hos de modeller som utvecklas.

Louvieres synpunkter understryks av det arbete som redovisas av Algers, Colliander och Widlert (1987). Betydelsen av goda indata kan inte nog understrykas. Många bristfälliga modellresultat förklaras helt av sådana problem. Självklart kan man inte vänta sig att en modell som egentligen inte fungerar särskilt bra i den omgivning där den är estimerad ska fungera i en helt annan omgivning.

Med det mer pragmatiska synsättet är den avgörande frågan om den befintliga modell som finns tillgänglig kan ge någon information som kan vara till nytta i en ny situation. Frågan är alltså inte om den gamla modellen är "riktig" i den nya situationen, utan om den kan förbättra möjligheterna att utföra prognoser i den nya situationen. Svaret på överförbarhetsfrågan kommer då även att bero på hur god information som finns i den nya situationen. Finns god information, och kanske till och med nya modeller, kan en modell från andra förhållanden försämma prognoserna. Finns endast mycket bristfällig information kan även en grov modell vara till stor nytta.

Det är således rimligt att anta att modeller i större eller mindre utsträckning alltid är beroende av den omgivning där de är estimerade. Det är en utopi att tro att samtliga de variabler som påverkar valet ska kunna fångas in i modellen. Dessa bakomliggande variabler (vilka alltså påverkar parametervärdena för de variabler som ingår i modellen) kommer att skifta från plats till plats. Hur statistiskt säkert en viss modell än är skattad kommer man därmed alltid att introducera en extra osäkerhet när modellen appliceras i en annan omgivning. En viktig del i en överförbarhetsstudie är därför att försöka utnyttja information från den nya omgivningen (platsen och/eller tidpunkten) för att anpassa modellen och därmed minska osäkerheten.

9.2 Tidigare erfarenheter

De tidigaste studierna inom området behandlade överförbarhet av enkla färdmedelsvalsmodeller (se t.ex. översikten i Ben-Akiva 1981 eller Daly 1985). Med det pragmatiska synsättet som diskuterats ovan visade resultaten i allmänhet att överföring av modeller var möjlig och att den gav värdefulla resultat på de nya orterna.

Tretvik (1989) studerar möjligheterna att föra över färdmedelsvalsmodeller mellan olika norska städer. Algers, Colliander och Widlert (1987) studerar överföringar av främst färdmedelsvalsmodeller, men även en modell för samtidigt val av färdmedel och färdmål, mellan svenska städer. Båda studierna visade att alternativens konstanter bör kalibreras om när modellerna flyttas och att trafikstandardvariabler är mer överförbara än andra variabler som förklarar färdmedelsvalet.

Det finns få studier som behandlar överföring av hela modellsystem. Gunn och Pol (1985) rapporterar en studie där ett komplett modellsystem flyttas från en region till en annan. Systemet var de så kallade Zuidvleugelmodellerna (1977-1981) som byggdes upp med data för området runt Utrecht. Slutsatserna av studien var att det är möjligt att flytta ett helt modellsystem, förutsatt att nya konstanter estimeras och att skalan på övriga parametrar anpassas. Separata skalfaktorer för trafikstandardvariabler och övriga variabler förbättrade resultaten enbart marginellt.

Widlert (1989) studerade överflyttning av en kombination av befintliga svenska modeller och Zuidvleugelmodeller till Helsingborg. Ett komplett system av färdmedels-, destinations- och frekvensmodeller för olika resänder byggdes upp med en kombination av överförda modelldelar och nyestimering av konstanter, skalfaktorer och logsummer. Resultaten visade på vikten av att skatta nya konstanter, att trafikstandardvariabler generellt sett har en god överförbarhet men att parametrarna för andra variabler kan avvika påtagligt.

9.3 Stockholmssystemets överförbarhet

De begränsade erfarenheterna av överföring av hela modellsystem tyder på att det är möjligt att tillämpa de modeller som utvecklats i detta projekt på andra orter. Tidigare erfarenheter tyder som diskuterats ovan på att trafikstandardvariabler, det vill säga parametrar för olika restidskomponenter och kostnader, bör vara väl överförbara till andra platser. Även parametrar för socioekonomiska variabler bör gå att överföra. Överförbarheten hos de parametrar som direkt speglar ortsspecifika förhållanden ("city-dummys", variabler för stormarknader) är naturligtvis sämre.

Modellerna i detta projekt är dock betydligt mer komplicerade än de som tidigare överförts och många av de samband som modellen täcker in kan se olika ut på olika orter. Modellerna behandlar explicit interaktioner och restriktioner inom hushållen. Sådana samband kan förväntas vara rimligt överförbara inom landet, men det är ännu en öppen fråga hur väl överförbara sambanden är mellan olika länder.

En inskalning på disaggregerade data kräver hushållsundersökningar av tillräcklig storlek. En fullständig test av överförbarhet kräver lika mycket data som en nyutveckling av modeller. Sådana tester är därför dyrbara och tidskrävande att utföra.

Under 1992 kommer delar av Stockholmsmodellerna att föras över till Trondheim i samband med projektet "Överföring av modeller i Norden" som finansieras av Nordiska kommittén för transportforskning, NKTF. Överföringen gäller i första hand arbetsresedelen av modellstrukturen. Projektet är av stort intresse eftersom Trondheim representerar en ort av avvikande storlek och dessutom ligger i ett annat land.

10. Slutsatser och generella kommentarer

10.1 Allmänna slutsatser

Vårt arbete har styrts av ett antal hypoteser om vad som påverkar individernas och hushållens beteende. De viktigaste av dessa formulerades ganska exakt i form av föreslagna modellstrukturer redan i Planeringsrapport 5 (Algers och Widlert 1986a). Hypoteserna handlar om att trafiksystemet och tillgängligheten har betydelse för resmönstret på fler sätt än vad som tidigare kunnat visas empiriskt och att flera viktiga val rör hela hushållet och att hushållssamband därför bör modelleras explicit.

Generellt kan vi konstatera att våra hypoteser har bekräftats. Detta tar sig uttryck i att de modellstrukturer som erhållits överensstämmer mycket väl med dem som föreslogs i planeringsrapporten. De parametrar som direkt speglar om våra hypoteser styrkts eller ej är bland annat de logsumparametrar som för vidare tillgänglighetens betydelse genom modellstrukturen. Dessa parametrar har kunnat estimeras framgångsrikt. Det är särskilt uppmuntrande att detta varit möjligt trots att modellernas komplexitet (främst antalet alternativ) tvingat oss att dela strukturerna på flera ställen och estimerat dessa delar sekventiellt. Att logsumparametrarna blivit signifikanta även i de delmodeller som behandlar hushållsinteraktioner visar på existensen av dessa val och att de påverkas av trafiksystemets utformning.

Andra indikationer på att modellarbetet lyckats är att modellerna innehåller de viktigaste förklaringsvariablerna som vi utifrån tidigare genomförd forskning har anledning tro påverkar trafikanternas beteende, samt att de dessutom innehåller ett vitt spektrum av förklaringsvariabler vars betydelse för beteendet kan förklaras a priori.

10.2 Modellsystemets svaghet och styrka

Processen

Modellsystemets fullständighet och realism har som baksida att det behövs omfattande bearbetningsprogram för att generera indata i form av alternativmängder och variabler och omfattande styrfiler till det estimeringsprogram som använts. Risken för fel är betydande - en stor del av arbetet har bestått i att rätta fel. Arbetet blir därigenom både tidsödande och mödosamt.

Under projektet har vi haft en mycket ovanlig möjlighet att lägga ner tid på att genomföra arbetet med den ambitionsnivå vi själva önskat. Vi har dessutom själva haft kontroll över hela processen, ända från planering av datainsamlingen till genomförandet av modellutvecklingen. Det har även varit möjligt att gå tillbaka till de enskilda intervjuformulären när utvärderingen av modellerna har visat på observationer vars beteende modellerna inte kunnat förklara. Inte sällan har detta lett till att bearbetningsprogram kunnat rättas eller att indata kunnat korrigeras.

Modellsystemet

Även om modellsystemet är omfattande och komplext innebär den realiserade strukturen avsevärda förenklingar i beskrivningen av beteendet.

När systemet beskriver hur hushållen formar olika typer av reskedjor antas arbetsresorna ligga fast och varken ändra målpunkt eller färdmedel. Endast kopplingen mellan arbetsresor och *ett* ytterligare ärende beskrivs i modellerna.

Kopplingen mellan de olika resärendena är förhållandevis svag i systemet. Endast via modellen för sekundära destinationer finns en koppling mellan arbetsresor och andra ärenden. Bilinnehavsmodellen påverkar tillgängligheten till färdmedel i alla modeller. Arbetsresornas tidsmässiga förläggning behandlas som fix i systemet och ger en förutsättning för möjligheterna att utföra olika ärenden för de olika hushållsmedlemmarna. Tidsbudgetrestriktionerna som sådana behandlas även de grovt i systemet. Det har inte funnits detaljerad information om butikernas öppettider utan genomsnittsvärden har använts vid beräkningen av möjliga alternativ. Motsvarande begränsning gäller övriga resärenden.

Även destinationsvalet behandlas grovt i modellerna. Valet av destination förklaras genomgående med *storleksmått* som primärt mäter antalet "elementaralternativ" som finns i respektive målområde. Modellerna innehåller däremot få mått på *kvaliteten* på utbudet i målområdena.

Fortsatt forskning inom dessa områden är både angelägen och möjlig. I vårt projekt har en övergripande restriktion varit att modellerna även ska gå att använda praktiskt. Denna restriktion innebär att de variabler som används måste finnas tillgängliga på områdesnivå när modellerna används och dessutom måste gå att prognosera. Restriktionen styr dessutom - i än högre grad - vilka modellstrukturer som är möjliga att förverkliga. Mycket omfattande ömsesidiga beroenden mellan resärendena leder snabbt till ohanterliga strukturer. Släpps kravet att utveckla praktiskt användbara modeller och intresset istället koncentreras mer till att förstå och förklara individernas och hushållens beteende är det redan idag möjligt att gå betydligt längre än vi gjort.

Som framgått är modellstrukturerna komplexa, en komplexitet som försvaras av våra empiriska resultat. Modellerna innehåller även ett osedvanligt stort antal förklaringsvariabler som kunnat visas bidra till att förklara individernas och hushållens beteende. Vi kan därför förvänta oss att specifikationen som sådan ger en mer realistisk beskrivning av trafikanternas beteende än tidigare modeller. En avigsida med denna komplexitet är dock att modellsystemet ställer stora krav på indata när det ska användas. Flera författare (Alonso 1968, Daly och Ortuzar 1990) har pekat på att de totala felen i modellen visserligen sjunker när specifikationen görs allt mer komplex, men att inverkan av mätfelen i variablerna samtidigt ökar när dessa blir fler. Slutsatsen har varit att det finns en "optimal" avvägning mellan komplexitet och mätfel - en avvägning som påverkas av kvaliteten på de datakällor som står till förfogande. I detta projekt har vi haft resvanedata av osedvanligt hög kvalitet. Vi har dessutom haft trafikstandarduppgifter med hög detaljeringsgrad och hög kvalitet. Slutligen är även tillgången till kompletterande information (t.ex. uppgifter om förekomst av olika aktiviteter i olika områden) generellt sett bättre i Sverige än i de flesta andra länder. Det kan därför hävdas att det i Sverige är meningsfullt att utveckla mer komplexa

system än i de flesta andra länder. Frågan om vår komplexitet är den "optimala" är dock omöjlig att besvara empiriskt utan att utveckla enklare modellsystem parallellt.

Det i föreliggande avhandling presenterade modellsystemet är oss veterligt det mest fullständiga som hittills utvecklats någonstans i världen. Systemet behandlar en stor del av det totala resandet och framför allt behandlas fler valbeslut än vad som tidigare varit möjligt. Bilalockeringsmodellerna är de första i sitt slag (såväl som på de förstudier som gjordes inför projektet). Modellerna för sekundära destinationer är de första som implementeras i ett system. Interaktionerna mellan olika hushållsmedlemmar vid valet av arbetsplats har inte behandlats tidigare. Reststypvalsmodellerna i tjänste- och icke-arbetsresemodellerna är nya, liksom modellerna för hur hushållet väljer vem eller vilka som ska utföra inköpsresor. Systemet har sammanfattningsvis sin främsta styrka i att det behandlar individernas och hushållens resande mer fullständigt och realistiskt än vad som tidigare varit möjligt. Samtidigt är det ett system som kan användas - och redan används - för praktisk trafikplanering.

11. Referenser

- Algers, S. (1974). Prognosmodell för bilinnehav, LAKU-information nr 4. Stockholms läns landsting.
- Algers, S. (1982). The Calibration and Application of Travel Demand Models in Sweden. I "Transportation Planning Models", ed Florian, M. North Holland, Amsterdam.
- Algers, S. (1985). Stockholmer Verkehrsbefragung 1985. Haushaltsbefragungen zum Verkehrsverhalten, B85 Schriftenreihe der Deutschen Verkehrswissenschaftlichen Gesellschaft e.v.
- Algers, S. (1991). Prognosmodeller för intercityresor, teknisk rapport. Transek AB.
- Algers, S. Daly, A. Widlert, S. (1990). The Stockholm Model System - Travel to Work. Fifth World Conference on Transport Research, Yokohama, Japan.
- Algers, S. Daly, A. Widlert, S. (1991). Modelling Travel Behaviour to Support Policy Making in Stockholm. 6th International Conference on Travel Behaviour, Quebec, Canada.
- Algers, S. Daly, A. Widlert, S. (1992). The Stockholm Model System - Shopping Trips. Sixth World Conference on Transport Research, Lyon, Frankrike.
- Algers, S. Colliander, J. Widlert, S. (1987). Logitmodellen - användbarhet och generaliserbarhet. Byggnadsrådet, rapport R30:1987.
- Algers, S. Tegnér, G. Widlert, S. (1984). Färdmedelsval för arbetsresor i Göteborg. Göteborgs Spårvägar
- Algers, S. Widlert, S. (1982). Förslag till prognosmodellsystem. Landstingets trafikkontor.
- Algers, S. Widlert, S. (1983). Modellsystem - planeringsrapport 1. Landstingets trafikkontor, AB Storstockholms Lokaltrafik.
- Algers, S. Widlert, S. (1985). Applicability and Stability of Logit Models in Sweden - Some Recent Findings with Policy Implications. 4th International Conference on Travel Behaviour, Noordwijk, Holland.
- Algers, S. Widlert, S. (1986a). Reviderat modellsystem - planeringsrapport 5. Landstingets trafikkontor, AB Storstockholms Lokaltrafik.
- Algers, S. Widlert, S. (1986b). Hushållens val av antal bilar, typ av bil och allokering av bilarna inom hushållet. Transek AB.

- Algers, S. Widlert, S. (1987a). Restricted Car Availability and Mode Choice - A Household Interaction Phenomenon. 5th International Conference on Travel Behaviour, Aix-en-Provence, Frankrike.
- Algers, S. Widlert, S. (1987b). Trafikplanering med logitmodeller - sammanfattning av svenska erfarenheter. Byggforskningsrådet, Rapport R29:1987.
- Alonso, W. (1968). Predicting best with imperfect data. *Journal Am. Inst. Planners*, 34(3).
- Anas, A. (1983). Discrete Choice Theory, Information Theory and the Multinomial Logit Gravity Models. *Transportation Research* 17B, sid 13-33
- Antonisse, R, Daly, A. Ben-Akiva, M (1989). A Highway Assignment Method Based on Behavioural Models of Car Drivers' Route choice. Transportation Research Board, 68th Annual Meeting, Washington D.C.
- Antonisse, B. Daly, A. Gunn, H (1986). The Primary Destination Tour Approach to Modelling Trip Chains. PTRC Summer Annual Meeting.
- Anderstig, C. Mattsson, L-G. (1991). An Integrated Model of Residential and Employment Location in a Metropolitan Region. *Papers in Regional Science: The Journal of the RSAI*.
- Axhausen, K. Gärling, T. (1992). Activity-Based Approaches to Travel Analysis: Conceptual Frameworks, Models and Research Problems. *Transportation Reviews* (under tryckning).
- Ben-Akiva, M. (1973). Structure of Passenger Travel Demand Models. Massachusetts Institute of Technology.
- Ben-Akiva, M. (1981). Issues in Transferring and Updating Travel Behaviour Models. I "New Horizons in Travel Behaviour". Lexington books, Lexington
- Ben-Akiva, M. Boccara, B. (1987). Integrated Framework for Travel Behaviour Analysis. 5th International Conference on Travel Behaviour, Aix-en-Provence, Frankrike.
- Ben-Akiva, M. Bolduc, D. (1991). A Multinomial Probit Formulation for Large Choice Sets. 6th International Conference on Travel Behaviour, Quebec, Canada.
- Ben-Akiva, M. Lerman, S. (1985). Discrete Choice Analysis - Theory and Application to Travel Demand. MIT Press, Cambridge, Massachusetts.
- Berglund, C-O. Tegnér, G. Widlert, S. (1977). Val av resmål och färd sätt vid inköpsresor - en beteendestudie. Byggforskningsrådet, Rapport R8:1977, Stockholm.
- Boyce, D. Chon, K. Lee, Y. Lin, K. LeBlanc, L. (1983). Implementation and Computational Issues for Combined Models of Location, Destination, Mode and Route Choice. *Environment and planning A* 15, 1219-1230.
- Brög, W. Meyburg, A. Wermuth, M. (1983). Development of Survey Instruments Suitable for Determining Nonhome Activity Patterns. *Transportation Research Record* 944.

- Bruzelius, N. (1978). *The Value of Travel Time - Theory and Measurement*. Skrift Nr 1978:3, Nationalekonomiska Institutionen, Stockholms Universitet
- Cigno, A. (1991). *Economics of the Family*. Oxford University Press, Oxford, England.
- Cambridge Systematics Incorporated (CSI). (1980). *Travel Model Development Project: Phase 2 Final Report*. Metropolitan Transportation Commission, Berkeley, Californien.
- Daganzo, C. (1979). *Multinomial Probit, The Theory and its Application to Demand Forecasting*. Academic Press, New York.
- Daly, A. (1982). *Estimating Choice Models Containing Attraction Variables*, Transpn Research 16B sid 5-15.
- Daly, A. (1985). *A Study of Transferability of Disaggregate Mode Choice Models from Grenoble to Nantes*. Cambridge Systematics, Haag.
- Daly, A. (1987). *Estimating "Tree" Logit Models*, Transpn Research 21B sid 251-267.
- Daly, A. Gunn, H. (1985). *Cost-effective Methods for National-level Demand Forecasting. I "Behavioural Research for Transport Policy"*. VNU Science Press, Utrecht, Holland 1986.
- Daly, A. Lindveldt, K. (1992). *Stockholm Demand Prediction Model - Technical Documentation*. Hague Consulting Group, Haag.
- Daly, A. Ortuzar, J.D. (1990). *Forecasting and Data Aggregation: Theory and Practice*. Traffic Engineering and Control, december 1990.
- Daly, A. van Zwam, H. (1983). *Travel Demand Models for the Zuidvleugel Study*. PTRC Summer Annual Meeting.
- Emme/2 user's manual. INRO consultants, Montréal 1991.
- Florian, M. (1982). *An Introduction to Network Models Used in Transportation Planning*, I "Transportation Planning Models", ed Florian, M. North Holland, Amsterdam.
- Fridén, L. Odén, N. Scheele, S. (1985). *Ekonomi, sysselsättning och befolkning i tillväxt respektive stagnation*. Underlagsmaterial till Skiss 85, Regionplanekontoret, Stockholms Läns Landsting.
- Gaudry, M. Wills, M. (1978). *Estimating the Functional Form of Travel Demand Models*. Transportation Research. Vol 12.
- de V. Graaf, J. (1967). *Theoretical Welfare Economics*. The Syndics of the Cambridge University Press, London.
- Gunn, H. Pol, H. (1985). *Model Transferability: The Potential for Increasing Cost Effectiveness*. 4th International Conference on Travel Behaviour, Noordwijk, Holland.

- Hansen, S. (1980). Utveckling av en lågtrafikmodell för Stockholm. Trafiknämnden, Stockholms läns landsting.
- Hansen, S. (1981). In Favor of Cross-Cultural Transferability of Travel Demand Models. I "New Horizons in Travel Behaviour". Lexington Books, Lexington.
- Hau, T. (1987). Using a Hicksian Approach to Cost-Benefit Analysis in Discrete Choice: An Empirical Analysis of a Transportation Corridor Simulation Mode. *Transportation Research Part B* Vol. 21B No. 5 1987 pp 339-357.
- Hensher, D. (1989). Behavioural and Resource Values of Travel Time Savings: A Bicentennial Update. *Australian Road Research* Vol. 19 Nr 3 1989
- Horowitz, J. Sparmann, J. Daganzo, C. (1982). An Investigation of the Accuracy of the Clark Approximation for the Multinomial Probit Model. *Transportation Science*, Vol. 16, nr 3.
- Hägerstrand, T. (1970). What About People in Regional Science? I "Papers of the Regional Science Association", Vol 24 sid 7-21.
- Hägerstrand, T. Lenntorp, B. Mårtenson S. Jenstav, M. Wallin, E. (1974). Ortssystem och levnadsvillkor. I SOU 1974:2.
- Jara-Diaz, SR. Videla, J. (1989). Detection of Income in Mode Choice: Theory and Application. *Transportation Research Part B*, Vol. 23B NO. 6 1989 pp 393-400.
- Jones, P. Dix M. Clarke, M. Heggie, J. (1980). *Understanding Travel Behaviour*. Oxford University, Oxford.
- Kmenta, J. (1986). *Elements of Econometrics*, second edition New York, Macmillan.
- Kollo, H. Purvis, C. (1988). *Regional Travel Forecasting Model System for the San Francisco Bay Area*. Metropolitan Transportation Commission, California.
- Lam, S-H. Mahmassani, H. (1991). Multinomial Probit Model Estimation: Computational Procedures and Applications, 6th International Conference on Travel Behaviour, Quebec, Canada.
- Lerman, S. (1981). A Comment on Interspatial, Intraspatial and Temporal Transferability. I "New Horizons in Travel Behaviour". Lexington books, Lexington.
- Louviere, J. (1981). Some Comments on Premature Expectations Regarding Spatial, Temporal and Cultural Transferability of Travel Choice Models. I "New Horizons in Travel Behaviour". Lexington books, Lexington.
- Luce, R. Suppes, P. (1965). Preference, Utility and Subjective Probability. I "Handbook of Mathematical Psychology". Vol 3. Wiley, New York.

Lundberg, S. (1982a). Resvaneundersökning i Jönköping - utförlig resultatrapport. Byggforskningsrådet, Rapport R94:1982.

Lundberg, S. (1982b). Resvaneundersökningar - metodbeskrivning. Byggforskningsrådet, Rapport R93:1982.

Lundberg, S. (1982c). Resvaneundersökningar - varför och hur? Byggforskningsrådet, Rapport R92:1982.

Lundgren, J. (1989). Optimization Approaches to Travel Demand Modelling. Linköping Studies in Science and Technology. Dissertations no. 207. Linköpings universitet.

Malinvaud, E. (1972). Lectures on Microeconomic Theory. North Holland, Amsterdam.

McCarty, P.S. (1981). Further Evidence on the Temporal Stability of Disaggregate Travel Demand Models. Department of Economics, Purdue University.

McFadden D. (1974). Conditional Logit Analysis of Qualitative Choice Behaviour. I "Frontiers of Econometrics". P Zarembka, ed. Academic Press, New York, sid 105-142.

McFadden, D. (1978). Modelling the Choice of Residential Location. I Karlquist et al, red, "Spatial Interaction Theory and Residential Location". North Holland Amsterdam, sid 75-96.

McFadden, D. (1981). Econometric Models of Probabilistic Choice. I Manski, C. och McFadden, D. "Structural Analysis of Discrete Data with Econometric Applications", MIT Press, Cambridge, Massachusetts.

McFadden, D. Talvitie, A. (1977). Demand Model Estimation and Validation. The Urban Travel Demand Forecasting Project. Phase 1 final report series, Volume V. Berkeley, Californien.

Morikawa, M. Ben-Akiva, M. McFadden, D. (1990). Incorporating Psychometric Data in Econometric Travel Demand Models. Paper till Transportation Science.

Mouwitz, J. (1983). RVU 80/81 - Resandet över dagen-veckan-året. AB Storstockholms Lokaltrafik.

Mouwitz, J. (1991). Så reser vi i Stockholms län - Invånarna och pendlingen. Regionplane- och trafikkontoret, AB Storstockholms Lokaltrafik.

Mouwitz, J. (1992). Teknisk rapport för individundersökningen. Regionplane- och trafikkontoret, AB Storstockholms Lokaltrafik.

Nylander, P. (1992). Inskalning av trafikmodeller i Örebro. Transek AB.

Nylander, P. Widlert, S. (1991). Inskalning av Fredrik-modeller på Oslo-data. Transek AB.

- Ortuzar, JD. Willumsen, LG. (1990). *Modelling Transport*. John Wiley & Sons Ltd. Chicester, England.
- Payne, J. Bettman, J. Johnson, E. (1988). Adaptive Strategy Selection in Decision Making. *Journal of Experimental Psychology: Learning, Memory and Cognition* No 3 534-532.
- Planeringsrapport 3. (1985). Provundersökning 1. Landstingets trafikkontor, AB Storstockholms lokaltrafik.
- Planeringsrapport 4. (1985). Provundersökning 2 och 3. Landstingets trafikkontor, AB Storstockholms lokaltrafik.
- Ramjerdi, F. (1990). Cost-Benefit Analysis and Distributional Consequenses of an Area Licensing Scheme for Stockholm. *Arbeidsdokument, Transportökonomisk Institut 1990, Oslo.*
- Rosenqvist, G. (1991). Så reser vi i Stockholms län - Teknisk rapport och en jämförelse mellan hushållsundersökningen och individundersökningen. *Stockholms läns landsting, AB Storstockholms Lokaltrafik.*
- Råbäck, G. (1991). Automatic Traffic Counts in Stockholm. *Storstockholms Lokaltrafik.*
- Resvaneundersökningen 1984/1985. (1984). SCB. Statistiska meddelanden T11 SM 8701.
- Schwertner, B. Fallast, K. Katteler, H. Sammer, G. (1983). Selected Results of a Standardized Survey Instrument for Large-Scale Travel Surveys in Several European Countries. *New Survey Methods in Transport, 2nd International Conference, Australien.*
- Stockholmstrafiken rapport nr 93. (1985). *Stockholms gatukontor.*
- Theil, H. (1971). *Principles of Econometrics*. Wiley, New York.
- Thurston, L. (1927). A law of comparative judgement. *Psychological review* 34:273-286.
- Tomth, J-E. (1992). Så reser vi i Stockholms län - Resvanor, tidsanvändning och besöksmönster. *Regionplane- och trafikkontoret, AB Storstockholms Lokaltrafik.*
- Tretvik, T. (1989). Logitmodeller för transportplanlegging - Matematiska modeller för reiseatferd basert på kvalitativ valghandlingsteori. *Norges Tekniske Høgskole, Trondheim.*
- TU 71 - Trafikundersökningar i Stockholmsregionen. (1974). *Stockholms Läns Landsting, Trafiknämnden.*
- Widlert, S. Tegner, G. Berglund, C-O. (1977). Val av resmål och färdssätt vid inköpsresor - En beteendestudie. *Statens Råd för Byggnadsforskning, Rapport R8:1977.*
- Widlert, S. (1989). Resbehov med buss i tätort - Teknisk rapport. *Transek AB.*
- Widlert, S. (1990). Trafikanternas krav på tågresor - Sammanställning av resultat från olika undersökningar. *Transek AB.*

Williams, H.C.W.L. och Senior, M.L. (1977). Model based transport policy assessment: (2) Removing fundamental inconsistencies from the models. *Traffic Engineering and Control* 18(10), 464-9.

Young, W. Brown, H.P. (1983). A Revealed Importance Elimination-by-Aspects Model of Mode Choice. PTRC Summer Annual Meeting.


Zuidvleugel-studien. (1977-1981). Cambridge Systematics, Rapport 1-7, Haag.

Denna rapport R36:1992 som utges av Bygghorskningsrådet är en av tre rapporter som tillsammans utgör författarnas avhandlingar vid Institutionen för Trafikplanering, KTH, samt slutrapportering av projektet. Projektet har gemensamt finansierats av Bygghorskningsrådet, Transportforskningsberedningen, AB Storstockholms Lokaltrafik och Stockholms Läns Landsting Regionplane- och Trafikkontoret.



KUNGL
TEKNISKA
HÖGSKOLAN

TRAFIKPLANERING

Transportforskningen 



AB Storstockholms Lokaltrafik



Regionplane- och trafikkontoret

Art.nr: 6812036

Abonnemangsgrupp:
X. Samhällsplanering

Distribution:
Svensk Byggtjänst
171 88 Solna

Cirkapris: 90 kr exkl moms

R36:1992

ISBN 91-540-5488-5

Bygghorskningsrådet, Stockholm