A micro-simulation model of shipment size and transport chain choice[1]

Gerard de Jong
Significance and Institute for Transport Studies, University of Leeds
Moshe Ben-Akiva
Significance and Massachusetts Institute of Technology

Abstract
Practically every international, national or regional freight transport model system in the world lacks explicit treatment of logistics choices (such as shipment size considerations or the use of distribution centres). This paper deals with the development of a new logistics model and its application within the national freight model systems of Norway and Sweden. This logistics model operates at the level of individual firm-to-firm (sender-to-receiver) relations and simulates the choice of shipment size and transport chain for all (several millions) these relations within the country, export and import. A logistics model with deterministic cost minimisation has been constructed for both Norway and Sweden. The full random utility logistics model has not yet been estimated on disaggregate data, but this is planned for both countries. For Sweden, more limited disaggregate models for the choice of mode and shipment size have been estimated.

1. Introduction

Many existing model systems for freight transport at the international, national or regional level use the conventional four step (production/attraction, distribution, modal split and assignment) approach, originally developed for passenger transport. Often, value-to-weight transformations and vehicle load factors are added as additional sub-steps (see de Jong et al., 2004). Usually, all steps are handled at the aggregate (zonal) level. Practically all these models (in section 2 we discuss exceptions to this rule) are lacking logistics elements, even though in recent decades logistic changes, such as the adoption of just-in-time delivery, have been (re-)shaping freight transport significantly. Logistics elements include the determination of shipment size and its influence on mode choice, or the use of consolidation and distribution centres. Here we define consolidation centres as locations where goods are transshipped (and possibly stored), with small loads getting in and larger loads getting out. Distribution centres are locations where goods are transhipped (and possibly stored), with large loads getting in and small loads getting out. Both consolidation and distribution centres exist not only in road transport, but can also be ports, airports or rail terminals.

In this paper we put forward a model that includes the determination of shipment size and the use of consolidation/distribution centres, within a behavioural framework, that can be estimated on disaggregate data and applied in micro-simulation. This model can be regarded as the logistics module within a larger freight forecasting and policy simulation system for a country, group of countries or large region within a country.

[1] The authors would like to thank Inge Vierth of SIKA/VTI and two anonymous referees for their very useful comments, but remain solely responsible for all the material presented in this paper.

The other parts –besides the logistics module- of the freight transport model system would be:

- Production-consumption (PC) matrices that give flows of goods by commodity type between two zones (municipalities for domestic zones, more aggregate zones abroad). Wholesalers can be included both at the production and the consumption end. In this case we call these matrices 'PWC matrices'. These matrices can for instance be generated by spatial input-output models (either multi-regional or regionalised national input-output models, see Marzano and Papola, 2004, Hunt and Abraham, 2005) or by spatial computable general equilibrium (SCGE) models.
- Assignment to the networks.

The PWC flows are input for the logistics module, after disaggregation of the zone-to-zone flows to the level of firm-to-firm (sender-to-receiver) flows. The outputs of the logistics model consist of origin-destination (OD) vehicle flows, which are used in aggregate network assignment. OD flows differ from PWC flows in that a PWC flow can consists of multiple legs, each with a different mode and with transshipments between the modes (e.g. a truck-ship-truck transport chain). At the transshipment points there can not only be changes of mode, but also consolidation of shipments together with other shipments, and de-consolidation. More information about the model as a whole can be found in Ben-Akiva and de Jong (2007). This paper focuses on the choices that are modelled in the logistics module:

- Frequency/shipment size (so inventory decisions are endogenous).
- The number of legs in the transport chain (direct transport, two legs, etc…)
- Use (and location) of consolidation and distribution centres for road and rail transport, but also including ports and airports.
- Mode (road, rail, sea, air) used for each leg, including choice of vehicle/vessel type and loading unit (unitised or not).

The latter three choices together are called 'transport chain choice'.

The focus of the paper is on the presentation of the general structure of the logistics model; there are no estimation results for the full transport chain choice on disaggregate data as of yet (this work is planned for 2007 and later). We restrict out attention to the pure transport of goods; vehicle movements for the delivery of services are not included.

In section 3 of this paper we present the inventory logistics part of the proposed logistics model, which focuses on the determination of shipment size. In section 4 the transport logistics part (choice of transport chain, including the number of legs and the modes for each leg) is discussed. Both inventory logistics and transport logistics are based on a minimisation of a full logistics cost function. This could be regarded as the behaviour of integrated shipper-carrier operations (there are no conflicts of interests between shipper and carrier in the model, as can occur when interaction is made explicit, as in Holguin-Veras et al., 2007). The treatment of empty vehicles is described in section 5 and in section 6 we discuss the required data for model estimation. Section 7 deals with an application of this framework to Norway and Sweden and describes the available data in these countries as well as our ideas on how to estimate the logistics model for each of the two countries, whereas section 8 deals with progress made so far (development of prototype or 'version 0' and a 'version 1') in Norway and Sweden. Related work concerns the estimation of a model for mode and shipment size on disaggregate data from the Swedish 2001 Commodity Flow Survey, starting from the same cost function. This is reported in section 9. This model is more limited in scope than the proposed logistics model, because it does not explain the use and location of consolidation and distribution centres. Finally, in section 10 conclusions are drawn and directions for future work are discussed.

2. Existing freight transport models that include logistics

Practically every international, national or regional freight transport model system in the world is lacking the above-mentioned logistics elements. Exceptions are the SMILE and SMILE+ model in The Netherlands (Tavasszy et al., 1998, Bovenkerk, 2005), the SLAM model for Europe (SCENES Consortium, 2000), the EUNET 2.0 model for the Pennine Region in the UK (Yin et al., 2005) and the model for Oregon (Hunt, 2003, Hunt et al., 2001). The first three model systems are based on aggregate (zonal) data[2], but unlike other freight transport models, include the use of consolidation and/or distribution centres, so that routes between a production zone and a consumption zone can be either direct (one leg) or indirect (multi-leg; these models can distinguish various types of multi-leg transport chains). The logistics module that we propose in this paper is specified (though not yet estimated) at the level of the individual decision-maker (or rather that of an individual business relation between a sender and a receiver). A more aggregate specification of a logistics model for Sweden, following the SMILE model in several respects, can be found in Östlund et al. (2002). The logistics component within the Oregon model for commercial transport also operates at the disaggregate level (in this case that of shipments). The logistics model that we propose below differs from what is being done in the Oregon model. Many steps in the Oregon model use existing distributions (e.g. to generate discrete shipments, allocate shipments to establishments, generate transshipment stops) based observed data (such as the US Commodity Flow Survey), from which random draws are made. This does not give a causal model in which endogenous variables are explained by exogenous variables, but a random process that just tries to replicate observed outcomes (descriptive model) without explaining them. Also in the Oregon model approach there are few policy variables that can be used to perform a policy simulation (the location of transshipment facilities can be changed as a policy measure). The logistics model presented below can be regarded as a causal and policy-sensitive model.

3. Inventory logistics

Large inventories reduce the risk of not being able to serve demand or use the required inputs in a production process. On the other hand, small and frequent deliveries lead to higher transport and stockout costs but lower inventory costs. This trade-off between transport and keeping inventories is part of the logistics model.

There is a trend towards control of the chain by the retailer ('factory gate pricing' in the food sector; Potter et al., 2003). On the other hand there is a tendency towards control by the sender ('vendor managed inventories' in for instance the petrochemical sector; Waller et al., 1999; Disney et al., 2003). In both cases however, the

[2] The SMILE model was originally specified as a disaggregate (nested logit) model, but because of lack of disaggregate data, it was later calibrated to aggregate data.

information on which the size of the inventory is based stems from the receiver. The demand for his products or the peculiarities of his production process are the key determinant of stock size and shipment size. We shall assume that the inventory decisions (especially shipment size/frequency) are generally made at the C (W) end, by the receiver. Inventories at the P end are determined by production considerations ('production-smoothing inventories', see Shirley and Winston (2004) for more information about this distinction). In the logistics module, we need not go into production scheduling and the trade-off between production costs and inventories, since production is already modelled in the PWC flows. By assuming that the PWC flows –the quantities of the goods flows- are already determined in an earlier (aggregate) model, we effectively partition firm behaviour in two parts: one part on location choice, production technology and choice of supplier or receiver and one on logistics decisions including shipment size, which is of course a simplification of real-world behaviour, but in our view a reasonable one.

It is not required to model both the shipment size and the frequency of ordering (and thus transporting) a good. If the total annual demand Q for the good is known (from the PWC matrices and the disaggregation to firm-to-firm flows we know the annual flows by commodity type), then Q = f.q, in which f is the frequency per year of ordering and transporting the good and q is the shipment size. Here we seek to model shipment size; and frequency will follow once we determine the optimal shipment size. Alternatively, we could model frequency.

The shipment size to be determined is the size of the shipment as it arrives at the destination end C. We assume that corresponding amounts of this good are produced and lifted at the P end, but in transport from the P zone, these amounts (shipments) may be consolidated into larger vehicle loads. A shipment is then defined as a certain quantity of the good that is ordered together and delivered together. It can exceed a full truckload, and, in the case of road transport, can consist of several trucks ('convoy'). The model does not explicitly take into account shipment size selection for optimal inventories at transshipment locations between sender and receiver (we assume the receiver determines the shipment size from his own perspective), though the time and cost involved in storing the goods at transshipment points are included in the logistics costs function.

As the shipment size increases, transport costs decrease, while inventory costs increase. The trade-off between transport and inventories is modelled in the Economic Order Quantity (EOQ) model (first formulated by Harris of Westinghouse Corporation in 1915 (Winston, 1987)). The optimal shipment size is found by minimising the sum of the total logistics costs. The solution is called the 'economic order quantity'. Different inventory theoretic model specifications have been derived for this problem (see for example Baumol and Vinod, 1970; Chiang, Roberts and Ben Akiva, 1981; Vieira, 1990 or Park, 1995).

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 logistic chain l:

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

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

The purchase costs of the goods from different suppliers are not part of the optimisation, since the senders and receivers of the goods have already been determined. However, the purchase costs do play a role through the capital costs of the goods that are included in the equation above.

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

$$G_{rskmnql} = o_k.(Q_k/q_k) + T_{rskql} + d.j.g.v_k.Q_k + (d.t_{rsl}.v_k.Q_k)/365 + (w_k + (d.v_k)).(q_k/2) + a\ .\ ((LT.\sigma_{Q_k}^{2})+(Q_k^{2}.\sigma_{LT}^{2}))^{1/2} \quad (2)$$

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)
j: the fraction of the shipment that is lost or damaged (might vary between modes)
g: the average period to collect a claim (in years)
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.
a: a constant to set the safety stock in such a way that there is some fixed probability of not running out of stock. For medium/high frequency products, a common assumption is that the demand (and lead-times) follows a Normal distribution. a will then be: $a = F^{-1}(CSL)$, where F^{-1} is the inverse Standard Normal Distribution and CSL is the cycle service level, that is the probability that the stock will not be empty during a replenishment cycle.
LT: expected lead-time for a replenishment (time between placing the order and replenishment)
σ_{LT}: standard deviation for the lead-time
σ_{Q_k}: the standard deviation for the yearly demand

The optimal shipments sizes will in the standard cases not be influenced by the safety stock, or vice versa. However, different transport alternatives with different transit times will have an impact on the safety stock through the lead-time (and possibly through the standard deviation of the lead-time), and thereby also on the inventory cost (and the total cost). This may be the case for alternative modes. In principle, lead-time should then be a function of the mode (h): LT = LT(h).

The most general situation, which is valid for most commodity types, is optimisation of both inventory and transport cost based on a common cost function. Then optimal shipment size is determined as:

$$-(o_k * Q_k)/q_k^2 + (w_k + d * v_k)/2 + \partial T_{rskq}/\partial q_k = 0 \tag{3}$$

The last term takes into account the effect of economies of scale in transport operation (bigger vehicles and vessels have lower unit cost) on shipment size. However, the transport costs depend on the transport chain that is chosen, and this is not yet known. In the prototype (version 0) logistics module for Norway and Sweden (see section 8) we have assumed that transport costs do not matter in the determination of shipment size. For most commodity types, the optimal shipment size is then determined as:

$$q_k = \sqrt{\frac{(o_k * Q_k * 2)}{(w_k + d * v_k)}} \tag{4}$$

For other commodity types, where the shipment size is not so much a decision variable but a constraint on the transport costs optimisation (e.g. when inventory costs are low), we have in the prototype used an exogenous shipment size.

The assumption that transport costs do not influence shipment size has been relaxed in the current version (version 1) of the model for Norway and Sweden. The outcomes of eq. (4) are now only used as a starting point, and twenty different frequencies are generated from this, all smaller than the starting point, because transport cost is a force that will work to increase shipment size (by moving to larger vehicles or vessels one can obtain lower transport costs per tonne). All twenty points are evaluated using the full logistics costs function (including transport costs) and the shipment size and frequency of the one with the lowest total logistics cost is chosen. We assume that shipment size is the same for all legs in the transport chain, though a shipment may be combined with other shipments in the same vehicle or vessel at consolidation centres.

4. Transport logistics

The simplest, but not necessarily cheapest, option would be to transport the shipment directly from P to C without using consolidation or distribution centres or transshipment. In this case there would be no costs of consolidation, distribution and transshipment. However, the pure transport costs per tonne are decreasing clearly with increasing shipment size: larger road vehicles and rail and waterway modes usually have lower freight rates per tonne. Therefore, especially for less than full truckload shipments, it is quite likely that the savings from direct transport in terms of transshipment costs are smaller than the additional pure transport costs.

Another alternative for this transport involves consolidation; using consolidation centres in the neighbourhood of the production location (see Figure 1).

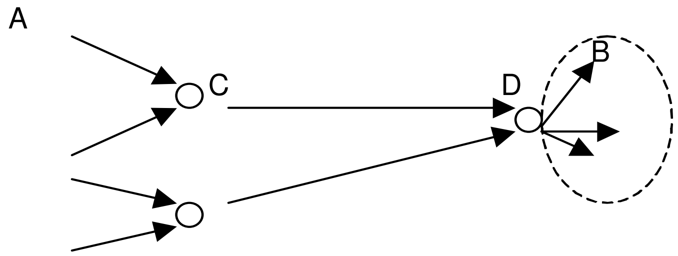

Figure 1 - Consolidation and distribution: the shipment of goods from location A to B is consolidated at C and distributed at D

Mode availability depends on the specific spatial relation studied, as well as on the commodity group k and the shipment size q. The total set of modes h in the logistics module consists of:

- Road transport (with different vehicle types/sizes);
- Rail (with different train types);
- Sea (with different vessel types);
- Air transport (with different types of aircraft).

Also the distinction between unitised shipments and non-unitised cargo is part of the mode definition h. We now introduce some further notation:

Table 1 - Symbols used for logistic chains

Sender	m
Receiver	n
Logistic chain	l
Commodity type (omitted)	k
Value	v
Mode/vehicle type/loading unit	h
Transshipment location	t
Leg	i
Number of legs	I_l

In the following we omit the subscript for commodity type. Commodity types are independent: all equations are simply repeated over all commodity types. We are also not using the zone subscripts r and s, because the locations of sender and receiver are implied in m and n.

The logistic chain *l* (l of Leonard) consists of a chain of modes and transshipment locations:

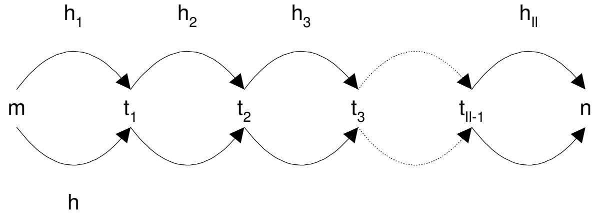

Figure 2 - Logistic chain

From sender m (producer P or wholesaler W) this is a transport of one or more (OD) legs to receiver n (C or W). We denote a leg of logistics chain *l* by i, and the number of legs of logistic chain *l* is I_l (I of Isaac, sub *l* of Leonard). The mode on the first leg is denoted h_1, the mode on the second leg h_2, etc., but there could be several modal alternatives at each leg (as an example we depicted two mode options per leg in Figure 2 above). Between the OD legs there are transshipment locations, which can be consolidation/distribution centres, ports, airports or intermodal rail terminals. At these locations (denoted t_1 till t_{Il-1}), goods change modes, but there can be temporary storage, i.e. to wait for large vessels with low frequency, as well.

The logistic chain can now be written as a series of mode-transshipment location points, one for each leg of the chain, with the last being a mode-receiver location pair (equivalently we could have started with the sender and the first mode):

$$l = \{(h_1, t_1), (h_2, t_2), \ldots, (h_{Il}, n)\} \quad (5)$$

Each pair indicates a leg i , i =1,..., I_l.

We can regard this as three sub choices within *l*:

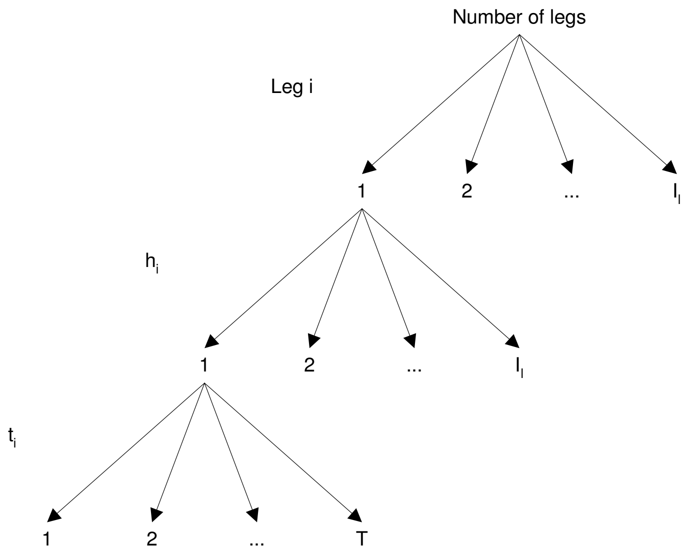

Figure 3 – Choices within logistic chain

The explanatory factors are included in the logistics costs function G_{mnl}

$$G_{mnl} = \sum_{i=1}^{I_l} specific\ to\ leg\ i + specific\ to\ the\ chain\ l$$

$$P(l|\{i, \ldots, L_{mn}\}) = P(l|L_{mn}) = \min_{l \in L_{mn}}[G_{mnl}] \quad (6)$$

L_{mn} = choice set of logistic chains per mn.

The choice of transport chain is determined on the basis of the same logistics costs function as used for shipment size. This function can be parameterised as:

$$G_{mnql} = \beta_{0ql} + \beta_1.(Q/q) + X_{mnql} + J_{mnql} + \beta_2.j.v.Q + \beta_3.(t_{mnl}.v.Q)/365 + (\beta_4 + \beta_5.v).(q/2) + a.((LT.\sigma_Q^2)+(Q^2.\sigma_{LT}^2))^{1/2} \quad (7)$$

Where:

β_{0ql} = alternative-specific constant

β_1 = o

β_2 = d.g

β_3 = d (in transit)

β_4 = w

β_5 = d (warehousing).

What we have done in eq. (7) is to include a number of items, such as order costs, storage costs and capital carrying costs, in the coefficients to be estimated. The reason for this is that data on these items will be very difficult to obtain. As a result, the coefficients have specific logistical interpretations. We could distinguish between the implied discount rate (d) of the inventory in transit (β_3) and of the inventory in the warehouse (β_5), because these need not be the same. Including revenues for the shippers and doing profit maximisation instead of cost minimisation is not required here, since the PWC flows (and therefore the sales) are already given. In the minimisation, we assume that firm size does not matter. This assumption may be relaxed to accommodate economies of scale in warehousing, ordering and transport. Also, variation of the discount rate for the inventory capital costs and of other preferences between firms could be included.

A random utility discrete choice model can be obtained by using minus the total annual logistics costs as the observed component of utility and by including random cost components ε that follow specific statistical distributions. These random components account for omitted variables that influence decisions on shipment size and transport chains[3], measurement errors and such.

$$U_{mnql} = - G_{mnql} - \varepsilon_{mnql} \qquad (8)$$

Where:

U_{mnql} : utility derived from logistic and transport choices

G_{mnql} : observed component of total transport and logistics costs

ε_{mnl} : random cost component.

Transport cost reductions for a shipment might be achieved by consolidating the shipment with other shipments in the same vehicles and vessels, at consolidation centres (including ports, airports and railway terminals). This possibility of sharing the transport costs with other shippers is included in the costs functions (by defining part of the transport costs as variable, and only including the shipment's part of the load of the vehicle/vessel for these). The degree of consolidation (and consequently the part of the variable costs that need to be borne by the shipment) depends on which vehicles are available in a transport chain and on the presence of other cargo that needs to be shipped from the same consolidation centre of that transport chain to the distribution centre of that chain. The load factor (vehicle load as percentage of vehicle

[3] For instance in an empirical specification without safety stock considerations, as used at the moment for Norway and Sweden, transport time reliability is missing from the logistics costs function, though this can be expected to influence shipment en transport chain choice.

capacity) therefore is an endogenous variable: the OD pattern that the logistics model produces, provides the required information on other cargo between terminals. This endogeneity problem can be solved by applying the model in an iterative fashion: the OD pattern of a previous model run is used to indicate whether for a shipment there will be other cargo with which it can be consolidated. In section 8 we describe how this was implemented for Norway and Sweden. This also includes consolidation by carrying out collection runs to different senders, without trying to predict the exact configuration of the collection tours at the P end and the distribution tours to several receivers at the C end; this would require solving numerous travelling salesman problems, which is clearly beyond the scope of this model.

5. Empties

Empty vehicles are an important output of a freight transport model, since these vehicles (or vessels) use the network, just as well as the loaded vehicles. In several freight transport models, the empty vehicles in some direction are made dependent on the goods flows in that direction. So every tonne transported from an origin zone r to a destination zone s will lead to the movement of an empty vehicle of the same size from the same origin r to the same destination s. However, this will get the directionality wrong; the empty flows are by nature reverse flows (Holguin-Veras and Thorson, 2002). It would be better to assign all product OD flows to vehicles first and then define another product: ‘empties’. The flows for these vehicles (similarly for vessels and aircraft) are the mirror image of the loaded OD flows: they go from s to r. This is what we do in the logistics module for Norway and Sweden: we calculate the loaded trips first and then use vehicle balances and let vehicles return to where they came from, with specific shares for empty and loaded return trips. In this formulation, the probability that some of the empty capacity will also be used for transporting goods in the opposite direction is taken into account.

6. Data requirements

To estimate the logistics module, as described in the sections above, requires information on the following items:

- Data on individual shipments (e.g. from a commodity flow survey or a shippers survey): sector of sender and receiver, origin and destination, value of the goods, modes and vehicle/vessel type and size used, cargo unit, shipment size/frequency, use of freight terminals (including intermodal terminals and marshalling yards), consolidation and distribution centres, ports and airports. Preferably this would be transport chain information: which shipments go directly from P to C, which use the above intermediate points?
- Data on where the freight terminals, consolidation and distribution centres, ports and airports are located;
- Data on transport and logistics costs: transport costs per vehicle/vessel kilometre, terminal costs, handling and storage costs for all available alternatives.

Only a few countries have available recent multi-modal information at the disaggregate level. Besides estimation on (partly) disaggregate data, the model can also be calibrated to aggregate data, such as data on modal shares by commodity type for aggregate zones and on the shipment size distribution.

7. Application to Norway and Sweden: data availability and ideas for estimation

The existing national model systems for freight transport in Norway (NEMO) and Sweden (SAMGODS) are relatively sophisticated models, judged by international standards in freight transport modelling. These models contain the conventional four steps (with mode choice and assignment being handled simultaneously in a multimodal assignment), as well as value-to-weight transformations and vehicle load factors. All steps are handled at the aggregate (zonal) level. Both models are lacking logistics elements, such as the determination of shipment size or the use of consolidation and distribution centres (though indirectly, transshipments between modes are included by using multi-modal network assignment).

The transport authorities of Norway and Sweden, working together as the Work Group for transport analysis in the Norwegian national transport plan and the Samgods group in Sweden, have commissioned RAND Europe[4] (now Significance) to develop a logistics module that will be part of their national freight transport forecasting systems and policy analysis. There will be two logistics modules, one for Norway and one for Sweden, but with a common structure and, if required because of missing data in one of the countries, also some common coefficients.

Sweden has available a Commodity Flow Survey (CFS) for the years 2001 and 2004/2005 with information on about a million individual shipments (SIKA, 2003). This includes the value and weight of the commodity, the sender and receiver location and the sequence of modes used. However, the locations of the transshipment are not included (except for ports in international transports). For Norway, there is no CFS, but probably the detailed shipment data from one of more of the biggest logistic service providers can be used. This data would include the locations of the consolidation and distribution centres used. Both for Norway and Sweden there might be targeted (e.g. for specific commodity types) additional data collection at the level of individual shipments

Data on the locations of the terminals, ports and airports are available in Norway and Sweden, though with some focus on the larger terminals. For the link-based transport we use transport distance and time from the networks as inputs to costs functions for each vehicle and vessel type. Transshipment cost information is also available for each type of transshipment and the information on the other logistics costs items comes from industry expertise.

7.1 Estimation for Sweden

In the CFS, transshipment locations are unobserved. For road we also do not know I_l (number of legs), since we do not have information on the use of consolidation and distribution centres. Because we have data on the sequence of modes used, we know

---

[4] The project for Norway and Sweden is carried in several phases, some of which involve partners: work on the logistics module has been carried out together with Solving International, Solving Bohlin & Strömberg, Michael Florian of INRO Canada and Stein Erik Grønland of SITMA.

how many transshipment locations have been used (with some observation error), in terms of changes between the nine modes distinguished in the CFS, but we do not know in which zones the transshipments took place. The problem can be stated as;

$$\tilde{G}_{mnl} = Min_{I_l, t_i, i=1,\ldots,I_l} \tag{9}$$

Given $\lfloor h_1, h_2, \ldots, h_{I_l} \rfloor$ Find the corresponding t_i's

For Sweden we shall use the full model specification (as above), but we note that the transshipment locations t_i are not observed. In the loglikelihood, we use the expectation of the outcome of the transshipment choice, which in the logit formulation will be the logsum (e.g. for the first leg):

$$\ln \sum_{t_1|h_1} \exp(-G_{mnl}) \tag{10}$$

By estimating the Swedish logistic chain choice model together with the Norwegian one (and if posible use new datasets in this at the same time), the parameters of all choices can be identified

7.2 Estimation for Norway

In the data from the logistic service provider(s) we will have information on the modes h and transshipment locations used t. But information on the sender and receiver (except their location) will be limited if not absent, and most importantly, we will not know commodity type k (only some information on the way of handling: e.g. refrigerated).

So, what we observe is a sum over commodity types. We may be able to link a commodity type (but not using a detailed classification) to information in the dataset, e.g. on the value (and/or weight) of the shipment:

$$P(l|\theta) = \sum_k P(l|\theta, k) \cdot P(k|value, \ldots) \tag{11}$$

Also, we can try to identify parameters by commodity type by combining the data with the data from Sweden and performing simultaneous estimation.

8. The 2005/2006 and 2006/2007 logistics models for Norway and Sweden

This concludes the specification of the logistics model, which is planned to be estimated on disaggregate data for Norway and Sweden in 2007 and beyond (when other shipment-level data will be available). In 2005/2006 a preliminary logistics model (a 'prototype' or 'version 0') for Norway and Sweden was developed on existing data, which assumes deterministic cost minimisation, but already uses micro-simulation. This prototype had a weak empirical foundation since it was not based on disaggregate estimation or aggregate calibration.

In 2006/2007 a version 1 model has been constructed; this is an improved deterministic cost minimisation model. In the version 1 model, the logistics cost minimisation takes place in two steps. In the first step (transport chain generation), the optimal transshipment locations (from the list of available terminals) are determined for each type of transport chain and origin and destination zone. In the second step (transport chain choice), shipment size and transport chain (number of legs, selection of modes and vehicle types) are determined by enumerating all available options for a specific firm-to-firm flow and selecting the one with the lowest logistics costs.

For Norway, the version 1 model uses all firm-to-firm flows, based on register data on the firms by number of employees and municipality, and no expansion is needed to go to the population of all goods flows in Norway. For Sweden a sample of firm-to-firm flows (for different size classes) is used for application of the disaggregate logistics choices, after which an expansion procedure needs to be used to arrive at population totals.

Unlike the prototype version, shipment size in the version 1 model depends on economies of scale in transport, through the transport cost function (a force leading to larger shipment sizes, because these have lower transport cost; see section 3). The degree of consolidation (or the load factor of the vehicles) between consolidation centres and distribution centres is determined in an iterative procedure which starts with an assumed average load factor, but in a subsequent iteration includes information on the availability of other cargo (based on the available transport chains and port statistics), and in an even further iteration uses the flows between consolidation centres predicted in the previous model iteration.

We have developed a procedure to calibrate parameters in the cost function to available aggregate data. A number of calibration parameters (e.g. for implied discount rates, mode-specific constants, constant for direct transport) is added to the cost function. We use observed OD data by mode, commodity type for aggregate zones (10x10 zones for a country) as calibration data. The calibration parameter values are then determined in an iterative process using the Box-Complex procedure (see Box, 1965, Balakrishna, 2006). This method belongs to the class of direct search methods, that do not require derivative methods (unlike gradient search), which is convenient given our highly nonlinear logistics cost function (which is a step function).

Below we give some results for the version 1 model for Norway. We assigned the P(W) side in Norway to more than 100,000 firms (senders) and the C(W) side to almost 400,000 firms (receivers). There are more receivers than senders because senders can only be firms producing goods or wholesalers whereas receivers include firms in all sectors (e.g. also including services). The number of firm-to-firm flows generated for Norway is 6 mln. This number refers to annual flows (business relationships), each of which can consist of several shipments. The program that was written for the 2006/2007 logistics model thus creates a file with 6 mln records. For each of those records we now have a sending firm (m) in some zone (r), a receiving firm (n) in some zone (s), a commodity type k and an annual total flow Q. After this, the shipment size (and frequency) and transport chain for this flow are determined on the basis of deterministic costs minimisation. These results are generated for each firm-to-firm flow (so we get a transport chain for every record) and added to the 6 mln firm-to-firm records. The 2006/2007 model therefore already is a micro-simulation model. From this large micro-level file, several more aggregate files can be derived. The version 1 model for Norway distinguishes 32 commodity types, transport chains of one, two, three and four legs, ten road vehicle types, 28 vessel types (including ferry), eight train types and two types of aircraft. The distinction between containerised/non-containerised is incorporated by defining container and non-container vehicles and vessel types. The runtime of this version 1 logistics model for Norway was up to 2 hours on a standard PC.

9. A mode and shipment size choice model estimated on the Swedish CFS 2001

The US Commodity Flow Survey has been analysed quite frequently (e.g. Sorratini, 2001, Vanek and Morlok, 1998, 2000). The Swedish CFS has not yet been used for transport modelling.

In a research project at the Institute for Transport Studies, University of Leeds, we have used the Swedish CFS 2001 to investigate mode choice and shipment size (the CFS 2004/2005 became available only very recently). More specifically discrete choice models have been estimated explaining these choices from characteristics of the shipper, the shipment and transport time and cost on the networks.

The CFS 2001 datafile we are using has 922,913 records. Each record is a shipment to or from a firm in Sweden, with information on origin, destination, modes used, weight and value of the shipment, sector of the sending firm, commodity type, access to railtracks and quays, etc. From this we selected a file of 748,952 outgoing shipments of Swedish firms (domestic transport and export, no import) for which we have complete information on all the endogenous and exogenous variables.

The endogenous (choice) variable in the model is a combination of mode and shipment size. The mode here refers to the mode used in Sweden. As mode alternatives we use: road transport, rail transport, sea transport and air transport.

Road transport includes transport chains that use passenger car, truck or minibus/van, but no non-road modes. Rail transport includes direct rail transport or chains with road access or egress to rail. Sea transport includes transport chains with direct sea transport as well as chains with road and rail access/egress to/from the port. Air transport includes chains with direct air connections and chains with road access/egress to/from the airport. Other mode chains are used only very seldomly (within Sweden), according to the CFS.

We classified the continuous weight variables 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 the mode and shipment size model are the 17 discrete combinations of mode and shipment size listed in Table 2. For air transport there were not enough observations in the three highest weight categories for inclusion in the model.

Table 2 also gives the number of observations by mode (the number of times this option is chosen in the CFS). Road transport is clearly dominant (in terms of the number of shipments). A road transport with a shipments size in the fourth and fifth category will be a convoy of several vehicles. Very large consignment sizes for sea are missing as imports are not included.

Table 2. Choice alternatives used in the model 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

The logistics cost function used in an approximation to (1), (2) and (7). It includes link-based transport costs, transshipment 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. 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). 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.

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
- Transshipment costs.

Furthermore, transport time was also taken from the networks, and used to calculate the capital cost on the inventory in transit.

The network information and the costs function information was assembled in the course of the 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), described earlier in this paper. 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), but simplified the costs functions to fit the 17 mode-shipment size combinations that we are using in the mode and shipment size model.

The estimation results for a multinomial logit model for mode and shipment size are in Table 3. Nested logit and mixed logit models on this data set (so far) did not lead to satisfactory results.

The pseudo rho-squared w.r.t. constants only is negative, but this statistic was calculated by comparing the log likelihood value of the estimated model (which has only three mode-specific constants) to a reference model with no less than 16 alternative-specific constants (the 17 choice alternatives minus 1).

The mode and shipment size choice model contains many very significant coefficients (owing to the very large sample size). If the sending firm has access to industrial rail tracks this greatly increases the probability of using rail, and if it has access to quay docking facilities this increases the chance of choosing water-based transport. Large senders are more likely to use rail, whereas 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. Building materials and minerals are unlikely to be transported in small road shipments, but larger the road shipment size the more likely it is to be chosen for this commodity type. Larger shipments of petroleum products, metal and chemical products are more likely to be transported by rail. Water transport is also more likely for chemical products, and even more so with large shipment sizes. For ores and metal waste, rail and water transport have a higher choice probability, except for the smallest shipment size. Machinery and equipment are more likely to use rail or water transport, except for shipments between 3.5 and 15 tonne. 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. Please note that the time- dependent link-based transport costs (labour and vehicle costs) have already been taken into account in the transport costs. The remaining time costs are related to the capital cost of the inventory in transit and maybe also to deterioration and safety stock considerations. For air transport these turn out to be substantial. Estimation of separate transport time times value coefficients for road, rail and water transport did not lead to significant coefficients.

Table 3. Estimation results for a multinomial logit model for mode and 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	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		

We also calculated some elasticities for the model presented in Table 3, using sample enumeration on the estimation data. We obtained road transport cost direct elasticities of around -0.5, rail and water transport costs direct elasticities that were all above 1 (in absolute values), air transport cost direct elasticities that were almost 0 and air transport time direct elasticities of around -2.

10. Summary, conclusions and directions for future work

In this paper, the general structure of a new logistics model has been specified. This structure has been worked out for Norway and Sweden, for inclusion in their national freight transport forecasting systems. Estimation of the logistics choices on disaggregate data is planned for both countries, but has not been carried out yet. The logistics model takes as inputs commodity flows from production to consumption zone. The logistics model then disaggregates these flows to firm-to-firm flows. After this disaggregation, the logistics decisions (shipment size, use of consolidation and distribution centres, mode and vehicle/vessel type and loading unit type choice) are simulated at this firm-to-firm level (micro-simulation). The basics mechanism for these decisions is minimisation of the total annual logistics costs function. The output of the model consists of flows between origins and destinations (OD-level), where consolidation and distribution centres (including ports, railway terminals) are also treated as origins and destinations. Furthermore, the model can provide information on total logistic cost between zones, which can be used in trade or spatial interaction models.

A prototype based on a deterministic function has been developed and applied in micro-simulation for Norway and Sweden in 2005/2006. This prototype has been extended and improved in 2006/2007, and is now calibrated to data on mode shares between aggregate zones and on the observed shipment size distribution. These implementations show that it is feasible to simulate the goods flows in a (not too large) country at the level of individual firm-to-firm flows and shipments within a reasonable run time. This micro-level logistics model does require aggregate flows between production zone and consumption zone (the disaggregation of these to the firm level can be part of the logistics model) as inputs as well as inputs from transport network models. It can only work in combination with models for these other tasks, such as multiregional input-output models and assignment routines.

The versions implemented so far use a deterministic logistic costs minimisation as the basic decision mechanism. The calibration to aggregate data (by commodity type) serves to obtain a better match to observed modal split data, and does this partly by including mode-specific calibration factors, that may represent generic modal preference for or resistance against some mode. Other than this, it does not take into account that there are more factors than logistics costs (including time components) that can influence decisions on shipment size and transport chains (as well as measurement errors) and will lead to variability in the outcomes. Such omitted factors may include reliability of travel time and flexibility in reacting quickly to unforeseen demands. To capture the influence of omitted factors and measurement errors, a model is required that not only includes mode-specific parameters, but that goes beyond deterministic optimisation by also including random components.

To estimate the random utility formulation of the logistics model, one needs data on individual shipments that give the transport chains (locations of transshipments, modes and vehicle types for each leg). For most countries, these are not available, and then the micro-level logistics model can be implemented in the version that is calibrated to aggregate data on mode shares.

Estimation of the random utility logistics model on disaggregate data is foreseen for future years for both countries. Some of these data are already available (e.g. the Swedish CFS), some still need to be collected. For Norway the new data might include the location and use of consolidation/distribution centres as part of transport chains by major logistics operators.

Besides this planned estimation exercise, there are a number of ways in which the proposed logistics model might be improved in the somewhat more distant future:

- In the logistics model there is consolidation of multiple shipments in the same vehicle/vessel but there are no explicit tours in the model. The model analyses freight transport looking at it shipment by shipment (or rather firm-to-firm flow by firm-to-firm flow). It is also looking at the amounts of other cargo at the same OD pairs (for consolidation), but does not try to predict which shipments will be combined together and how the vehicle tours will be made up. There is no explicit handler/carrier perspective. We think it would be very hard to include these things at the scale of a national model, but on the other hand there are interesting developments in the field of freight transport tour models, that might one day be combined with the shipment-based approach of our framework (see Stefan et al., 2005: micro-simulation of commercial vehicle tours, but no representation of shipments).
- Empty transports (see section 5) are dealt with in a way thay moves beyond common practice, but that can be further improved, especially if one would move in the direction of vehicle tour models (Holguin-Veras and Thorson, 2002).
- A national, and certainly an urban, transport model system also needs to include transport of services. The proposed logistics model does not include these. Ideas on how to incorporate such flows can be found in de Jong et al. (2004).

In a parallel research project, The Swedish Commodity Flow survey 2001 was used to estimate a model at the level of individual shipments. This model simultaneously explains mode and shipment size. Many commodity-type-specific variables were found to have a significant influence on these choices. Also, goods with a higher value per tonne are more likely to be transported in small shipments (up to 15 tonnes) because of inventory costs considerations (at the destination). Transport costs had a highly significant impact and the additional capital costs on the inventory in transit (also possibly representing some influences of deterioration of the goods and of keeping a safety sock) were only significant for air transport. The implied interest rate for these transports clearly exceeds the local capital market interest rates. Future work on this model may include re-estimation on the Swedish Commodity Flow Survey 2004/2005, models with other explanatory variables explaining some of the coefficients (observed taste variation) and with random (unobserved) taste variation in some of the parameters (e.g. time and costs) or with other more flexible substitution patterns.

References

Balakrishna, R., 2006. Off-line calibration of dynamic traffic-assignment models. PhD thesis, Massachusetts Institute of technology, Cambridge Massachusetts.

Baumol, W.J., Vinod, H.D., 1970. An inventory-theoretic model of freight transport demand. Management Science, 16, 413-421.

Ben-Akiva, M.E., Jong, G.C de, 2007. Choice analysis and market analysis for freight. Paper based on presentation at the colloquium in memory of Prof. Marvin L. Manheim, Antwerp, 2004.

Bovenkerk, M., 2005. SMILE+, the new and improved Dutch national freight model system. Paper presented at the European Transport Conference 2005, Strasbourg.

Box, M.J., 1965. A new method of constrained optimization and a comparison with other methods. Computer Journal, 8(1), 42-52.

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

Disney, S.M., Potter, A.T., Gardner, B.M., 2003. The impact of vendor managed inventory on transport operations. Transportation Research E, 39, 363-380.

Holguin-Veras, J., Thorson, E., 2002. Preliminary insights into the practical implications of modeling commercial vehicle empty trips. Paper presented at the European Transport Conference 2002, Cambridge.

Holguin-Veras, J., Jong, G.C. de, Xu, N., Maurer, H., 2007. An experimental economics investigation of shipper-carrier interactions on the choice of mode and shipment size in freight transport. Paper to be presented at the European Transport Conference 2007, Noordwijkerhout.

Hunt, J.D., Abraham, J.E., 2005. Design and implementation of PECAS: A generalized system for the allocation of economic production, exchange and consumption quantities. Chapter 11 in: Foundations of Integrated Land-Use and Transportation Models: Assumptions and New Conceptual Frameworks. Elsevier, Oxford, 217-238.

Hunt, J.D., 2003. Agent behaviour issues arising with urban system micro-simulation. European Journal of Transport Infrastructure and Research 2(3/4), 233-254.

Hunt J.D., Donnelly, R., Abraham, J.E., Batten, C., Freedman, J., Hicks, J., Costinett, P.J., Upton, W.J., 2001. Design of a statewide land use transport interaction model for Oregon. Proceedings of the 9th World Conference for Transport Research, Seoul, South Korea.

Jong, G.C. de, Gunn, H.F., Walker, W., 2004. National and international freight transport models: overview and ideas for further development. Transport Reviews, 24-1, 103-124.

Marzano, V., Papola, A., 2004. Modelling freight demand at a national level: theoretical developments and application to Italian demand. Paper presented at the European Transport Conference *2004*, Strasbourg.

Östlund, B., Janson, G., Burgess, A., Ruijgrok, C., Tavasszy, L., Petersen, M.S., 2002. A logistics module for SAMGODS, TFK –Institute för transportforskning, Stockholm.

Park, J.K., 1995. Railroad marketing support system based on the freight choice model. PhD dissertation, Massachusetts Institute of Technology.

PbConsult, 2002. Commercial transport and transport supply components of the TLUMIP second generation model. Third Oregon Symposium on Integrated Land Use and Transport Models, 23-25 July, 2002.

Potter, A., Lalwani, C., Disney, S.M., Velho, H., 2003. Modelling the impact of factory gate pricing on transport and logistics. Proceedings of the 8th International Symposium on Logistics, Seville, 6-8 July 2003.

RAND Europe, Solving, 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, 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, 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.

SCENES Consortium, 2000. SCENES European transport forecasting model and appended module, technical description. SCENES deliverable 4 for the European Commission DG-TREN, ME&P, Cambridge.

Shirley, C., Winston, C., 2004. Firm inventory behavior and the return from highway infrastructure investments. Journal of Urban Economics, 5, 398-415.

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).

Stefan, K.J., McMillan, J.D.P., Hunt, J.D., 2005. An Urban Commercial Vehicle Movement Model for Calgary, Alberta, Canada. Transportation Research Record 1921, 1-10.

Tavasszy, L.A., Vlist, M. van de, Ruijgrok, C., Rest, J. van de, 1998. Scenario-wise analysis of transport and logistics with a SMILE. Paper presented at the 8th WCTR Conference, Antwerp.

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

Vanek, F., Morlok, E.K., 1998. Reducing US freight energy use through commodity based analysis. Transportation Research D, 5 (1), 11-29.

Vieira, L.F.M., 1990. The use of stated preference data in freight transportation demand analysis. PhD thesis, MIT, Cambridge, Massachusetts.

Waller, M., Johnson, M.E., Davis, T., 1999. Vendor managed inventory in the retail supply chain. Journal of Business Logistics, 20(1), 183-203.

Winston, W.L., 1987. Operations research: applications and algorithms. Duxbury Press, Boston.

Yin, Y., Williams, I., Shahkarami, M., 2005. Integrated regional economic and freight logistics modelling: results from a model for the Trans-Pennine Corridor, UK. Paper presented at the European Transport Conference 2005, Strasbourg.