A model describing the develoment of road freight transport

Ole Kveiborg
National Environmental Research Institute, Frederiksborgvej 399, P.O. Box 358, DK-4000 Roskilde.
Phone +45 46301835, Fax +45 46301212, E-mail: olk@dmu.dk

Paper presented at the conference
"Trafikdage på Aalborg Universitet '98"

Abstract

In an ongoing cooperation between Statistics Denmark and NERI we are developing a satellite model to a Danish macroeconomic model ADAM. That is a model linked to the macroecenomic development in ADAM, deriving the freight transport demand on Danish roads and energy consumption from this. A first operational version of the model is now ready. In the paper an overall view of the model is given, with simple calculation examples. A general discussion of the possible developments of the model into a dynamical model is given with focus on a specific area of the model.

1. Introduction

As environmental aspects in present policy proposals become more and more important, the need for methods to help decision makers to make better decisions is evident. ADAM (Annual Danish Aggregated Model) is an econometric short to medium term model originally developed to predict the economic development in 19 different sectors in the Danish economy when different policies are implemented. In the 80's and 90's different models were developed, using the results from ADAM to predict the energy-related environmental impacts. Some of these models have been implemented in close connection with the ADAM model complex and have been named 'Satellite models'[1]. Other national models, developed in the same period work independently from the actual ADAM model complex[2].

A very large part of the energy-related impacts stem from the transport sector. Transport is included in ADAM via two sectors (sea transport and other transport services). These sectors deliver transport services to the other sectors through the economic input-output structure that ADAM is based upon. The existing environmental models are either connected to ADAM through the simple description of the transport sector in ADAM, or through the overall development in GDP in ADAM. These simple connections are then used to make predictions of the transport demand (in kilometers), and hence the derived energy consumption.

The transport demand varies between different sectors in the economy. This means that the overall development in GDP is not a very sophisticated way of combining the economy with the transport sector (as it is done in the Reference model), nor is the actual transport demand, measured as energy consumption, directly related to the economic development in the different sectors (as it is done in the Satellite models, see Andersen and Trier, 1995). Neither of these present models are satisfying ways to predict the future transport demand. The proposed model will combine these two models into one model, using the developed economic link from the Satellite models and the more detailed transport elements from the Reference model.

---

[1] These models are described in Andersen and Trier (1995).
[2] One of these models is the socalled Reference model, which has been developed by the Department of Transport in coorporation with Cowiconsult (Cowiconsult, 1990).

The model will be developed in several steps. Till this moment (mid July 1998) the first step has been completed, and the work on the succeeding steps has begun. This has led to a preliminary operational model describing the road freight transport in Denmark. Similar work have been carried out in the EU-financed project REDEFINE (see Cardebring et al., 1998), and also by McKinnon and Woodburn (1996) for GB. Neither of these have had as i prime target to develop an actual model.

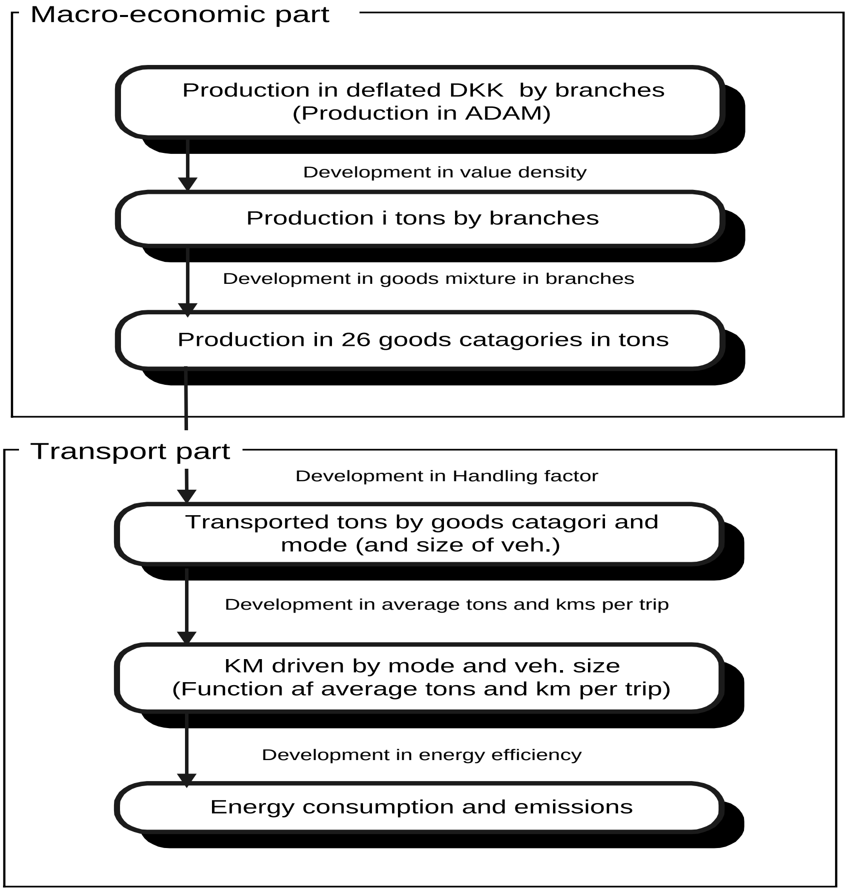

Figure 1 *The general structure of the model. The boxes represent different successive calculation steps of the model. The upper three boxes represent the "macro-economic" part of the model, where production is measured in DKK and tons respectively. Between the boxes the possibilities for time developments in the different coefficients are shown. The present version of the model is only linked through fixed coefficients.*

In section 1 the general outline of the model will be presented, focusing on the present fixed coefficient version of the model. In section 2 a simple calculation example is discussed. In the third section possibilities for transforming the model into a more dynamic model is described. The final section gives a brief discussion on the different limitations of the model and areas, in which the model in the future to incorporate.

2. The general model structure

The model consists of two parts, a macro-economic input-output submodel and a transport submodel. The interface between these two parts is the actual improvement on the Satellite models and the Reference model. In *Figure 1* the general outline of the model is presented.

The upper half represents the economic part of the model and the lower half is the actual transport demand and resulting energy consumption.

In the following the two parts of the model are described separately.

2.1 Macroeconomic submodel

The input to the model is the production in 19 different sectors or branches in the Danish economy, measured in deflated DKK (1980-DKK) in each of the forecast years. The size of productions in each sector are calculated in ADAM based on a predefined forecast. Imports and exports are calculated separately, as it is assumed that international trade influences the national transport in another way than the national production.

The production in monetary values is linked to production measured in tons through a direct coefficient:

$$m_i^P = a_i * fX_i \qquad (1)$$

Where fX_i is the production in branch *i* in deflated 1980-DKK
m_i^P is the production in branch *i* in tons, *P* indicates production
a_i is the weight/DKK-coefficient (the value density).

The problem in using this direct connection is, that some of the production (in DKK) are services and not physically produced goods. The implication is that the coefficient a_i is zero in some branches. In sectors with both "production" of services and physical goods, the intra-branch distribution of these two types of production is assumed fixed through time. To give a possibly more accurate value of $a_{i,}$ the value of services could be deducted before the calculation of produced tons. This has however not been done in the present version of the model.

In the next step the production in the branches (in tons) is distributed on 25 different goods categories (*Crops, Potatoes, Crude Oil* etc.) and a supplementary mixed cargo category. This includes both a distribution within each branch, and an aggregation on the different goods categories:

$$m_v^P = \sum_i b_{iv} * m_i^P \qquad (2)$$

where v is an index of goods category, and
b_{iv} is the share of the production in branch *i*, that is of goods category *v*, measured in tons.

This is the macro-economic sub-model in broad terms.

2.2 The transport submodel

The central element of the transport submodel is the connection to the economic submodel. In the present version of the model constructed this is a simple fixed between production and tons lifted in the different goods categories. The tons lifted are further distributed on transport mode, with a handling factor for each mode. The term mode is in this model interpreted widely. It includes trucks in several categories (vans, trucks below 6 tons gross vehicle weight, from 6 to 16 tons, 16 to 32 tons and above 32 tons) as well as a differentiation between *own transport* and *haulage contractors* is performed. Each of these is treated as a separate mode:

$$m_{jv}^{T} = c_{jv} * m_{v}^{P} \qquad (3)$$

where j is the index for transport mode, and
T indicates transported tons

The factor c_{jv} expresses 'how many times a specific good is transported with a specific mode'. It can in that respect be interpreted as an indicator on how many intermediate links a good must pass through before it has reached its final place of consumption. It is however not a measure of the modal split.

The tons lifted is then transformed into traffic performance (vehicle kilometers), with and without load. This calculation consists of two parts: the *average load* per trip and the *average distance* per trip:

$$vkm_{jv} = \frac{\overline{s_{jv}} * m_{jv}^{T}}{\overline{m_{jv}}} \qquad (4)$$

where $\overline{m_{jv}}$ is the average tons per trip per mode j and goods category v,
$\overline{s_{jv}}$ is the average trip length per trip per mode j and goods category v, and
vkm_{jv} is the vehicle kilometers driven per mode j and goods category v.

The energy consumption and the emissions of CO_2, SO_2 and NO_X are then simply calculated by multiplying the vehicle kilometers with the corresponding energy- and emission coefficients. This is done for each mode with and without load, but not separately for each good category. A separate calculation of the load factor within each category could improve the accuracy of this calculation. This has not been done in the present version of the model.

It is evident from the preceding, that the model is merely an advanced calculation model in its present version. In sector 4 a discussion of how this can be altered is presented.

3. An example illustrating the structure of the model

In this section the type of calculations or forecasts, that can be made with the present version of the model is illustrated. The forecast is based on an Agenda 21 economic forecast made by the Danish Department of Finance in 1995. The resulting production values in 1995 and in 2005 in the different sectors of ADAM are shown in *Table 1*.

In general there has been a 23.6% increase in production measured in DKK and a 22.4% increase measured in tons. This is due to the fact that production in some sectors cannot be measured in weight terms. There is however a huge variation between the different sectors. The contribution to the general increase from the sectors varies in respect to the actual size of the production in the sectors. The *transport industry* sector has a large increase (44.5%), but the importance of this sector is limited because the production is relatively small, measured in weights. In the proceeding calculations, the service sectors (indicated by shaded areas in *Table 1*) do not have any direct influence on the actual amount of transport demanded. This may seem a bit strange as two of these sector are transport sectors. The explanation is that the transport sectors in themselves do not demand any transport (or have any physical goods that need to be transported). Instead the transport sectors delivers transport to the other sectors.

The values in the transport sectors are thus derived from the activities in the other sectors, which generally speaking, is the way ADAM works[3]. As explained in the second section of the paper the development in the model is from the production in 25 different goods categories to transport by different modes. The development in production in the 25 categories is illustrated in *Figure 2*.

Sector	Production 1995 (bill. Dkr.)	Production 2005 (bill. Dkr.)	Production 1995 (1000 T)	Increase prod.	Increase imports
Agriculture	49769	55622	20996.22	11.8%	34.6%
Crude oil	20969	30492	21905,84	45,4%	24,4%
Fuel oil	15799	15368	9842,01	-2,70%	-77,4%
Electricity	18808	17262	1852,48	-8,20%	-4,7%
Suppl. construct.	15313	19365	44981,39	26,5%	48,6%
Transp. industry	8927	12896	548,99	44,5%	67,8%
Chemical ind.	31876	37386	4525,18	17,3%	56,9%
Construction	48468	66176	0,00	36,5%	66,3%
Trade	78088	100766	0,41	29,0%	55,8%
Sea Transport	38830	71758	0,00	84,8%	92,2%
Other Transport	46456	60178	0,00	29,5%	12,3%
Public Services	133001	148540	0,00	11,7%	13,5%
Sum (all sectors)	843902	1042725	129994	23,6%	51,6%

Table 1 *The production values in selected sectors of ADAM in 1995 and 2005 measured in DKK and in tons. Due to the fixed coefficients the developments in production are the same in DKK and tons, except sectors with no physical goods produced. Imports have not been separated out in production values, but the increase in imports are illustrated in the last column.*

Shaded areas indicate sectors producing services and no physical goods.

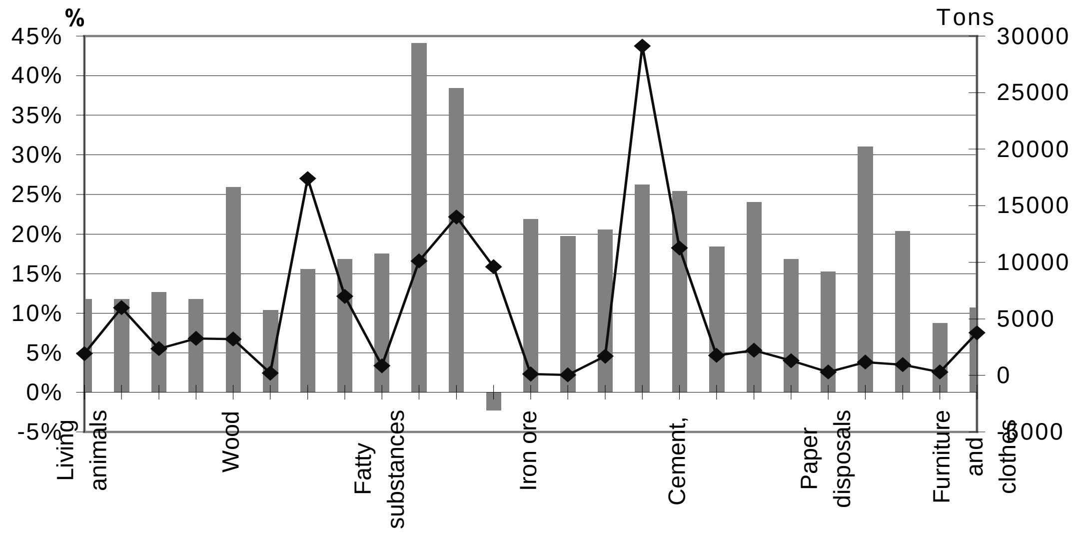

Figure 2 *The development in production in the 25 goods categories from 1995 to 2005 in percent (columns). The curve indicate the actual level in tons in 1995.*

[3] In general ADAM is constructed upon an input-output structure with fixed coefficients (α *),* so that an increase in e.g. the agriculture branch by one DKK results in an increase in the transport sectors by α DKK.

In *Table 2* the result from the proposed model is compared with the two similar models in Denmark. From this two things can be deducted: the new model calculates lower levels of energy consumptions, and the level of development is in line with the existing Satellite models (which is due to the similar linkage to the economic development).

	New model			**The Reference model**			**The Satellite model**
Year	Small veh.	Trucks	Total	Small veh.	Trucks	Total	Total*
1995	17.9	12.9	30.8	30.8	20.4	51.2	61.7
2000	20,0	14,0	34,0	31,1	22,2	53,3	69,9
2005	22,1	15,4	37,5	33,1	23,0	56,1	75,3
Growth	23.5%	19.4%	21.8%	7.5%	12.7%	9.6%	22.0%

Table 2 *A comparison of energy consumption measured in PJ between the new satellite model, the Reference model and the existing Satellite model for three selected years. The bottom line is the overall increase in energy consumption from 1995 to 2005. The figures for the Reference model and the Satellite model are from Andersen and Trier (1995).*

* The calculations are not distinguished between the transport modes.

Two obvious reasons for the lower level of predicted energy consumption are:

- The new model has not yet included all road freight transport.
  - A large part of the international transport (Danish vehicles crossing Danish borders) have not been included because detailed information about this are sparse.
- Another reason is the applied energy coefficients.
  - The coefficients in the new model are based on international experiments reported in the Copert database on energy- and emissions coefficients, and these are generally lower than the coefficients used in the Reference model and the Satellite model. For a strict comparison corrections for this should be made.

With the coming inclusions of these missing elements of the model, it is expected that the proposed model will give accurate and plausible predictions of the energy consumption from the national transport.

4. Time dependent development of selected coefficients

The model, as it has been described in section 2, is a very rigid model with fixed relations between each element. The implications of the fixed coefficients are that changes in the resulting transport demand is determined in the macro-economic development in ADAM. However, it is evident that this rigidity has to be loosened to incorporate changes in production compositions, technological changes, and changes in legislation on vehicle size etc. This section illustrates how this will be included in the model. The focus will be on the transportation side of the model, as this is the only area where data are available at this point. Note that the contents in this section are preliminary as elaborate investigations into the data have not yet been performed.

4.1 Theoretical discussion

The values of the variables in the model change due to changes in the production input from ADAM, and through the relations explained above. A dynamic development thus arise through time dependencies in the different coefficients in the model. In general this can be described as:

$$\alpha = \alpha(t, \mathrm{X}) \qquad (5)$$

Where α represents any coefficient in the model (the handling factor),
t denotes time, and
$\boldsymbol{X}$ is a vector of other explainable variables, either variables from ADAM or strictly exogenous variables.

Often this relation is specified as a linear (or log-linear) combination of the explainable variables at any point in time:

$$\alpha_t = \gamma_1 + \gamma_2 \mathrm{X}_{1,t} + \gamma_2 \mathrm{X}_{2,t} + \ldots + \gamma_n \mathrm{X}_{n,t} \qquad (6)$$

where $\gamma_{0,}\ \gamma_1, \gamma_2$ to γ_n are parameters that have to be estimated.

α_t can be dependent on lagged variables of the explainable variables. It is furthermore possible to use other more general mathematical forms. When time lags are added to the above formulation it is formulated as follows:

$$\alpha_t = \beta_0 + \beta_1 \alpha_{t-1} + \Gamma_0 \mathbf{x_t} + \Gamma_1 \mathbf{x_{t-1}} + \varepsilon_t \qquad (7)$$

where Γ_0 and Γ_1 are vectors of parameters according to the vector of explainable variables $\boldsymbol{x}$, lagged 0 and 1 period respectively, and
ε_t is the error terms, distributed identical and independent, $\varepsilon_t \sim \mathrm{IID}(0,\sigma^2)$[4].

This type of equation is called an *autoregressive distributed lag* with one lag in both exogenous and endogenous variables (or simply ADL(1,1)). More lags on both sets of variables could have been included, but the very limited amount of data makes this impossible. Without the error terms and with proper sizes of the β-parameters, the long run solution to this ADL is:

$$\alpha^* = \frac{\beta_0}{1-\beta_1} + \frac{\gamma_{1,0} + \gamma_{1,1}}{1-\beta_1} x_1^* + \frac{\gamma_{2,0} + \gamma_{2,1}}{1-\beta_1} x_2^* \qquad (8)$$

where $\gamma_{i,j}$ is the parameter associated with $x_{i,}$ lagged j=0 or j=1 period respectively, an asterisk indicate the steady state, long run equilibrium value of the variable (only two explainable variables x_i have been included here).

The interesting issue of the present model however, is to know the development over time towards the long run equilibrium. This can be introduced with a rearrangement of (7)

$$\alpha_t - \alpha_{t-1} = \beta_0 + \left[(\beta_1 - 1)\left(\alpha_{t-1} - \lambda_1 x_{1,t-1} - \lambda_2 x_{2,t-1}\right)\right] + \gamma_{1,0}\left(x_{1,t} - x_{1,t-1}\right) + \gamma_{2,0}\left(x_{2,t} - x_{2,t-1}\right) + \varepsilon_t \qquad (9)$$

where $\lambda_1 = \dfrac{\gamma_{1,0} + \gamma_{1,1}}{1-\beta_1}$, and $\lambda_2 = \dfrac{\gamma_{2,0} + \gamma_{2,1}}{1-\beta_1}$

[4] The errorterm ε_t could also be described as a time dependant proces (moving average or autoregressive). This has not been done here.

(9) is also called an *Error-correction model (ECM)*. The term in square brackets in (9) is the *error-correction term,* measuring the extent to which the values of the preceding period are differing from the long run relationship between α_t and $\boldsymbol{x}_t$. (β_1-1) is the proportion of the disequilibrium that is reflected in the movement of α_t in one period. The succeeding terms are impacts of the short run adjustments in the exogenous variables.

The ECM are common in ADAM and in the Satellite models because of their simplicity, and because the estimation is easy due to the fact that the model is a linear model reparametrised in a nonlinear way. The estimation of the ECM is simply an OLS regression.

For a further practical description of the ECM, see Statistic Denmark (1995), and Andersen and Trier, (1995), and for a more theoretical discussion see Davidson and MacKinnon (1993) and Harvey (1990).

The ECM estimation structure can however only be used when an adequate number of time periods are used. This is the case for the transport specific elements of the model (the lower three boxes in *Figure 1*). It will be more doubtful that the number of observations are adequate within some of the macro-economic elements in the model though. Instead some kind of cross-section estimation will be applied, looking at differences between different groups rather than looking at actual time developments. This is not further elaborated at this point.

4.2 Example of an estimation of a coefficient in the model

The handling factor c_{jv} will act as an example of how the ECM formulation is implemented in the model. First it has to be determined which variables act as explainable variables for the development in c_{jv}.

In McKinnon and Woodburn (1996) the development in the handling factor in GB between 1983 and 1991 is analyzed. They give the following reasons for changes (increases) in the factor:

1. Increase in the number of separate links in the supply chain.
   - more retailers and subcontracters
2. Increased weight loss during the production process
   - at each stage in the production chain some part of the products are discarded, and therefore no longer apparent in the tons lifted, whereas product weight is measured at the point of consumption.
3. Changes in the amount of packaging.

These explanations are supplemented in Cardebring et al (1998) with:

4. Changes in mergers and acquisitions in retail and wholesale sectors.

The listed explanations can however not be directly used to estimate the developments in the handling factors, as no data are available on a macro level.

The locational pattern of producers and consumers could also be an interesting explaining factor. This is however a factor influencing the long run, whereas ADAM is a model giving predictions on the short to medium term.

Another interesting variable is the price development of transport. There are of course a lot of different prices in the different branches and services in the transport sector. Ideally a different price for each different transport should be produced and used in the estimation. This is however not possible. As the model will be implemented as a satellite model to

ADAM, only prices within ADAM can be used, hereby securing a consistent estimation throughout the model complex. This means that instead of the various prices from the different transport services and branches, prices in the two transport branches in ADAM, prices on import goods, and prices in other relevant sectors of ADAM are the only prices that are possible to use. These price variables are semi-exogenous variables, as they are determined in the overall model complex. Together with the semi-exogenous prices, strictly exogenous variables should be used. These exogenous variables can lead to different developments between e.g. different modes or goods categories. In the example given here, variables indicating changes in government regulations on vehicle sizes would be such strict exogenous variables that may have an impact on the development of the coefficient. The problem with this type of variable is that it is not directly a quantitative variable. Some kind of transformation into a quantitative (continuous) measure would therefore be necessary before the actual estimation.

The ***X***-vector of explainable variables (see equation (5)) is then

X*={price Other Transport, prices of selected import goods, number of vehicles of mode j, number of vehicles of modes* $i \neq j$*}.*

where *Price Other Transport* is the sector price from ADAM in the branch *Other transport*.

The estimation using one step OLS will according to Davidson and MacKinnon (1993) lead to consistent results if $|\beta_1|<1$ (the stability condition). The estimation will also indicate which variables that are significant in explaining the development of the coefficient. Non-significant variables will be left out of the final model. The preliminary results are shown in *Table 3* below. The estimation is for the goods category *fertilizer* on trucks between 16 and 32 tons gross vehicle weight by haulage contractors. Similar estimations are performed for the other vehicle categories and for the own transport. These are however not reported here.

A stepwise selection method has been applied to decide on included and excluded variables. The last column in the table reports the number of times a specific variable has been entered in the estimation based on the stepwise selection.

Explainable variable	Estimated	t-values	final estimates[a]	Times entered in model[a]
Constant	78,62	0,789	20,62	25
α, lag=1	-0,67	-2,621	-0,65	11
Price *other transport*	-4,30	-1,060		0
Price *other transport*, lag=1	-4,92	-1,187	-2,64	2
Price *sea transport*	1,20	1,643		2
Price *sea transport*, lag=1	2,49	2,856	1,99	5
No. Of vans	0,00	0,585		3
No. Of vans, lag=1	-0,00	-0,461		3
No. Of small trucks	-0,01	-0,737		2
No. Of small trucks, lag=1	0,00	0,481		3
No. Of heavy trucks	0,00	0,527	-0,00	4
No. Of heavy trucks, lag=1	-0,00	-0,603		4

Table 3 *Preliminary estimation results of the one step OLS estimation of the handling factor(c_{ij})for the good fertilizers by haulage contractors on trucks between 16 and 32 tons gross vehicle weight. t-values are from estimations including all explainable variables.*

[a] Entries based on a stepwise selection, and only for variables entered

One surprising thing, that can be seen form the table, is that the price of *other transport* (which it is a conglomerate of all types of transport in the air and on the ground) has not been significant in any of the estimations on a 10% significance level. One explanation of this is that the market is slow in reacting to a price change. The influence of the price can thus be seen after one time period.

Another important implication from *Table 3* is that the number of significant parameters in each estimation is very low. In most cases only one or two[5]. A number of different explanations for this can be found. First of all the exogenous explainable variables are very aggregate and general, and do not have the variability needed for the different estimations. Secondly the variables have not been analyzed for multi-collinarities and other important statistical properties. Thirdly the general ECM estimation structure may be wrong, and the mathematical structure of the estimation equations may be erroneous, and finally the chosen exogenous variables are not the ones explaining the development or alternatively: there are no development in the specific coefficients. Contradicting the low number of explainable variables entering each estimation, are the small number of observations. The fewer the number of observations the larger is the possibility for different explainable variables entering the estimation.

One other important thing to note is that the signs of the parameters all seem to have the right sign. This is unfortunately not the case in all the estimations, implying that further investigations have to made in these estimations.

5. Discussion

In this paper a first version of a proposed satellite model to the Danish macro-economic model ADAM has been presented. The model is a combination of two existing models each being superior to the other in some areas. This first version is so far a very simple and rigid model. The aim of the model is to give an overall indication of the development of the transport demand as a result of different macro-economic initiatives. The supplementary possibilities specifically directed at the transport sectors are thus mainly included for the possibility of taking in relevant changes in the transport sectors.

All the calculations are made very mechanically due to the fixed coefficient structure in the model. This means that the present version of the model cannot be used for analysis of the impacts of policy proposals directly influencing the transport sector. The only kind of analysis that can be performed is impact analysis of different economic policy proposals influencing the economic sectors defined in the macro-economic model ADAM. With the refinements of the coefficient structure it will however be possible also to make simple impact analysis of initiatives specifically on the transport sector. As an example legislations on vehicle sizes have been mentioned, but a number of other possibilities could also be incorporated. One thing that will not be directly included is the possibility of making analysis of policies that influence the diesel or gasoline price. This is due to the pre-defined price on transport in ADAM. Introducing a new price structure or element into the model could lead to inconsistencies in the model complex. The only way of introducing the fuel price is to change the price on transport in ADAM. This change will however influence all the different subsectors defined in the proposed model in the same way, making it impossible to implement a tax aimed at separate transport sectors.

In the present version of the model only road transport modes (trucks and vans) are included. This is so due to very poor data on the other transport modes (rail, sea transport, and transport services such as coach, bus, and taxi driving). It is also because especially rail

[5] This is also due to the selection procedure chosen. Using for example a forward selection result in a significantly higher number of significant parameters. This method does not allow exclusion of variables, that have been entered at an earlier stage in the estimation.

and sea transport only contribute with a very small share of the overall goods transport. These modes will be included in future versions of the model to give a complete description of the transport sectors in the economy.

Literature

Andersen, F.M. and Trier, P. (1995) *Environmental satellite models for ADAM,* CO_2*,* SO_2 *and* NO_X *emissions.* NERI Technical Report No.148. National Environmental Research Institute, Denmark. 200 pp.

Cowiconsult (1990) *Reference model for the Danish Transport Sector 1988-2030*. Cowiconsult, Lyngby, Denmark (In Danish)

Davidson, R. and MacKinnon, J.G. (1993) *Estimation and Inference in Econometrics.* Oxford University Press, New York

Energistyrelsen, (1996) *Energy statistics 1995.* Copenhagen, Denmark (In Danish).

Harvey, A. (1990) *The Econometric Analysis of Time Series.* Phillip Allan, London.

Statistics Denmark (1995) *ADAM - A model of the Danish Economy, March 1995.* Statistics Denmark, Copenhagen (In Danish)

McKinnon, A.C. and Woodburn, A. (1996) A logistical restructuring and road freight traffic growth. - An empirical assessment. *Transportation* Vol.23 pp. 141-161

Cardebring, P.W. et al. (1998) *Linking economic activity and road freight traffic performance. Findings of the EU-sponsored project REDEFINE*. Proceedings of Economics and Institutions of Transport 1998, KFB and Högskolan Dalarna, Sweden.