TØI report 766/2005
Authors: Anne Madslien
Jens Rekdal
Odd I Larsen
Oslo 2005, 88 pages Norwegian language

Summary:

Regional travel demand models for Norway

Introduction

Molde Research Institute and the Institute of Transport Economics (TØI) have developed a model system for regional travel demand in Norway, commissioned by the Norwegian National Public Roads Administration, the Norwegian National Rail Administration, AVINOR (AVIation NORway), and the Norwegian Coastal Administration. The Ministry of Transport and Communications has provided supplementary funding. A national long distance model (for trips longer than 100 km one way) has earlier been developed in Norway, and the last version, exhibiting 1400 zones nationwide, was completed and implemented in 2002. The newly developed regional model system covers trips shorter than 100 km one way.

Molde Research Institute and TØI have been responsible for determining the overall model structure, processing the data, and estimating and implementing all of the different submodels. The clients were responsible for providing the data input to the estimation process. The main tasks of the data collection were to conduct a comprehensive travel survey in Norway (NTS2001), to establish the network models (which includes coded road network and public transport routes for the whole country) and to produce data reflecting level of service of the networks. Some of this work was contracted out to consultants.

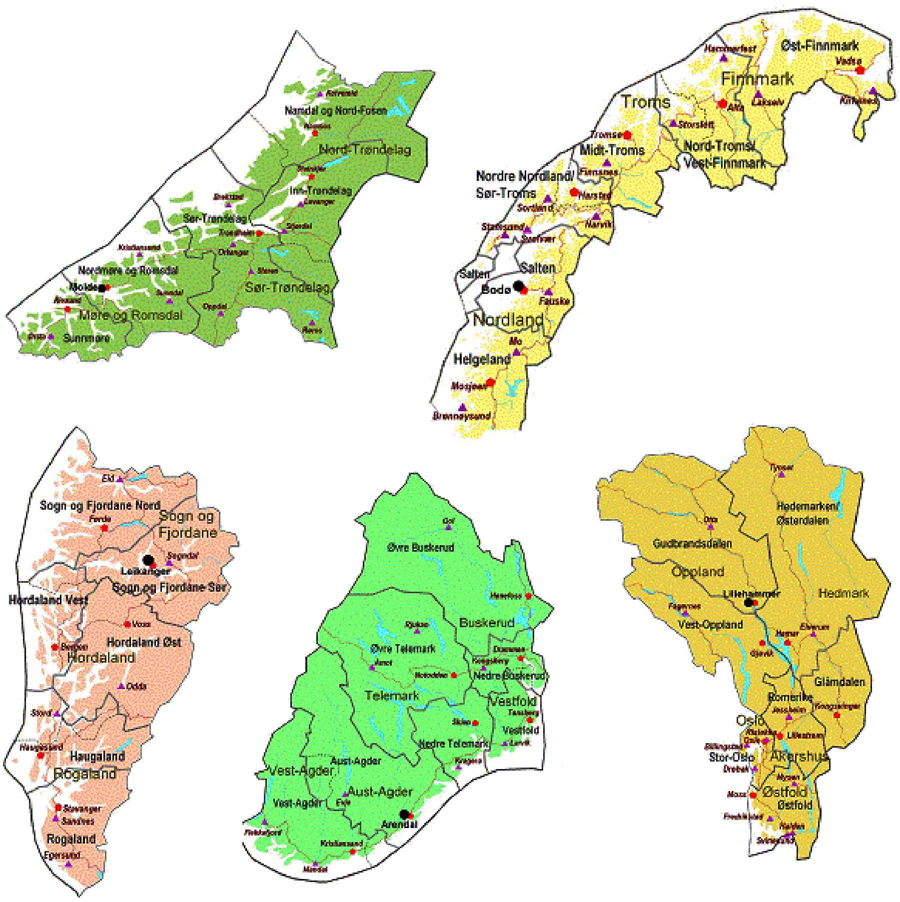
The Model System

The model system consists of five regional transport models, together covering all of Norway (figure 1). The models basically have the same structure, but they use different data sets and include a limited set of model specific geographic dummy variables.

A model run normally starts by producing *Level of Service (LoS) matrices* from the network models. These are matrices containing information about travel times, distances and costs between all zones in the region and for all modes of transport.

The report can be ordered from:
Institute of Transport Economics, PO Box 6110 Etterstad, N-0602 Oslo, Norway
Telephone: +47 22 57 38 00   Telefax: +47 22 57 02 90

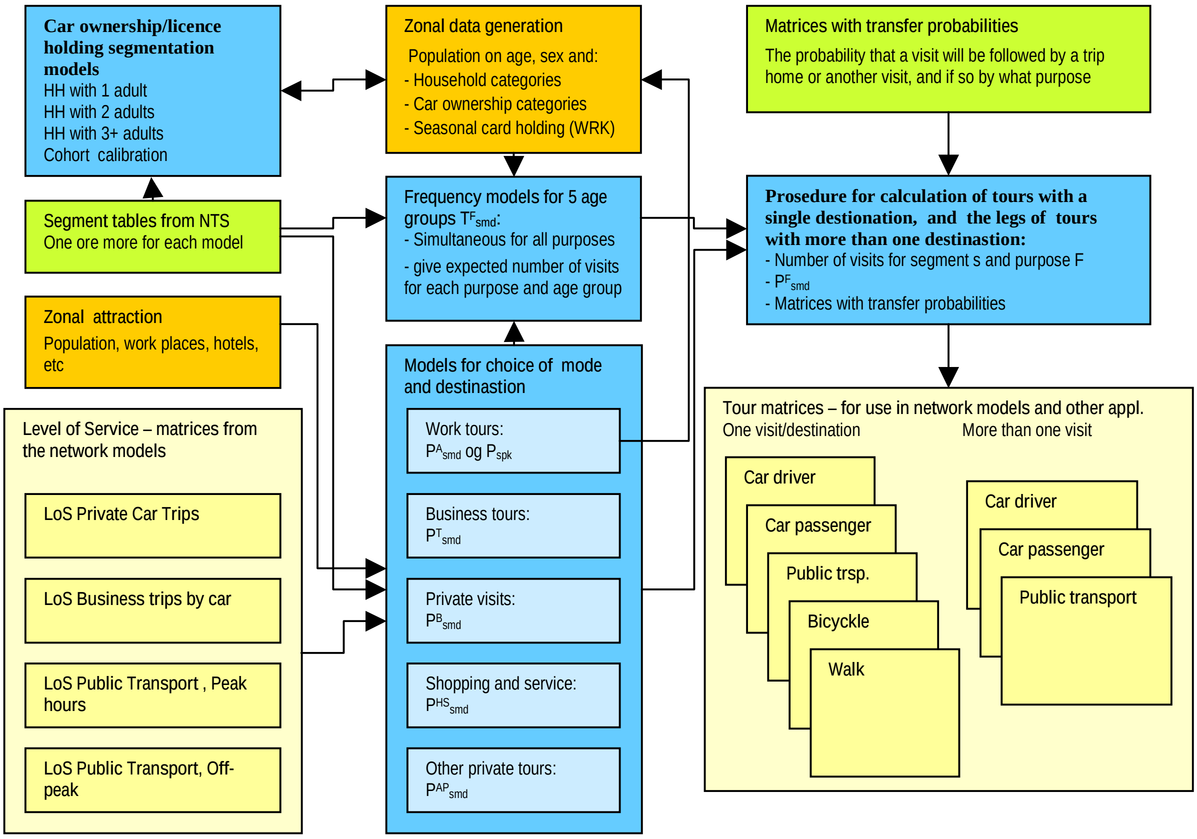
Figure 1. The five regions in Norway



TØI-report 766/2005

The general structure of the system of models is shown in figure 2.The LoS-matrices are input to a run of the transport demand model. Other inputs are *demographic zonal data* for generation of trips, zonal data for attraction of trips and different *segment tables* from the National Travel Survey (NTS2001). These tables give average values for segments of the population for variables where zonal data do not exist in any public register. Data describing the *almost 14 000 zones nationwide* were mainly supplied by Statistics Norway.

Figure 2. Structure of the model system



TØI-report 766/2005

Each of the models will be described in short in the following sections.

Car ownership and driver license holding segmentation models

The demographic data available for the model system partitions the population in the zones by age (5-year intervals), gender and (three) household categories:

- households with one adult (age 18 +)
- households with two adults
- households with three or more adults

The *car ownership and license holding situation* (together forming the *car availability* information used by other submodels) in the household are important aspects of an individual's possibility to make certain trips. Changes in car availability over time are probably also quite important for the overall increase in car traffic. Therefore, we wanted our models to capture the effects of changes in car ownership and license holding. The purpose of these models is to segment the population further than age, sex, and household types, so that information about the car availability for different population groups can be used in all parts of the model system. Every segment, as defined by gender, age group and household type in a zone, is subdivided into five car availability segments:

S1: Individuals without driver's license and with no car in the household.

S2: Individuals without driver's license, but with one or more cars in the household.

S3: Individuals with driver's license, but with no car in the household.

S4: Individuals with driver's license and at least as many cars as licenses in the household.

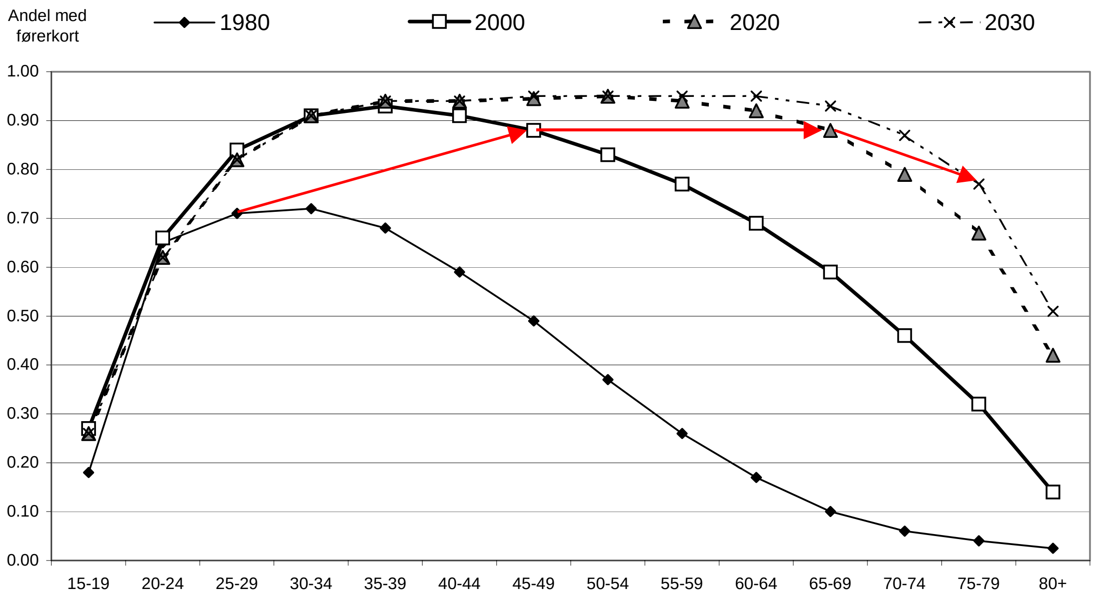
S5: Individuals with driver's license but with less cars than licenses in the household.

People in the various segments will have unequal possibilities of making a trip by car, e.g. individuals in segment S4 have full car availability while those in segment S1 will have no access to cars as driver and low access to cars as passenger.

For each of the three household categories, models splitting the population between the five car availability segments were estimated and implemented. The estimation was done by means of a maximum likelihood procedure coded in GAUSS. "Utility functions" were specified consisting of alternative specific constants and variables describing the zone of living and characteristics of the individual and the household (data from the National Travel Survey 2001). Important variables are age, sex, family type, population density and household income.

When implemented the models were calibrated to base forecasts (calculated by use of cohort effects) for license holding in the years 2010, 2015, 2020, 2025, 2030 and 2040. For women these temporal effects are illustrated in figure 3.

Figure 3. Cohort effects in driver license holding. License holding rate for women.



TØI-report 766/2005

The red arrows between the curves indicate how the rate of license holding for women who were in the age group 25 to 29 in 1980, will change over time. In 2000 these same women were between 45 and 49 years old, exhibiting a significant increase in license holding rate from 1980. From then on these women's driver's licence penetration rate does not change much until 2020. Then it decreases, but the rate is still higher in 2030 than in 1980. When the models are used for long-term traffic forecasts the calibration will take care of these cohort effects.

Models for choice of destination and mode

Five models have been estimated to simulate the simultaneous *choice of mode and destination*, one model for each of the travel purposes:

- work trips
- business trips
- private visits
- shopping and service trips
- other private purposes

The main data source for estimation of the models has been the National Travel Survey 2001, where some 20 000 respondents were interviewed. This data source was supplemented with data from a travel survey in the Oslo region that was carried out in the same period and with a very similar questionnaire (8000 respondents). The major advantage of pooling the two surveys was a much better coverage of public transport trips. Both surveys include trip diaries for one day, together with information on background variables for the respondents. For both surveys a considerable effort was spent on accurate coding of the geographical origin and destination of trips. Some of the data were also recoded for use in the estimation of mode/destination models. A set of rules was designed and used to recode tour chains with more than one destination into tour/return trips from home to a single main destination and back. The rules were based on time spent on each destination, distance from home to each visited destination, total distance for the trip chain and total distance for any possible simplification of the trip chain. Complex chains with no distinct main destination/purpose and/or many destinations were excluded from the data set used for estimation. This left us with about 23 000 tours in the data set for estimation.

For each observation (round trip) a random sample of 249 alternative destinations (within 100 km from the residential zone) was drawn to allow for simultaneous estimation of mode and destination choice on a subset of alternatives. Zonal data for attraction and LoS-data were connected to both the chosen destination and each of the randomly drawn, alternative destinations.

The five mode/destination models are home and tour based, and they use segmentation from the car availability and license holding models as direct input. The models are estimated for the following five possible travel modes:

- Car driver
- Car passenger

- Public transport (rail, subway, tram, boat, or bus)
- Bicycle
- Walk

The bicycle mode has only a small market share nationwide in Norway, varying from 3 per cent for shopping and service trips to 8 per cent for the private visits. Information from the coded network for public transport indicates that a significant amount of the observations did not have public transport as an available mode for the actual trip (as much as 60 per cent for some of the purposes). We believe that this percentage is higher than revealed by LoS-data, indicating that the network for public transport may have been coded with somewhat insufficient accuracy.

The data show significant differences between men and women with respect to mode choice as well as trip length distribution. Women are more often car passengers than men, and they use public transport more frequently. They also seem to walk more than men do. On average, women travel shorter distances than men, especially for work trips and business trips.

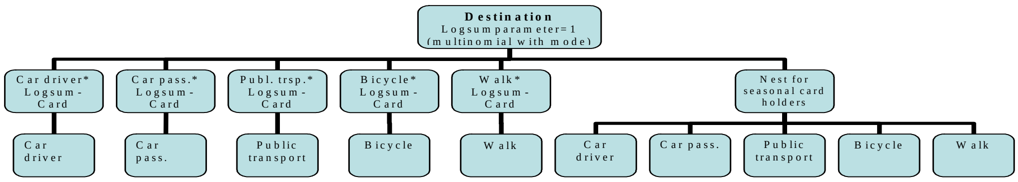
The models for mode and destination are logit models, describing how individuals choose between different alternatives (combinations of mode and destination) under the utility maximization assumption. In a multinomial logit model the utility of each alternative is composed of a deterministic and a random component, and the random components of all alternatives are assumed independent and Gumbel distributed. We want to study the choice of mode and destination simultaneously, and in this case the assumption of independent random components may be unrealistic. We can then introduce structured (or nested) logit models, allowing for different random terms at different levels in the three structure. In our work we have tested alternative tree structures for all travel purposes. For most purposes the variant with mode choice on the upper level had a somewhat better log-likelihood-value than the simultaneous multinomial model, but the value of the cost coefficient was significantly smaller and it had lower significance, which led to higher implicit values of travel time. Based on the fact that the values of time were already higher than expected, we chose to use the simultaneous multinomial model for three of the purposes (business, private visits, shopping/service). The model for “other private trips” is structured with destination above mode, while the model for work trips has a special structure, the acquisition of seasonal cards for public transport being estimated simultaneously with the mode and destination choice.

The model for “other private trips” was the most difficult to estimate with satisfactory results, probably because the travel purpose is a quite heterogeneous one, including leisure and recreational travel (often quite long trips), giving someone a ride (often very short trips, e.g. kids to and from the kindergarten), and several other minor travel purposes.

The model for work trips has a somewhat more complicated structure than the other models, involving simultaneous estimation of mode choice, destination choice, and seasonal card possession (zero cost for marginal trips) for public transport. The simple (and usual) way to handle seasonal cards in modelling is to use the price of the card divided by the number of days travelled as the public transport fare applicable to seasonal card holders (same fare assumed for all possible destinations). This has several disadvantages, since for seasonal card holders the cost of public transport use is perceived as zero on any given day. According to the travel survey, a small portion of the seasonal card holders use their car to go to work on the day of registration. These travellers, who normally would go by public transport, rarely own discount cards for road tolls or ferries, making the perceived cost of choosing the car even higher. In our work trip model these aspects are captured, in a far more correct way than in traditional mode/destination models.

In the model for work trips seasonal card holding is placed in a separate nest with its own logsum parameter (figure 4). This implies that we get five additional utility functions, one for each mode. The other utility functions are multiplied by this same logsum parameter to bring them to the same level (a “dummy-nest” for each mode). This is necessary to get the correct generic parameters. The model is simultaneous multinomial except for the nest for seasonal cards, implying a logsum parametre equal to one, as stated in the figure.

Figure 4. The model structure for work trips.



TØI-report 766/2005

In the nest reflecting alternatives having a seasonal card, the cost of using car as driver or passenger is defined as the sum of the ordinary cost of using the mode, and the per day cost of the season card. The cost of using public transport is the per day cost of the season card only. *Inside* the nest for seasonal card holders, public transport will be perceived as free, thus having a high probability of being chosen. *Between* the nests, public transport will be less expensive with a seasonal card than without. At the same time other modes will be more expensive for those having a seasonal card. The cost of a seasonal card varies by origin and destination, and is most often dependent on the distance, but there are also unit fares for certain regions (inside some of the cities) and zonally differentiated fares (in the larger Oslo area).

This approach to modelling the choice of mode, destination and seasonal card acquisition has several advantages, such as:

- More correct specification of the choices, in a way that reduces the correlation between travel time and travel cost.
- The model for seasonal card holding can be used for the four other travel

purposes as long as the model for work trips is run first.

- It becomes possible to analyse effects of different pricing schemes of single tickets and seasonal cards. Changes in prices will have an impact on both seasonal card holding and mode choice.
- It becomes possible to differentiate the costs for car drivers and car passengers depending on whether they use these modes sporadically or regularly (through discounts for toll costs and ferries). This will also serve to reduce certain cumbersome correlations.

While space does not allow a complete description of mode/destination choice models estimated for each of the five travel purposes, but some “highlights” which are common to all the models are worth mentioning:

- The coefficients associated with the waiting time variable for public transport were hard to estimate, due, probably, to problems and inaccuracies in the network coding. Lots of different variants of the variable were tested, as “open” and “hidden” waiting time, waiting time per boarding, various non-linear transformations, the sum of waiting time and access time, etc. In the final models waiting time is entered as a square root function. The coefficients have the right sign and size, and are significantly different from zero. Compared to a linear specification, the square root formulation implies that the per unit waiting time effect will vary with the time spent waiting, and that the impact of very short or very long waiting times is reduced.
- A dummy introduced for observations with at least two destinations on the tour reveals an increased probability of car use on such trips.
- The five different categories of car availability are important variables in all five models. In general, full car availability implies increased probability of being a car driver. In households with more license holders than cars, women seem to lose the “battle to use the car”.
- Several dummies for age and type of region (large city etc) have been tested and to some extent included in the models. We have, however, tried to avoid having too many segments in the final models (a high number of segments increases the computer time dramatically), and kept only the most significant ones.
- In Norway, the available information about parking opportunities (accessibility and costs) at a zonal level is very limited. We have therefore calculated indices based on the density of jobs at the respective destinations. The indices enter the models as dummy variables. For most purposes (all except private visits) a high job density reduces the probability of car use. The size and sign of the coefficients of these dummy variables reveal, as seen in comparison with the cost coefficients, a reasonably sized costs of parking when the duration of the visits (parking time) for the respective travel purposes is taken into account.
- Model calibration revealed certain problems in reproducing the observed trip length distribution. At first, the models appeared to underestimate the frequency of short and long trips, while overestimating the frequency of medium length trips. Distance dummies had to be introduced to overcome this problem.

One interesting application of this type of model is the revealed value of travel time savings, as represented by the ratio between the time coefficients and the cost coefficient. In our model, the implicit values of travel time savings appears to be on the high side in all models, except for work trips. One reason may be that the travel times used in estimation are calculated in a network model without capacity problems (giving too short travel times). There is also correlation between time and cost, which might give a tendency for one of the coefficients to reflect the total effect of distance. The time values are particularly high for the category "other private trips", probably because this travel purpose is very heterogeneous, ranging from long recreation trips to short private errands, such as picking up the children in the nearby kindergarten. However, the majority of these trips take place in weekends, when the value of time is considerable lower than for the rest of the week (see "in addition, weekend" in table 1). This may reflect a "looser time budget" in the weekends. This tendency is also found for private visits. The implicit values of travel time from the models are shown in table 1.

In all five models some of the values of time are higher for women than for men. It is not obvious how this should be interpreted. It might reflect the fact, confirmed by time use surveys, that women are generally busier than men, enjoying less "pure" leisure time or that they have a stronger perception of travel time discomfort.. Alternatively, it may have to do with the fact that women generally spend less time doing work for pay or for profit, have lower wages, and generally enjoy a larger, job independent household income. They may thus have a lower marginal utility of money, resulting in a smaller cost coefficient.

Table 1. Implicit values of travel time from the mode/destination models. NOK/hour (1 Euro = appr. 8 NOK)

	Work	Business	Private visits	Shopping/serv	Other
Car driver (CD):					
Men		135	63	66	198
Women		242	89	87	253
In addition, large city		58			
In addition, weekend			-19		-84(m), -77(f)
Men, less than 50 year	55				
Men, 50 year +	73				
Women, less than 50 year	84				
Women, 50 year +	106				
In addition, destination Oslo rush	32				
Car passenger (CP):					
All car passengers			61	55	
Men	92	64			198
Women	154	147			253
In addition, large city		58			
In addition, weekend			-19		-84(m), -77(f)
Public transport (PT):					
Invehicle time			33	28	92
Invehicle time, men	38	56			
Invehicle time, women	48	125			
In addition, weekend					-32
Access/egress time	43	168	61	57	133
Waiting time, when waiting 5 minutes	60	369	70	84	124
Waiting time, when waiting 30 minutes	24	150	29	34	51
Waiting time, when waiting 60 minutes	17	106	20	24	36
Transfer (NOK per transfer)		39		16	

TØI-report 766/2005

School trips (trips to and from school) are not covered by the mode/destination choice models, since the school bus routes have been not coded in the network model. These trips are dealt with through simple gravity models distributing children and adolescents to schools and universities. No mode choice model has been estimated for these trips.

Trip frequency models

For each travel purpose (including school trips), the *frequency models* produce the expected total number of sojourns (not trips or tours) as a function of age, gender and other background variables, together with logsums taken from the mode/-destination choice models. The data input comes from the National Travel Survey

2001, with supplementary information on the number of visits made during the registration day for each of the travel purposes.

The concept is based on a multinomial logit models that take care of the distribution of sojourns between *different travel purposes*, combined with Hurdle-Poisson models simultaneously predicting the *total number of visits*. Poisson models are a natural choice when the dependent variable is a non-negative integer (0,1,2,3,4,5,….), while the Hurdle-Poisson formulation is a convenient choice when the zero mass point has a probability which deviates from the standard Poisson distribution. The Hurdle-Poisson model provides the expected total number of visits for the six travel purposes, while the distribution between purposes is given by the logit model.

We first tried to estimate a joint model for the entire population, but this resulted in a huge amount of segments with dummy variables for different age groups. Instead we estimated independent models for five different age groups (13-24, 25-34, 35-54, 55-66 and 67+). The share of each travel purpose differs significantly between the age groups (table 2), which is why different models were needed.

Table 2. Distribution of visits between travel purposes for each age group/model. Week day, all of Norway.

Purpose/age	**13-24**	**25-34**	**35-54**	**55-66**	**67+**	**Total**
Work	12 %	26 %	31 %	28 %	3 %	24 %
Business	3 %	9 %	11 %	9 %	1 %	8 %
Shopping/service	19 %	24 %	24 %	29 %	53 %	26 %
Private visits	15 %	10 %	7 %	11 %	14 %	10 %
Other private trips	26 %	29 %	27 %	23 %	30 %	27 %
School	24 %	3 %	1 %	0 %	0 %	6 %
Total	100 %	100 %	100 %	100 %	100 %	100 %

TØI-report 766/2005

The model for each age group is specified by one utility function for each travel purpose, formulated with alternative specific constants and variables describing sex, age, family type, type of residence, and a logsum from the corresponding mode/destination model (which implies that the travel frequency is influenced by the accessibility described by the mode/destination models, i.e. level of service, attractiveness of destinations, and car availability). In addition we had to introduce some regional dummy variables.

Procedure for calculation of “intermediate” trips

The estimated models for mode/destination are based on tours starting and ending in the respondent’s own home with only one main destination, while the models for travel frequency gives the total number of sojourns made. If the results from these models were used directly, the model system will produce too many trips (because in practice some of the visits are “intermediate” trips, and not a round trip). To correct for this we have used a *procedure that takes care of round trips with intermediate destinations*. The procedure uses the information from the mode/destination models and the frequency models in such a way that the system produces the correct number of outbound trips, intermediate trips and return trips, based on matrices of "transfer probabilities". These express the probability that a visit of a certain purpose is followed by either a return trip or a new visit, and, in the latter case, for a specified purpose.

The matrices of "transfer probabilities" are calculated based on data from the National Travel Survey 2001. For all trips that are not return trips, the purpose of the given trip and the next trip is registered. All such trip pairs are counted and put into a table, and the probabilities are calculated based on this.

Some simplifying assumptions are used:

- All tours have one or two destinations.
- All legs in a trip are by the same mode.
- Only the attractiveness relative to the home zone matters for the choice of destinations 1 and 2.

The procedure used maintain full consistency in the sense that the calculated expected number of visits for each purpose is included in the OD-matrices produced by the model, and for each zone the numbers of starting and terminating trips are the same.

Preliminary results also indicate that the procedure will reproduce the total number of trips in the travel survey quite accurately.

Implementation

A run of the model results in tour matrices for simple tour/return trips, together with matrices for outbound trips, intermediate trips and return trips for tours with more than one destination. For each travel purpose there are separate matrices for each of the modes. The matrices can then be used to compare the transport situation in different scenarios for a given year, to forecast traffic growth for different modes in a future year etc. If we want to know the composition of the traffic on road links and on different public modes (the demand model calculates public transport as only *one* mode), the matrices have to be read into a network model. The trips will then be assigned on links and routes based on the algorithms of the network model.

The model system is programmed in C++ and implemented as a stand-alone program that can use different network models as platform as long as they can produce the same type of LoS-data. The model system has been implemented for five non-overlapping regions. Users can also define smaller regions that may be more suitable for their purpose and thus save on execution time and on the amount of input and output.