Arbeidsdokument av xx. juli 1999 PT/xxxx/99
O2430 UIPRO
Forskn.ass. Berit Grue

Uncertainty in the Norwegian national transportation model

Paper prepared for Trafikdage på Aalborg Universitet 31.08.99

Uncertainty in the
Norwegian national transportation model

Research assistant Berit Grue
Institute of Transport Economics, Oslo

1. Introduction

The Norwegian national transportation model (NTM-4) consists of several submodels adding up to a relatively large model system. The main model parts can be characterised by types of output data predicted:

- Possession of driver's licenses in the population
- Car ownership
- Number of trips performed on an average day, divided into short (<100 km) and long trips

A four-stage sequential model calculates the number of trips between all combinations of zones. The model system is based on 438 zones – one for each municipality - which gives an amount of geographical travel relations in the magnitude of >190,000. With the number of trips specified by trip length, different transport modes and travel purposes, the total number of values for trip variables is in the range of 7-8 million.

It is obvious that all these values cannot be predicted with perfect precision, and that examining the uncertainty and errors has to be related to certain types and levels of aggregation.

The following classification is made for the different sources to uncertainty in the NTM-4:

- *Basic data.*
  The model system is estimated on the Norwegian travel survey. Systematically biased data or errors from reporting will affect the model results.
- *Input data.*
  The coding of all details in the transportation system will indubitably be infected with some undiscovered errors and uncertainty. These types of data are used both when estimating and executing the transportation model. An inevitable source of uncertainty, especially connected to short trips, is the zone structure itself with municipalities defined as the minimum geographical units.
- *Model specification* is about the specification of relations between variables, the selection of independent variables and their appearance in the model, and the presumptions made about the distribution of the stochastic term. Limitations to model specification can appear from unavailability of data of supposed relevancy. A more basic question is if the use of a logit model system estimated on individual survey data provides the preferred results.

- *Statistical uncertainty* appears in all models based on econometric estimation methods. This type of uncertainty is indicated by the variance-covariance matrix and by T-values of estimated parameters.

The statistical uncertainty is not necessarily the most important source of error in a model, but is the only type of uncertainty that can be dealt with by analytical methods. This document describes a simulation project with the NTM-4 where effects on model results from variances in estimated parameters are examined.

2. Method

Every part of the model is estimated with the maximum likelihood method, which imply that the parameter estimates have asymptotic multinormal distribution. With this assumption, random normal variables can be generated and transformed to a multinormal distribution, using Cholesky decomposition of the symmetric and positive definite variance-covariance matrix of the estimated parameters.

We have drawn 100 sets of values each for two main groups of parameters in the NTM-4:

- Parameters for income or variables where income is included in the construction of the variable.
- Parameters for variables describing quality of transport (travel time, cost etc.)

The parameters related to income are important for long term prognoses, as changes in travel demand normally depend on income development. The second group of parameters is connected to more instant effects from changes in transport standard and travel costs.

Monte Carlo simulations is used for handling the model executions with the changing sets of parameters. A simulation with 100 model runs involves an enormous amount of calculations, but the output from one simulation is restricted to 100 sets of aggregated values for a selection of important result variables.

The sets of result values are analysed in different ways. A single simulation gives the effects from the different parameters on the levels of model estimators. To examine the elasticities from an exogenous variable, a change in the related constant value is applied before the simulation is run again with the same set of parameters. We have therefore run the simulation with the 100 sets of income parameters both with original constant factors and with a change in income development. In the same way, the simulation with the transport standard parameters is run four extra times with alternative value for one exogenous variable – air cost, train frequency, travel time in train or variable car cost.

3. The model system and the parameters

There are in total 14 parameters related to income located in the different parts of the model system. The parameters are connected to different segments, for instance trip length, travel purpose or transport mode, and are likely to show significant effects on related result variables calculated in the same part of model. It is possible though, that a parameter located in only one module can have significant effects on all result variables, if it affects values used as input to other parts of the model system. Table 1 shows a description of the 14 income parameters with estimates and localisation in the model system. Figure 1 is a brief picture of the NTM-4 and the interchange of data between the different model parts.

Table 1: The 14 parameters related to income. Coefficient estimates, and localisation in the NTM-4.

	Original coefficient estimate (present model)				Travel segment (long trips)		
	Coefficient value	**T-val.**	**Std. deviation**	**Model**	**Travel purpose**	**Transport mode**	**Parameter description**
1	-0.010009	-8.3	0.001204819	Driver's license possession			household income, no driver's licenses in hh
2	0.004603	8.5	0.000541176	Driver's license possession			HH income, both in hh have driver's licenses
3	0.2611	10	0.02611	Car ownership			log(remaining income)
4	0.6585	9.1	0.072362637	Frequency, long trips	Work		Ln(personal income)
5	0.0008633	3.6	0.000239806	Frequency, long trips	Private		Household income
6	0.001008	6.4	0.0001575	Frequency, long trips	Leisure		Household income
7	0.0004882	3	0.000162733	Frequency, long trips	Visit		Household income
8	0.007817	5.3	0.001474906	Mode, long trips	Leisure	Car	Income
9	-0.1873	-12.3	0.015227642	Mode, long trips	Business	Car,air,train,bus	Cost/personal income
10	-0.2074	-8.5	0.0244	Mode, long trips	Visit	Car	Cost/ hh income
11	-0.03337	-5	0.006674	Mode, long trips	Visit	Air	Cost/ hh income
12	-0.04631	-4	0.0115775	Mode, long trips	Visit	Train,boat,bus	Cost/ hh income
13	-0.01449	-4.2	0.00345	Mode, long trips	Leisure	Train,boat,bus	Cost/ ln(hh income)
14	-1.324	-6.3	0.21015873	Mode, long trips	Leisure	Air	[1]

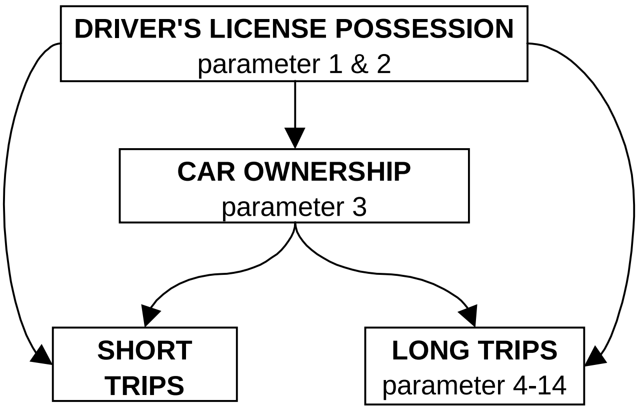

Figure 1: Interaction between the parts of the NTM-4 system and the localisation of income parameters.

[1] $\sqrt[2/3]{Cost} / hhincome$

The parameters related to transport standard are far more numerous than the income parameters in the model system. Important indicators for quality of transport, are among others travel time, cost, access and egress time and frequency. For this project we have selected 46 transport standard parameters (12 located in the submodel for short trips and 34 in the long distance model) mainly connected to transport time and cost for different travel segments and the logsums for each travel purpose.

4. Results

Table 2 shows that the effects from variances in income parameters are of less importance to the short trips compared to long trips, and that the car trips are less affected than the results for public transport trips.

Table 2: Effects from variances in income parameters. Boundaries for 95% confidence intervals compared to results with original parameter estimates.

Result variable	Relative standard deviation	95% confidence interval compared to original estimate (%)	
		Lower limit	Upper limit
Number of cars	15,345	- 29,6	+ 20,3
Short trips			
Total	0,447	- 0,7	+ 0,8
Car	2,518	- 4,6	+ 4,7
Public transport	6,866	- 11,0	+ 11,5
Long trips			
Total	4,609	- 7,9	+ 8,6
Car	6,859	-13,2	+ 12,8
Bus	18,312	- 22,5	+ 43,4
Boat	16,429	- 22,3	+ 33,0
Train	18,116	- 24,2	+ 38,2
Air	12,472	- 20,2	+ 26,5

Elasticities on model results from the estimated parameters can be examined by applying ln(result variables) as dependent variables in linear regression analyses, and calculate the equations coefficients(B) * estimated parameter values.

The results are shown in table 3, and we can see that the parameter no.3 has large impact. This parameter operates in the car ownership model, and represents the log value of the remains of income after car ownership costs are covered for. This parameter gives the largest elasticity for all result variables, also with respect to short trips, but the coefficients for short trips are, as expected, relatively small.

Table 3: Elasticities from income parameters with respect to transportation model results.

Variable	Parameter number													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
No. of cars	.074	.100	1.664											
Short trips	.020	.023	.027											
Short, car	.037	.050	.251											
Short, PT	-.037	-.055	-.697											
Long trips	.018	-.008	.283	.265	.038	.147	.057	.023	-.034	-.003	.004	-.003	.001	.004
Long, car	.063	.024	.622	.214	.041	.177	.051	.030	.031	-.019	.000	.007	.014	-.002
Long, bus	-.282	-.115	-1.589	.085	.021	-.010	.096	-.021	-.069	.081	.023	-.049	-.076	.089
Long, boat	-.207	-.083	-1.466	-.015	.100	.033	.062	-.071	-.049	.058	.013	-.048	-.136	.128
Long, train	-.243	-.123	-1.488	.671	.009	.001	.069	-.020	-.120	.047	.010	-.046	-.087	.077
Long, air	-.133	-.069	-.990	.554	.007	.024	.027	-.030	-.489	.022	-.047	.007	.027	-.125

The two sets of results from Monte Carlo simulations, of which one is run with increased income growth factor, are used to examine the income elasticities and the corresponding confidence intervals - in other words a valuation of the uncertainty due to income changes.

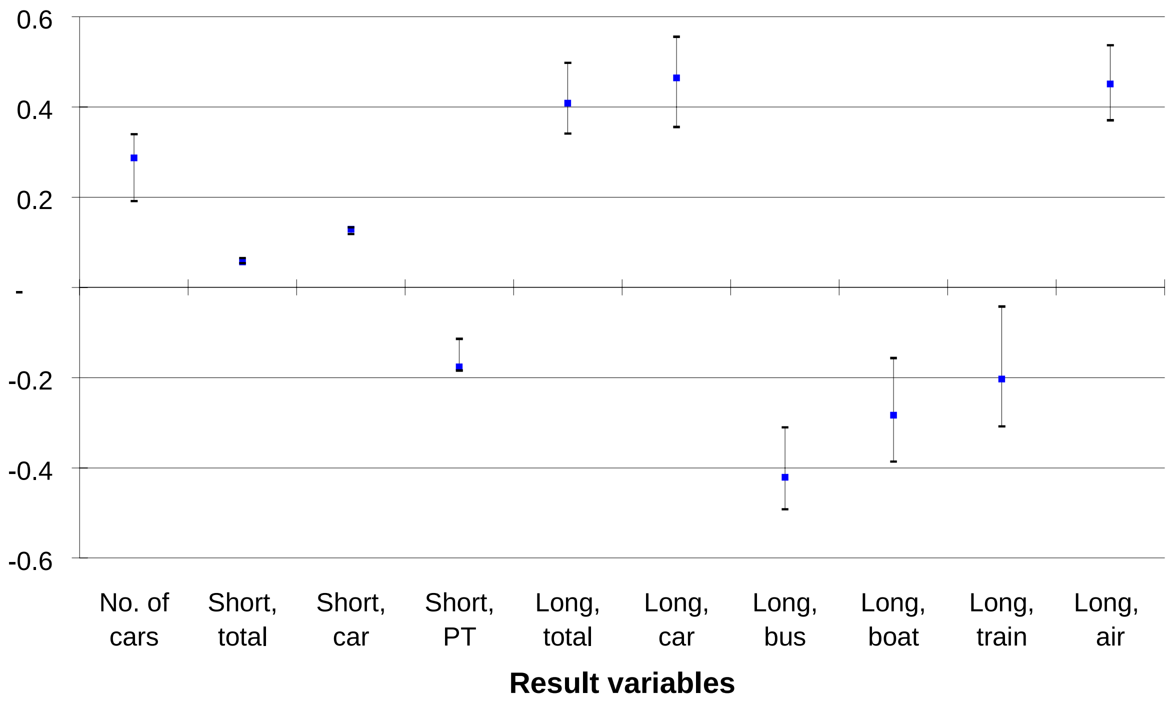

Figure 2: Income elasticities for result variable from the NTM-4. Elasticities from original model estimates are plotted with 95% confidence intervals.

Figure 2 shows the income elasticities from the set of drawn parameters, with 95% confidence intervals. The results show significantly negative income elasticities relating to long bus, boat and train trips, and that income growth leads to high transition from these travel modes to car and air transport.

After simulation with the income parameters, the long distance model was re-estimated (Rekdal & Hamre 1998), resulting in positive income elasticities for all transport modes and higher values relating to car and air transport than in figure 2. Even though the levels of elasticities from the new version seem more evident, we assume that the simulation results assimilate the re-estimated model concerning size of confidence intervals.

The re-estimated model version was used for the simulations with transport standard parameters. This part of the project made no demand for running the driver's license and car ownership models, as the selected parameters only appear in the models for trip generation. The short and long trips are modelled in separate segments with no interchange of data (as figure 1 shows), and have no common transport standard parameters. Effects caused by one parameter, is therefore connected only to results from the submodel in which it is located.

Linear regression analyses on the simulation results reveal that the logsum parameters have the most significant effects on the number of short trips. For long trips, we get significant effects from almost all transport standard parameters in the long distance model.

From table 4, we see that transport standard parameters in general cause less variation in model results than the income parameters. Again short trips come out less affected than the long distance trips, and there are smaller variations connected to short distance car trips than to public transport. Concerning long distance, the car trips are less influenced than other transportation modes by uncertainty in the parameters tested with this simulation.

Table 4: Effects from variances in transport standard parameters. Boundaries for 95% confidence intervals compared to results with original parameter estimates.

Result variable	Relative standard deviation	95% confidence interval compared to original estimate (%)	
		Lower limit	Upper limit
Number of cars			
Short trips			
Total	0,090	- 0,2	+ 0,2
Car	1,635	- 3,1	+ 3,0
Public transport	9,791	- 12,2	+ 24,8
Long trips			
Total	1,250	- 2,6	+ 2,3
Car	1,433	-2,8	+ 3,3
Bus	12,129	- 19,1	+ 26,2
Boat	6,654	- 10,1	+ 16,1
Train	5,684	- 12,9	+ 12,8
Air	7,121	- 11,3	+ 16,5

In the same way as with the income parameters, extra model simulations are run to lay the basis for calculations of elasticities from exogenous factors. For these simulations, we have used four exogenous variables concerning transport quality – air transport cost, travel time in train, train frequency and variable car cost. In figure 3, the direct elasticities for trips on the related modes of transport are plotted with 95% confidence intervals.

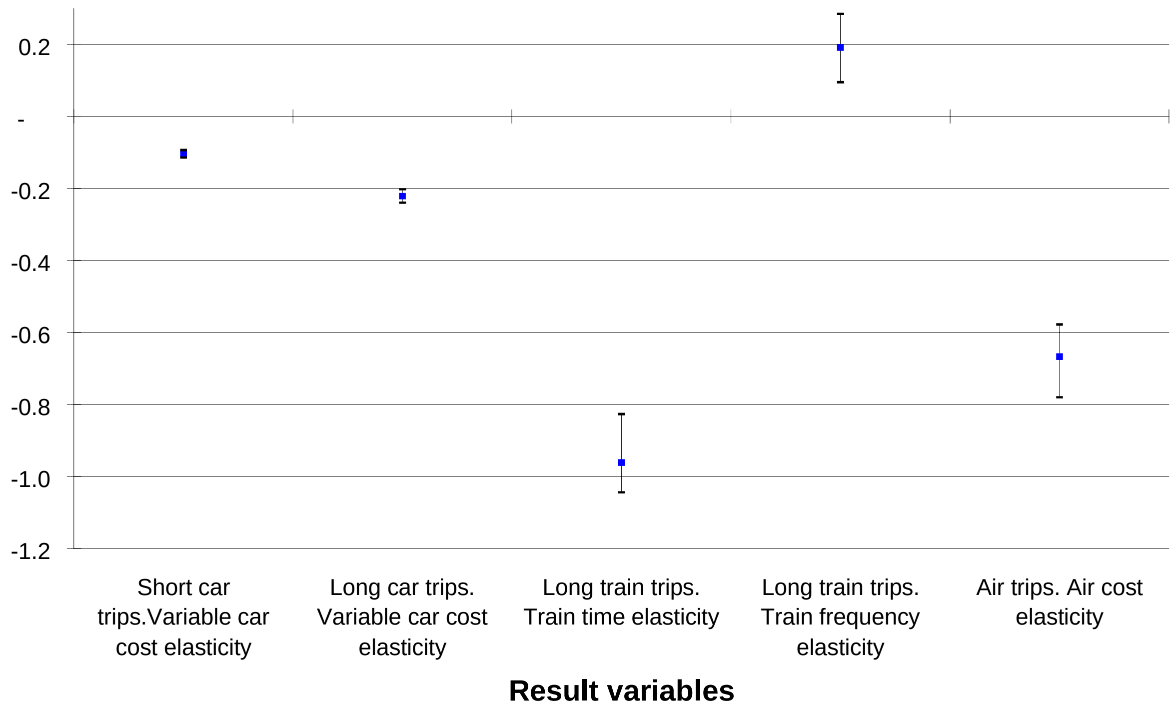

Figure 3: Elasticities from transport quality factors with respect to number of trips on related modes of transport. Elasticities from original model estimates are plotted with 95% confidence intervals.

5. Conclusion

With a model system built up by connections between different submodels, it would be impossible to picture the general uncertainty in result variables by studying specific coefficient's standard deviations or T-values separately. Drawing parameter values with distribution according to estimation method, and then simulating by running the model system with a large number of drawn parameter sets, makes it possible to achieve a description of statistical uncertainty in the model system as a total.

The income parameter in the car ownership model is the most important income related parameter in the NTM-4. This parameter value affects a large number of result variables because the car ownership influences on the travel mode choice. In general, uncertainty in income parameters has small effects on short trips, which appears from the low values and small confidence intervals for income elasticities. For results related to long trips, the income elasticities have higher estimated values than short trips, and the confidence intervals are in the range of +/-15%.

The confidence intervals for elasticities related to transport standard variables (figure 3) are located within an area that would be difficult to achieve from time series data, except with very long time series.

A re-estimation of the long distance model with some changes in model specifications was accomplished after finishing the simulations with the drawn income parameters. The re-estimation resulted in differences in levels of estimated income elasticities, but the confidence intervals calculated with the previous version are assumed to be representative for the new model results as well.

After re-estimation, the income elasticities for long trips seem more evident regarding all transport modes, and the relatively small confidence intervals indicate that statistical uncertainty does not induce major problems to long term prognoses. The model specification is probably a more important source of uncertainty and errors in the model results, indicated by the changes in levels of estimated values after re-estimation.

References:

Grue, B. 1999
Usikkerhet i den nasjonale persontransportmodellen.
Oslo, Transportøkonomisk institutt. TØI notat 1125/1999

Judge, G et al. 1985
The Theory and Practice of Econometrics. John Wiley & sons, New York.
ISBN 0-471-84572-8

Rekdal, J. 1998
Logsummer i logitmodeller estimert på transformerte data.
Oslo, Transportøkonomisk institutt. TØI arbeidsdokument PT/1236/98

Rekdal, J og Hamre, T. 1998
Den nasjonale persontransportmodellen. Versjon 4c.
Oslo, Transportøkonomisk institutt. TØI arbeidsdokument PT/1282/98