DISCRETE MODE AND DISCRETE OR CONTINUOUS SHIPMENT SIZE CHOICE IN FREIGHT TRANSPORT IN SWEDEN

Gerard de Jong
ITS Leeds, Significance and NEA
Daniel Johnson
ITS Leeds

1. INTRODUCTION

Mode choice in freight transport is usually studied in isolation (or in combination with network assignment, as multi-modal assignment). However, mode and shipment size are closely linked decisions. Large shipment sizes usually coincide with higher market shares for non-road transport, whereas there is a high correlation between road transport and small shipment sizes. Previous models that combined both choices in a single model include McFadden et al. (1985), Abdelwahab and Sargious (1992), Abdelwahab (1998) and Holguín-Veras (2002). These authors estimated joint discrete-continuous models, where mode choice is the discrete and shipment size the continuous choice. Other authors (e.g. Chiang et al. 1981, de Jong, 2007) have studied this joint decision-making problem as two simultaneous discrete choices, classifying shipment size into discrete categories.

This paper investigates whether it makes a difference for the main policy outcomes of the model (notably the transport time and cost elasticities per mode) whether the disaggregate model is specified as:

- An independent discrete mode choice model (which is the most common formulation)
- A joint model with discrete mode and discrete shipment size choice
- A joint model with discrete mode and continuous shipment size choice.

All three models will be estimated on the same data using standard regression and discrete choice model estimation software, to yield the models as they would be estimated by researchers in practice, but starting from the three different specifications listed above. The question then is whether the three models will provide the same elasticities

The Swedish Commodity Flow Survey (CFS) is a unique data source in Europe. The CFS 2001 details about almost 1 million individual shipments to or from a company in Sweden, with information on origin, destination, modes used, weight and value of the shipment, sector of the sending firm, commodity type, access to rail tracks and quays, etc.. Whilst the US Commodity Flow Survey has been analysed several times (e.g. Sorratini, 2001; Vanek and Morlok, 1998, 2000), its Swedish counterpart has barely been used for model estimation so far. Using this Swedish CFS mode and shipment size choice at the individual shipment level can be explained from characteristics of the shipper, the shipment and transport time and cost on the networks.

Earlier work (De Jong, 2007) estimated both mode and shipment size as discrete choices, but clearly shipment size is a continuous variable. This paper further develops this earlier work to simultaneously estimate shipment size (as a continuous variable) and mode choice (as a discrete variable) using full information maximum likelihood estimation methods for sample selection problems applied to the Swedish 2001 Commodity Flow Survey.

Both assuming independence between mode and shipment size choice and discretising the continuous information on shipment size may be interpreted as forms of specification error.

The current paper presents estimation results for all three specifications distinguished above. Mode-specific transport time and costs elasticities of the mode shares and values of time from the joint discrete/continuous model are compared against those from the independent discrete mode choice model and the joint discrete choice model. This shows whether these differences in specification lead to differences in the elasticity values - the model outputs that are typically used to evaluate transport policies.

In section 2, the CFS 2001 data are described, as well as the data transformations and data linkages that were carried out to establish the estimation data set. Section 3 gives the estimation results for a pure mode choice model independent of shipment size. In section 4 are the estimation results for models where both the mode and the shipment size are treated as discrete choices. The outcomes for the joint discrete-continuous model are in section 5. Section 6 contains elasticities for the three different types of models. Finally in section 7 are the conclusions from this research.

2. DATA USED

The CFS 2001 data file (SIKA, 2003) we are using has 922,913 records. Each record is a shipment to or from a company in Sweden, with information on origin, destination, modes used, weight and value of the shipment, sector of the sending firm, commodity type, access to rail tracks and quays, etc. From this we selected a file of around 749,000 outgoing shipments of Swedish production and wholesale firms (domestic transport and export, no import) for which we have complete information on all the endogenous and exogenous variables.

The (endogenous) mode choice variable in our models refers to the mode used in Sweden. We use four modes in all models:

- Road transport
- Rail transport
- Water transport
- Air transport.

The CFS contains information on mode chains (e.g. truck-train-truck). If a chain contained an air transport mode, we classified it as air transport; if it contains water transport modes (but not air transport), we labelled it as water transport, and if the chain comprised a train transport mode (but not air of water transport), it was regarded as rail transport. Road transport consists of mode chains with road transport modes only. Models that explain the choice of mode chain, together with discrete shipment size, estimated on the latest Swedish CFS (2004/2005) can be found in Windisch (2009).

For the discrete mode and discrete shipment size model, we classified the continuous weight variable from the CFS into five categories:

- Up to 3,500 kg
- 3,501-15,000 kg
- 15,001-30,000 kg
- 30,001-100,000 kg
- Above 100,000 kg.

The choice options in this model are the 17 combinations of mode and shipment size listed in Table 1. For air transport there were not enough observations in the three highest weight categories for inclusion in the model.

Table 1 also gives the number of observations by mode and shipment size alternative (the number of times each option is chosen in the CFS). Road transport is clearly dominant (in terms of the number of shipments). A road transport with a shipment size in the fourth and fifth category will be a convoy of several vehicles.

Table 1. Choice alternatives used in the model with discrete mode and shipment size; and observed CFS frequencies

Choice alternative	Mode	Shipment size	Number of observations in estimation data set
Road1	Road transport	Up to 3,500 kg	649,683
Road2		3,501-15,000 kg	42,042
Road3		15,001-30,000 kg	16,737
Road4		30,001-100,000 kg	13,720
Road5		Above 100,000 kg	1,233
Rail1	Rail transport	Up to 3,500 kg	4,453
Rail2		3,501-15,000 kg	995
Rail3		15,001-30,000 kg	1,433
Rail4		30,001-100,000 kg	1,771
Rail5		Above 100,000 kg	1,318
Water1	Water transport	Up to 3,500 kg	5,486
Water2		3,501-15,000 kg	1,489
Water3		15,001-30,000 kg	1,541
Water4		30,001-100,000 kg	458
Water5		Above 100,000 kg	644
Air1	Air transport	Up to 3,500 kg	6,011
Air2		3,501-15,000 kg	388
Total			748,952

In the CFS, the origins of the shipments are coded by municipality. The domestic destinations are also given in terms of municipalities, for the foreign destination there is information in terms of the zones in the STAN national freight transport model. Within Sweden, the STAN model uses municipalities as well. The municipality codes and foreign STAN codes were used to append network information to the CFS records. From the STAN networks we took distance between origin and destination and time between origin and destination by mode. This information was used to calculate transport costs, including:

- Distance-based link costs (e.g. fuel)

- Time-based link costs (e.g. labour)
- Initial loading and final unloading costs
- Access and egress costs to/from the main mode
- Transfer costs.

The network information and the costs function information was assembled in the course of a project to develop a logistics module for the Swedish and Norwegian national freight transport models (for the Samgods group in Sweden and the NTP group in Norway). We used the information on this that was incorporated in the versions 0.1-0.3 of the logistics model (RAND Europe and SITMA, 2005, 2006)[1] but simplified the costs functions to fit the mode-shipment size combinations that we are using in this paper.

3. AN INDEPENDENT DISCRETE MODE CHOICE MODEL

Several specifications were tried. The estimation results of the best-performing multinomial logit (MNL) model are in Table 2. The MNL model contains many very significant coefficients (owing to the very large sample size), which have the expected sign. The estimation implies that shipments of big firms are more likely to be shipped by rail transport. Shipments with a low value density (the value of the goods divided by the shipment weight) are more likely to be transported by rail and water transport, whereas the highest value density goods have higher road and especially higher air transport probabilities. The transport cost of a mode has a negative influence on the probability of being chosen. Air transport is much more time-sensitive than the other modes, followed by road transport and then rail transport. For water transport (the slowest of the modes), transport time was not significant. The constants indicate that after correcting for the above influences, road is the most attractive mode (relative to air, which is the reference here), and rail and water transport are less attractive.

The above results are all quite plausible, and many investigations might stop here. However, in the above the endogenous nature of shipment size was ignored, whereas in reality this is a choice variable as well. We now therefore move on to models that also explain shipment size.

4. A MODEL FOR DISCRETE MODE AND DISCRETE SHIPMENT SIZE CHOICE

This model has been published before (De Jong, 2007, De Jong and Ben-Akiva 2007). It is based on a logistics costs specification.

The total annual logistics costs G of commodity k transported between firm m in production zone r and firm n in consumption zone s of shipment size q using mode (or rather transport/logistic chain) l:

$$G_{rskmnql} = O_{kq} + T_{rskql} + D_k + Y_{rskl} + I_{kq} + K_{kq} + Z_{rskq} \qquad (1)$$

Where:
G: total annual logistics costs
O: order costs
T: transport, consolidation and distribution costs

Table 2. Estimation results for multinomial logit model for independent mode choice

Variable	Relevant alternatives	Coefficient	t-ratio
Road constant	Road	3.925	167.0
Rail constant	Rail	-1.818	-33.7
Water constant	Water	-0.997	-22.8
Company is in biggest size class (sector-dependent)	Rail	0.417	13.8
Value density (value/weight) in bottom 20%	Rail	1.440	62.9
Value density (value/weight) in quantiles 20%-40%	Rail	0.229	7.4
Value density (value/weight) in bottom 20%	Water	0.193	7.6
Value density (value/weight) in top 10%	Road	0.563	15.7
Value density (value/weight) in top 10%	Air	1.089	18.1
Value density (value/weight) in top 5%	Air	2.323	50.1
Commodity type is metal products	Rail	-0.724	-17.5
Commodity type is chemical products	Rail	0.288	6.2
Transport cost in SEK/shipment	All	-0.403E-05	-52.9
Transport time in hours (*10)	Road	-0.642E-02	-39.1
Transport time in hours (*10)	Rail	-0.134E-03	-6.6
Transport time in hours (*10)	Air	-0.0155	53.6
Number of observations: 749,062			
Final log likelihood value: 131876.5			
Pseudo rho-squared w.r.t. zero: 0.873			
Pseudo rho-squared w.r.t. constants: 0.068			

D: cost of deterioration and damage during transit
Y: capital costs of goods during transit
I: inventory costs (storage costs)
K: capital costs of inventory
Z: stockout costs

Equation (1) can be further worked out (see RAND Europe et al, 2004; RAND Europe and SITMA, 2005):

$$G_{rskmnql} = o_{k.}(Q_k/q_k) + T_{rskql} + D_k + (d.t_{rsl}.v_k.Q_k)/365 + (w_k + (d.v_k)).(q_k/2) + Zr_{skq} \qquad (2)$$

Where:
o : the constant unit cost per order
Q: the annual demand (tonnes per year)
q : the average shipment size.
d: the discount rate (per year)
v: the value of the goods that are transported (per tonne).
t: the average transport time (in days).
w: the storage costs per unit per year.

The logistics cost function used in the mode and shipment size model is an approximation to (1) and (2). It includes link-based transport costs, transfer costs, but –for the air transport options- also the value of the shipment times the transport time, to represent the capital cost on the inventory in transit (Y). For the other modes, the coefficient for this last term was not significant in estimation. The value of the goods (per weight unit: value density) is included for the two smallest shipment sizes for each mode to represent that for high value goods, small shipment sizes are more likely (to keep the inventories down: influence of I and K). The estimated model does not include specific terms for order costs, deterioration of the goods and for the safety stock (because information on deterioration and annual demand is missing). However, the latter two components will be proportional to value of the goods and shipment time, so the terms for value density and value times transport time will also be picking up some of these influences.

Several model specifications were tried for the logistics cost specification, including the multinomial logit (MNL), and several variants of the nested logit model and the mixed logit model. The nested and mixed logit models on this data set did not lead to satisfactory results (so far). The estimation results for the multinomial logit model with a logistics cost specification are in Table 3.

Many estimated coefficients imply similar behaviour as in the transport cost specification, and are not discussed here again. Products with a high value-density are more likely to be shipped in small quantities (shipment sizes up to 15 tonne), to keep the inventory costs down. Transport cost and the variable for inventory costs during air transport (transport time times value of the shipment) have the right (negative) sign, and are highly significant. The numerical outcomes imply that for a shipment worth 1 mln SEK, these costs are 10 SEK per hour. This implies a 9.4% interest rate on an annual basis, which is clearly higher than the interest rates at capital markets in Scandinavia. As in the transport costs specification, the time-dependent link-based transport costs (labour and vehicle costs) have already been taken into account in the transport costs. Estimation of separate transport time times value coefficients for road, rail and water transport did not lead to significant coefficients. This variable, representing capital costs on the inventory in transit, is only relevant for air transport.

Table 3. Estimation results for a multinomial logit model for discrete mode and discrete shipment size choice

Variable	Relevant alternatives	Coefficient	t-ratio
Road constant	Road	5.652	420.6
Rail constant	Rail	-0.788	-22.4
Air constant	Air	1.686	69.3
Access to industrial rail track at origin	Rail	2.562	108.7
Access to quay at origin	Water	1.514	40.1
Company is in biggest size class (sector-dependent)	Rail	0.592	17.9
Value density in SEK/kg (truncated at 1,000,000)	All modes: smallest 2 shipment sizes	0.0404	121.6
Commodity type is minerals, building material	Road2	-1.142	-53.9
Minerals, building material	Road3	0.050	1.8
Minerals, building material	Road4	5.147	169.5
Minerals, building material	Road5	15.12	133.9
Petroleum products	Rail4, Rail5	7.250	76.1
Metal products	Rail1, Rail2	-1.514	-20.2
Metal products	Rail3	1.520	19.9
Metal products	Rail4	6.229	75.6
Metal products	Rail5	17.96	158.7
Chemical products	Rail1	-0.616	-7.2
Chemical products	Rail2	-2.058	-8.0
Chemical products	Rail3	2.178	20.6
Chemical products	Rail4	7.486	89.8
Chemical products	Rail5	17.96	148.7
Chemical products	Water1	1.238	33.6
Chemical products	Water2	0.257	3.4
Chemical products	Water3	2.107	20.3
Chemical products	Water4	4.750	32.7
Chemical products	Water5	13.86	78.7
Ores and metal waste	Rail2-5	5.525	76.0
Ores and metal waste	Water2-5	2.447	13.0
Machinery and equipment	Rail1	1.196	47.4
Machinery and equipment	Rail2	-2.116	-13.6
Machinery and equipment	Rail3	1.542	11.7
Machinery and equipment	Rail4	5.043	24.5
Machinery and equipment	Rail5	15.40	87.2
Machinery and equipment	Water1	0.502	16.8
Machinery and equipment	Water2	-0.687	-10.4
Machinery and equipment	Water3	2.208	29.3
Machinery and equipment	Water4	1.684	3.4
Machinery and equipment	Water5	12.59	53.8
Transport cost in SEK/shipment	All	-0.000128	-312.8
Transport time (in hours) times value of goods (in mln SEK)	Air	-0.00136	-109.1
Number of observations: 748,952			
Final log likelihood value :-689146.3			
Pseudo rho-squared w.r.t. zero: 0.675			
Pseudo rho-squared w.r.t. constants: -0.431			

5. A JOINT DISCRETE-CONTINUOUS MODE AND SHIPMENT SIZE CHOICE MODEL

In this model, we use the shipment size directly as recorded in the CFS, that is as a continuous variable. This leads to a joint model with continuous shipment size and discrete mode choice (with four modes). These are regarded here as simultaneous decisions. The model can in principle be estimated simultaneously, using Full Information Maximum Likelihood. However most software packages do not allow such simultaneous discrete-continuous estimation, especially with more than two discrete alternatives. Two-step methods for such problems have been developed as well, the most famous being Heckman's (1979) two-step estimator for the case with normal error terms. Here we use a two-step or indirect estimation first suggested and applied by Holguín-Veras (2002). The mode choice utility is given by:

$$U_l = \boldsymbol{\beta}_l \mathbf{W}_l + \phi y_l + \varepsilon_l \quad (3)$$

Where:
U_l is the utility derived from mode l
W_l: a vector of independent variables explaining mode choice
β_l: vector of parameters to be estimated
y_l: shipment size
ϕ: parameters to be estimated
ε_l.: error term

The continuous shipment size is given by:

$$Y_l = \alpha_l \mathbf{X}_l + \eta_l \quad (4)$$

Where:
$\mathbf{X}_l$: a vector of independent variables explaining shipment size
$\alpha_{ll:}$ parameters to be estimated
$\eta_{l\,:}$ error term

Following Holguín-Veras (2002), we now first estimate eq. (4) for shipment size on the basis of exogenous variables only, without distinguishing by mode l. Then we use the estimated shipment size y^e (instead of what was observed on shipment size in the CFS) as an instrumental variable in eq. (3) in each of the utility functions for mode choice. The estimation results of the regression for shipment size are in Table 4.

Table 4. Estimation results for continuous shipment size (in kg)

Variable	Estimated coefficient	t-ratio
Constant	35.05	53.4
Value density (value/weight) in quantiles 20%-40%	-31.74	-47.9
Value density (value/weight) in quantiles 40%-90%	-35.20	-62.9
Value density (value/weight) in top 10%	-36.33	-36.3
Value density (value/weight) in top 5%	-35.95	-33.8
Metal products	-10.40	-13.9
Chemical products	2.21	2.3
Company is in biggest size class (sector-dependent)	-0.54	1.0
Number of observations: 748824		
F statistic 615 (significant at 0.001%)		

In the mode choice model, the estimated shipment size (which is observation-specific, but not mode-specific) is used in two places: in the calculation of the costs per shipment (which depend on shipment size, through the number of vehicles needed) and in a variable for that measures the suitability of the mode for the shipment size at hand:

$$V_l = ABS(M_l - y^e) \quad (5)$$

Eq. (5) gives the absolute difference between the average observed shipment size for a certain mode l and the estimated shipment size. We hypothesise that at its average observed shipment size, the capacity of the mode and the shipment match very well (also assuming that most shipments are not consolidated with others). When the shipment deviates more from this average (either smaller or larger), the probability of choosing that mode for this shipment will decrease. We thus expect a negative estimated coefficient for this variable. The estimation results for the mode choice model are in Table 5.

Table 5. Estimation results for multinomial logit model including estimated shipment size at instrumental variable

Variable	Relevant alternatives	Coefficient	t-ratio
Road constant	Road	3.945	165.4
Rail constant	Rail	0.891	2.5
Water constant	Water	1.972	4.2
Company is in biggest size class (sector-dependent)	Rail	0.414	13.7
Value density (value/weight) in bottom 20%	Rail	0.381	2.4
Value density (value/weight) in quantiles 20%-40%	Rail	0.147	4.5
Value density (value/weight) in bottom 20%	Water	-0.606	-4.0
Value density (value/weight) in top 10%	Road	0.471	11.5
Value density (value/weight) in top 10%	Air	0.964	14.4
Value density (value/weight) in top 5%	Air	2.215	43.4
Commodity type is metal products	Rail	-0.654	-10.3
Commodity type is chemical products	Rail	0.269	5.7
Absolute difference between estimated and average observed shipment size V_l	All	-0.0178	-7.5
Transport cost in SEK/shipment	Road	-0.716E-06	-28.2
Transport cost in SEK/shipment	Rail	-0.745E-07	-10.4
Transport cost in SEK/shipment	Water	-0.542E-07	-19.3
Transport cost in SEK/shipment	Air	-0.687E-07	-13.5
Transport time in hours (*10)	Road	-0.527E-02	-35.0
Transport time in hours (*10)	Rail	-0.203E-03	-6.6
Transport time in hours (*10)	Air	-0.166	-58.3
Number of observations: 748,825			
Final log likelihood value: -133859.1			
Pseudo rho-squared w.r.t. zero: 0.871			
Pseudo rho-squared w.r.t. constants: 0.053			

When comparing the outcomes in Table 5 with the independent mode choice model (Table 2), we see that after including estimated shipment size, the coefficients for the rail constant, water constant, value densities for rail and water, transport time and transport costs have changed considerably. The estimated coefficient for V_l for the influence of the match between shipment size and vehicle size has the correct negative sign. Air transport is the most time-sensitive mode, followed by road transport and then rail. Mode comparisons between the various models will follow in Section 6.

6. COMPARING ELASTICITIES FROM THE THREE MODELS

We calculated time and costs elasticities from the three types of models, using the estimated results presented above. The outcomes are in Table 6 below.

Table 6. Own cost and time (per shipment) elasticities of mode choice (modal market shares in shipments) for three model specifications

	Independent mode choice	Discrete shipment size and mode	Continuous shipment size and discrete mode
Road cost	-0.004	-0.030	-0.003
Rail cost	-0.515	-0.126	-0.044
Water cost	-0.892	-0.073	-0.092
Air cost	-0.280	-0.001	-0.088
Road time	-0.040	-	-0.035
Rail time	-0.285	-	-0.434
Air time	-1.383	-0.871	-1.484

The elasticities in Table 6 only give the impact of substitution between modes. In the second model (with discrete shipment size choice), the changes in costs and time also have an impact on the choice of shipment size. If this is included in the elasticities, the road cost elasticity or demand for road transport for instance becomes around –0.5. Windisch (2009), using the Swedish CFS 2004-2005, found that there is much more substitution between shipment sizes than between transport chains (modes). She also found small own road cost elasticities, as we do in all three model specifications. All three specifications also show that air time has the highest elasticity. Nevertheless, three are important differences between the three specifications. For road cost, road time and air time, the independent mode choice model comes quite close to the more correct model that includes an instrument for continuous shipment size. However, for rail cost and road cost, the independent model leads to substantial overestimations of the elasticities. The direct reason for this is probably that we use the same (relatively high) cost coefficient for all modes in the independent mode choice model, but mode-specific coefficients in the model with continuous shipment size as instrument. This therefore is not directly the result of accounting for the endogenous influence of shipment size, but indirectly it is, because in the pure model choice model, separate cost coefficients by mode did not lead to significant estimates for all modes. In the simultaneous model it did.

The coefficients of the pure mode choice model are clearly different from those of the model with both choices included as discrete variables. The latter model also leads to quite different elasticities than the model that uses continuous shipment size.

7. CONCLUSIONS AND FUTURE WORK

The Swedish Commodity Flow survey 2001 was used to estimate different model specifications at the level of individual shipments. One model only explains mode choice, ignoring the endogenous nature of shipment size. The other two models simultaneously explain mode and shipment size, where shipment size can either be a discrete or a continuous variable.

All three model specifications lead to models with significant estimated coefficients that have the right sign. The elasticities for changes in cost and time by mode from the three specifications are however quite different. The only similarity in the elasticity outcomes is that road costs changes have a small impact on the market share of road in all three models (mode choice effect only) and that the direct elasticity of air time is the biggest of all direct elasticities. Starting from the pure mode choice model, accounting for endogenous shipment size leads to different direct elasticities for road and water transport. Also, a model where continuous shipment sizes have been converted into discrete categories leads to a model that has different behavioural responses than the model with continuous shipment size.

The model with discrete mode choice and continuous shipment size choice can be seen as the preferred model because it takes account of the endogenous nature of shipment size and uses the shipment size data as they come from the survey. However, standard software for simultaneous estimation is not generally available, and in the two-step estimation, the explanatory variables of shipment size need to be exogenous, so cannot include transport time and cost by mode. This model therefore cannot give an impact of time and cost of the choice of shipment size. From the model with discrete shipment size and mode, we conclude that shipment size is rather sensitive to changes in time and costs (so changing road transport will lead to bigger road shipments, not so much to substitution from road to rail and water-based transport). Therefore, in practical applications, a model with two discrete choices might be preferable, certainly when the classification into discrete shipment size categories would not contain just a few (as we did) but many categories.

Future work on this model may include re-estimation on the CFS 2004/2005 data, with new costs model input (also extending Windisch, 2009). It may also include simultaneous estimation of the discrete/continuous model, models with two discrete choices but more shipment size categories and models with other explanatory variables explaining some of the coefficients (observed taste variation) and with random (unobserved) taste variation in some of the parameters or with other more flexible substitution patterns.

NOTES

[1]The STAN model used here is not the latest version of the SAMGODS national freight transport model system for Sweden. The latest version includes a logistics module,version 2.0 explaining transport chain and shipment size choice at the micro-level (see for instance Ben-Akiva and de Jong, 2008 or Significance, 2008). This version was used to calculate the logistics cost input in Windisch (2009).

REFERENCES

Abdelwahab, W.M. (1998) Elasticities of mode choice probabilities and market elasticities of demand: evidence from a simultaneous mode choice/shipment-size freight transport model, **Transportation Research E**, 1998, 257-266.

Abdelwahab, W. M. and M. A. Sargious (1992) Modelling the Demand for Freight Transport, **Journal of Transport Economics and Policy** 26(1), 49-70.

Ben-Akiva, M.E. and G.C. de Jong (2008) The aggregate-disaggregate-aggregate (ADA) freight model system, in M.E. Ben-Akiva, H. Meersman and E. van de Voorde(Eds.), **Recent developments in transport modelling: lessons for the freight sector,** Emerald, Bingley, UK.

Chiang, Y., P.O. Roberts and M.E. Ben-Akiva (1981) Development of a policy sensitive model for forecasting freight demand, Final report. Center for Transportation Studies Report 81-1, MIT, Cambridge, Massachusetts.

Heckman, J.J. (1979) Sample selection as a specification error, **Econometrica**, 47, 153-161.

Holguín-Veras, J. (2002) Revealed Preference Analysis of the Commercial Vehicle Choice Process, **Journal of Transportation Engineering, American Society of Civil Engineers** 128(4), 336-346.

Jong, G.C. de (2007) A model of mode and shipment size choice on the Swedish commodity flow survey, Paper presented at UTSG 2007, Harrogate.

Jong, G.C. de and M.E. Ben-Akiva (2007) A micro-simulation model of shipment size and transport chain choice, Special issue on freight transport of **Transportation Research B**, 41, 950-965.

McFadden, D.L., C. Winston, and A. Boersch-Supan (1985) Joint estimation of freight transportation decisions under non-random sampling, in E.F. Daughety (Ed.) **Analytical Studies in Transport Economics**, Cambridge University Press, Cambridge.

RAND Europe, Solving and INRO (2004) The specification of logistics in the Norwegian and Swedish national freight model systems, Model scope, structure and implementation plan. Report for the NTP and Samgods group,TR-225-SIKA, RAND Europe, Leiden.

RAND Europe and SITMA (2005) The development of a logistics module in the Norwegian and Swedish national model systems, deliverable 4: final progress report on model development, PM-1968-SIKA, RAND Europe, Leiden.

RAND Europe and SITMA (2006) Documentation and clarification of deliverable 4 and the associated program delivery for a logistics module in the Norwegian and Swedish national model systems, deliverable 4a, PM-2055-SIKA, RAND Europe, Leiden.

Significance (2008) Method report - logistics model in the Swedish national freight model system (version 2), deliverable 6B for the Samgods group, Significance, The Hague.

SIKA (2003) Commodity Flow Survey 2001, Method report, SIKA Report 2003:4, SIKA, Stockholm.

Sorratini, J.A. (2001) Estimating state-wide truck trips using commodity flows and input-output coefficients, **Journal of Transportation and Statistics**, 3 (1).

Vanek, F. and E.K. Morlok (1998) Freight energy use disaggregated by commodity: comparisons and discussion, **Transportation Research Record**, 1641, 3-8.

Vanek, F. and E.K. Morlok (2000) Reducing US freight energy use through commodity based analysis. **Transportation Research D**, 5 (1), 11-29.

Windisch, E. (2009) A disaggregate freight transport model of transport chain and shipment size choice on the Swedish Commodity Flow Survey 2004/05, MSc Thesis, Delft University of Technology.