PREFACE

This report is volume one from a study of data requirements and methodologies for goods transport modelling. The study is intended to provide the basis for a general discussion about the application of goods transport models in Denmark. Volume one provides an overview of different types of models and data availability. Volume two discusses in detail the problems associated with data collection for discrete choice models.

The writers are responsible for the conclusions and recommendations in this publication, which are not necessarily shared by the Danish Transport Council.

The Danish Transport Council, February 2002.

1 SUMMARY AND CONCLUSIONS

This chapter provides a summary of the main findings of the goods transport modelling study. The study has had a two-fold purpose;

1) Development of a common framework for goods transport modelling.
2) Identification of data collection methods that can provide the necessary inputs.

The study has been reported in two volumes. The first provides an overview of different types of models and data availability. The second provides a detailed discussion of the problems associated with data collection for discrete choice models. In this chapter findings from both volumes are summarised.

The summary is divided into three sections.

The *first section* points out some conclusions and recommendations concerning the modelling and data requirements, based on available studies.

The *second section* is devoted to recommendations related to data collection for discrete choice models.

Finally, the *third section* suggests some improvements or modifications of the present Danish statistics.

1.1 MODELLING AND DATA REQUIREMENTS

The study has been looking into different types of models and into methodologies applied in different studies. The report discusses the possible application of such models in a Danish context. An important source has been the Swedish goods model, which has been developed as part of the Swedish long-term infrastructure planning. In Denmark long-term infrastructure planning is not institutionalised. Therefore, the design and application of models should support the Danish planning needs within the context of goods modelling.

The Danish needs for a recurrent model basis include physical infrastructure maintenance. The most important wear of the infrastructure is attributable to the goods transport on road and rail. Therefore, there is a need to predict with a reasonable accuracy the future traffic on road and rail in terms of vehicle/wagon types and their volumes.

Another important aspect that Danish goods transport models should be able to address is the impact of changes of the overall framework for goods transport services. Effects of changes of cost relations between transport modes, and effects of changes of weight and dimensions of goods transport modes, are aspects where a goods transport model could bring about necessary decision support.

A number of EU research projects has dealt with the development of freight transport models. The most recent advances have been research on how to integrate new logistical concepts in the models. Several EU projects, among them REDEFINE and SCENES, have suggested ways of doing this. The very close relationship between commodity classifications and modelling methodologies is stressed in these two projects.

The findings of REDEFINE have been applied in the present Swedish SAMGODS model. Therefore, an important knowledge about the strong and weak sides of integrating the 'logistical family' and 'transport chain' concepts are already available.

Data for modelling:

A main focus of the present study has been on how to relate models with relevant statistics. In several cases the specific models have been developed using data that were collected during the establishment of the model. However, a number of Danish goods transport models have relied on available data concerning Danish import and export, and transport statistics by transport mode. Due to a recent change in the registration procedures concerning statistics on intra-EU trade, the trade data have deteriorated. As a consequence, it is not any longer possible to establish an adequate description of import and export by transport mode.

A number of EU research projects have dealt with the discrepancies between the needs for data and the traditional statistics collated over the last 30 years. This involves the statistical needs related to a description of the presently more complex transport chains, use of terminals, transport needs for different commodity types, etc. The recommendations address the following issues:

- The linking of trade and transport statistics, in order to identify modal distribution and transit flows.
- The harmonisation of data classification including freight flow indicators.

A summary of some of the recommendations is provided on the following pages.

TRANSITIE:

Most of the TRANSITIE pilot study recommendations and findings, aiming at improving Dutch transport statistics, are of some relevance to the Danish situation as well. The aim of the study was to check whether it is possible to get a consistent and reliable picture of the international goods flows to, from and through the Netherlands by other means than the present trade and transport statistics. Among the recommendations of the study are:

- Development of a procedure to elaborate transport data on the basis of monthly files on international trade, along with a harmonisation between transport and international trade;
- Conversion table to be provided of net to gross weights, on the basis of recent data calibration at a detailed commodity specification level;
- Improved procedure to establish transport chains from international trade data on disguised transit;
- Analysis of the overlap between entrepôt clearance to the country (the Netherlands) and entrepôt storage as well as clearance to other EU Member States.

The TRANSITIE study points out the need for a methodology linking import/export data and data on entrepôt transport with the transport database. The remaining transport flows should then relate to transit flows. For the determination of transport chains in intra-European transit, a separate methodology will have to be developed that excludes transit in relation to third countries.

The study suggests that the following variables should be considered when linking trade and transport data:

- Mode of transport (sea, rail, road, pipeline, air).
- Direction (incoming/unloading or outgoing/loading in the country in question).
- Country of origin/loading or destination/unloading.
- Country region of origin/loading (probably aggregated at the NUTS 3 level, i.e. the county level).
- Commodity group (NST/R-4 digit level).
- Container indicator (yes, no or unknown).

The TRANSITIE pilot study has revealed several procedural differences between the various data sources. Among these differences are the registration of the goods type, definition of the country of origin and destination, the port of registration in the case of a ship visiting several ports, and the degree of containerisation of the goods transported.

However, the pilot study managed to obtain a linking of 81% of both export and import flows with the trade data. This was done by taking into consideration the direction (incoming/unloading or outgoing/loading in the Netherlands), the country of origin/loading or destination/unloading, and the Dutch region of origin/loading or destination/unloading. It is even suggested that a 90% linking could be reached by taking into account some trade flow unbalances concerning the ports of Amsterdam and Rotterdam, oil imports from Saudi-Arabia etc.

A quality check on the transport/trade data is recommended for all countries.

A Danish study, similar to the TRANSITIE case study, should be considered, linking trade and transport statistics in order to identify the geographical distribution pattern including transit flows.

MESUDEMO:

The EU funded MESUDEMO research project recommends that the following aspects related to goods transport databases should be kept in mind when aggregating the necessary transport information:

- A harmonised classification should be used to enable keeping hold of as much detailed information as possible, such as information on the load (consignment).
- Some individual items should be considered as new variables, such as number of trips, number of consignments etc.
- Some additional variables should be calculated, e.g. weight multiplied by consignment distance to arrive at tonne-kilometres, capacity multiplied by trip distance (loaded or unloaded), etc.
- Consistency should be ensured with the classification used in trade databases for building transport chains.
- Consistency should be ensured with the classification used in other parts of the transport policy information system (socio-economic data, passenger flows, infrastructure links and nodes, impacts).

These aspects are considered in the proposal for goods transport indicators that is elaborated in Chapter 5.

Potential applications in Danish transport modelling:

As mentioned above, the development and application of goods transport modelling in Denmark have up till now relied on the availability of trade and transport data. Since information about the links between trade data and mode of transport has deteriorated, it is clear that in the future more emphasis will be placed on the transport specific statistics, such as the truck surveys, elaboration of port databases, etc.

However, trade data are continuously collected and processed. The information that is being lost is about the mode of transport used at the border crossing. This particular aspect is becoming less relevant, due to the increased use of transport chains. In the latter case several different transport modes are linked together in a transport chain, which brings the goods from the origin to the final destination. Therefore more information is required about the composition of the transport chains.

The Swedish SAMGODS model suffers from a similar lack of transport chain information. The modal and route choice model of the SAMGODS model (STAN99) is able to establish these transport chains based on a calibration of calculated loads per infrastructure link against the observed load per link. However, data on transport chains are not available for the calibration and validation of the model.

In the Fehmarn Belt transport model the purpose was to establish the mode and route choice across a precisely defined screen line. The modal split on this screen line could be observed by means of a traffic analysis. It is, however, not possible to distinguish different transport chains in this model. Therefore in this case the modal and route choice model is much simpler than in the SAMGODS model.

The trade data collected in EU countries are basically registered based on the Combined Nomenclature (CN). The CN commodity classification forms the basis for different classifications, e.g. the NST/R, which is used throughout EU for the description of commodities in transport statistics. It should be ensured that the CN classification enables a conversion to the NST-2000 classification to be eventually applied by the EUROSTAT.

The mode specific data that are collected in Denmark have been analysed, with a view to possible improvements that allow logistical characteristics to be included in the statistics.

A future grouping and classification of transport statistics, with a view to freight transport modelling, could take the point of departure from the INFOSTAT and MESUDEMO research project recommendations that are summarised in Chapter 5.

A preliminary grouping of the areas of *strategic transport model application* could be as follows:

- Effects of national transport policy measures on traffic pattern and modal distribution.
- Effects of specific transport infrastructure investment projects on traffic pattern and modal distribution, and the socio-economic and environmental impact of the projects.
- Assessment of transport network development plans.

A distinction should be made between bulk and goods organised in loading units, and between low-value, medium-value and high-value goods. Also the REDEFINE study points out the need for density information, in order to catch volume characteristic consignments. This is of particular relevance to the modelling of modal choice.

1.2 DISCRETE CHOICE MODELS

The present study has also dealt with the particular data catch problems related to establishment of discrete choice models. These constitute a specific category of the 'disaggregate' type models that are described in Chapter 3. A more thorough description of the problems illustrated with problems identified in Danish modelling work has been provided in volume two of this report. Some of the main findings are repeated below.

However, for a more thorough discussion of discrete choice models and the related data, the reader should consult volume two.

Discrete choice models have been in use since the beginning of the 1980s. They are probabilistic models where the probability of an individual person choosing a given option is a function of his/here socio-economic characteristics and the relative attractiveness of the option (Ortuzar, J. and L.G. Willumsen, 1994/2001).

Some useful properties of discrete choice models are indicated below.

Discrete choice models are based on theories of individual behaviour. Thus it is likely that these models remain stable in time and space.

Discrete choice models are estimated by using individual (singular transport) data.

Discrete choice models are probabilistic, i.e. they yield the probability of choosing an alternative from the choice set and do not indicate a priori which one that is selected.

Discrete choice models have improved the travel demand modelling due to the following innovations:

- The importance of individual behaviour in travel demand modelling has been recognised.
- Compared to the aggregate models, the disaggregate models is able to capture more information from the data. Furthermore, disaggregated data is not a subject to significant statistical biases, which may occur in aggregated data.
- New theories for travel demand modelling, such as the theory of *random utility*, have been developed.
- The application of new types of data, such as Stated Preference (SP) data, has been developed.

The theoretical background of the discrete choice models seems to be sound. Establishing the functional relationship between the respondent's preferences and the attributes of model travel alternatives (e.g. travel cost and travel time) in the form of a *utility function* has found no critics in the literature. The *logit models*, which are the mostly applied types of discrete choice models, seem to cover the whole variety of needs from these types of models.

Discrete choice models are estimated based on data for individual transports. Data about actual transports are referred to as Revealed Preference (RP) data. An important aspect of estimating discrete choice models is the use of data about hypothetical behaviour (Stated Preference (SP) data). The introduction of SP-data has to a certain degree circumvented the problem of correlation between variables in RP-data, e.g. transport time and transport costs could be interrelated, and the use of SP data makes it possible to estimate the effect of completely new measures, e.g. a new piece of infrastructure.

Discrete choice models have been used for the last 10 years in Danish goods transport modelling. Experience has shown that the developed models are increasingly sophisticated as are the use of the models. However, the developed models are not without problems. Many problems are related with data about the individual transports, and in the present work a thorough analysis of three different Danish projects in which the authors have participated have been analysed in order to identify possible shortcomings and possible ways of improving the collection and use of data.

The three projects are quite different, but for all projects following general observations concerning data could be deducted.

- The quality and the amount of available data play an important role in discrete choice models. For a number of reasons, many important observations are difficult to obtain from respondents. This applies first of all to transport costs, which might be a company's commercial secret or might not be known (in the cases where transport is part of the production process of the company in question). The uncertainty of goods forecasts is therefore strongly related to the uncertainty of measuring the importance of the cost parameter.

- It is shown that rating and ranking exercises (RP-exercises) produce different results compared to SP-experiments. Particularly the importance of transport costs seems to be strongly increased in the SP-experiments compared to ratings and rankings. One reason is that the SP-experiments do not reflect, in a true manner, the way that the companies comprehend some of the important parameters in ratings and rankings. Therefore, there is a major need to reconsider the way SP-experiments are carried out, particularly the methods for evaluation of parameters related to the quality of the transport services.

- It has also been shown that the measurement of 'soft variables' (e.g. 'information system' and 'customer service') is rather difficult in discrete choice models. Such variables are of no use unless they are decomposed and included in great detail in the SP-experiments.

- In many cases the interviewed companies have never applied more than one transport mode, despite that the type of shipment they work with could be an inter-modal transport. The presented mode or route alternative in a SP-experiment needs to be based on the calculated values for transport costs etc. for the particular shipment and the zone-pair. Assessment of alternative transport costs and transport times for the SP-interviews are usually quite simple and not so accurate. Therefore, there is a need to improve the methods for assessing level of service, that is transport costs, transport time and possible other parameters. E.g. assessment of level of service needs to take into account regulations (e.g. regulations for working and resting time), empty trips and possible discounts that the respondent might obtain for example from ferry operators. Only if the hypothetical situations in SP-experiments are realistic, they could lead to a realistic trade-off between the offered alternatives, which is a necessary requirement for a successful survey.

- Goods transport matrices might change fast because goods transport is a function of the demand for different commodities in the market, companies' degree of competitiveness, changes in technology, and price changes. Therefore there is a need for a permanent updating of goods OD-matrices.
- There is an obvious need for more goods data of the disaggregated form. This is particularly related to the revealed preference (RP) data. It should be considered to carry out analyses similar to the continuous analysis on goods flows carried out by the Swedish Statistical bureau. In this way a large number of RP data will be made available for both model estimation and forecasting. It should be noted that the RP data constitute the core of model forecasts. The problem of the required amount of data is also reflected in the issue of segmenting the goods models, i.e. estimating different models for different commodity groups or different logistical families. Each new segment demands further data. If the available data are too few, the model's accuracy (i.e. the quality of the obtained estimates) could be so poor that the forecasts achieved are unreliable.

RP data requires specific data collection. Although the present goods surveys comprising Danish trucks in national and international transports are based on individual trips, many pieces of information concerning the goods is not collected, e.g. transport price and transport time. This information is a requirement for RP-data.

The possible users of discrete choice goods models are both planning bodies (e.g. transport ministries, directorates), transport producers (carriers) and transport buyers (shippers). It is therefore an imperative to establish a platform where data can be obtained at a broad level, in order to facilitate the future goods transport planning in Denmark.

From a theoretical point of view, there is a need for a permanent theoretical and, more importantly, methodological improvement of the discrete choice goods models. The present study has underpinned the need for improvement of the surveying methods, improvement of the design of SP-experiments, and improvement of the assessment of level of service (i.e. better assignment models).

1.3 PROPOSALS FOR IMPROVEMENT AND MODIFICATION OF DANISH STATISTICS

With a view to improving freight transport information, there is a particular need to modify some of the transport operators' reporting to the Statistics Denmark. This includes a harmonisation of ferry transport reporting and supplementary information about rail transport. Such modifications and improvements are not very comprehensive and costly.

In relation to freight transport modelling, the following suggestions for modifications of the statistical sources would immediately improve the data situation considerably.

Ferry traffic statistics:

All ferry operators should be requested to provide a breakdown on types of trucks and trailers, in compliance with the data breakdown already provided for several national and international routes.

Periodically (e.g. every five years) there would be a need to estimate and calibrate the average loading figures per transport unit based on data samples, and to carry out ad hoc surveys to obtain some basic O/D information.

Port and maritime transport statistics:

The existing statistics is well structured and quite detailed. In general it appears sufficient if combined with other statistical sources. For specific purposes, it is possible to obtain detailed computerised data extracts from the operational databases of the individual ports and terminals. A set of data access criteria and guidelines should be formulated in relation to the aggregation level required, in order to avoid violating the confidentiality of commercial information.

It is recommended to establish a database for monitoring maritime transport. The database would provide details concerning the ship traffic through the Danish straits, including origin and destination as well as type and size of the ships.

Rail transport:

A supplementary geographically specification is required on goods flows by rail via the few major import/export outlets, and via the Great Belt link and other major sections. This specification could also include throughput figures via the largest combined terminals.

Terminal data extracts and traffic counts:

With a view to obtain more specific information about the spatial flow pattern and other characteristics of freight transport, it should be considered to introduce regular and cost-efficient data extracts from container, ferry and major combined transport terminals.

Automatic traffic counting devices should be installed at major road network nodes and sections, in order to obtain more specific information about the lorry traffic volumes and composition. These counts would need to be supplemented by periodically performed manual counts to obtain more detailed and specific information, and by a few sampled stop-interviews to obtain some basic O/D information.

Major outlets and locations for data catch:

Some major locations and outlets for potential freight transport and traffic data catch are indicated on the following overview map (Figure 1.1).

Advanced information and communication technology (ICT) could be applied.

OD-surveys:

The lack of O/D information is still a major drawback in relation to the application of freight transport modelling. Even taking into account that the present statistical reporting of goods transport flows would be considerably improved, the future description of goods transport flow patterns will need, to an increasing extent, to rely on modelling in combination with ad hoc OD-surveys and data sampling analyses.

A national regulation should be issued, to ensure that data obtained from publicly financed OD-surveys are made available to all interested stakeholders. Notably this applies to the surveys carried out in connection with the analysis of large-scale infrastructure projects. Any restrictions

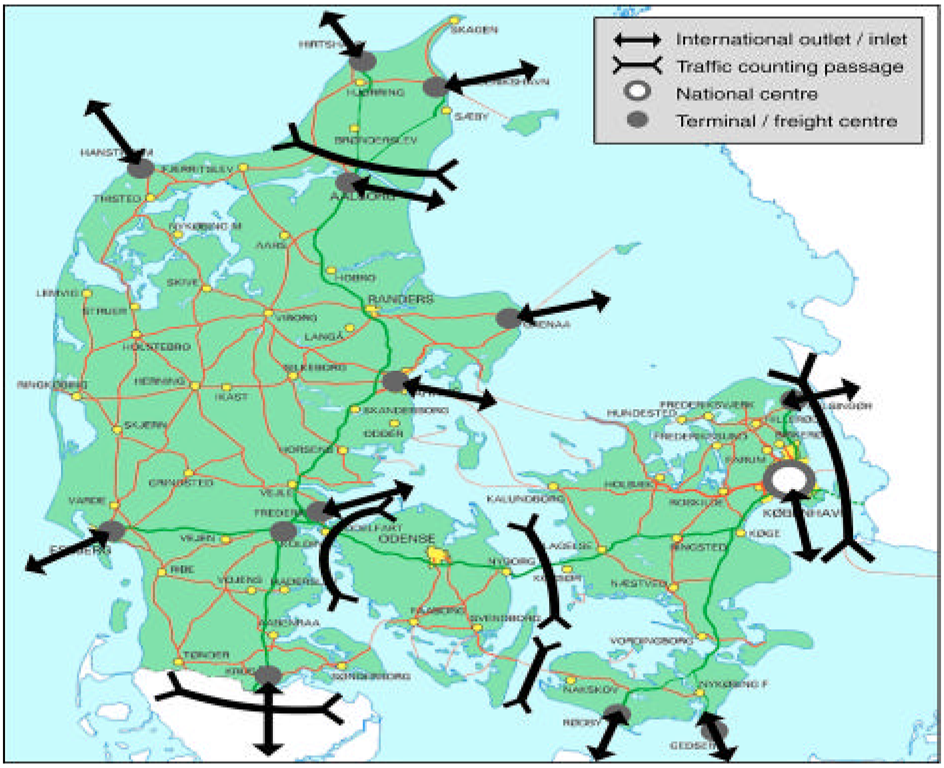


Figure 1.1: Major terminal locations and outlets for freight transport flows and traffic

should be avoided, except for those needed to ensure commercial confidentiality at the company level.

To facilitate the development of aggregate models comprehensive national OD-surveys could be carried out at for instance 5 years interval. The models could be used to calculate synthetic OD data for the intermediate years.

Pilot studies:

There is a need to carry out a few pilot studies in order to obtain the necessary empirical basis for introduction and specification of the modifications suggested above.

A Danish pilot study similar to the Dutch TRANSITIE study would contribute to the development of a methodology for harmonisation of data, identification of modal distribution, and estimation of transit flows.

In order to develop the methodology for estimating regional trade flows, a study could be carried out in order to analyse how to extract in a cost-efficient way the geographical information that is embedded in the INTRASTAT reporting (see Chapter 5).

Interviews and other relevant survey methods should supplement the pilot studies, in order to link the statistical analyses with O/D- and modal distribution information.

2 INTRODUCTION

This introductory chapter *first* describes the aim, scope and focus of the study.

Secondly, the structure and contents of the report is presented.

Thirdly, an overview is provided of the abbreviations and terms used in the report.

2.1 AIM, SCOPE AND FOCUS OF THIS STUDY

During the recent years, a number of research projects and studies have been carried out about the goods transport pattern in Denmark. The resulting models have included both national and international goods transport. In relation to these models, research efforts have focused on different topics, such as elements to be included in the description of modal choice and route choice for different types of goods. Studies have also been carried out on establishing values of travel time for goods transport.

These efforts have not yet lead to any presentation of a complete and open framework for the continued research and development on freight transport modelling. Thus the *prime target for the present study* has been to contribute to the development of a common framework for goods transport modelling. This should take a form, which facilitates that international experiences could be effectively exploited within a Danish national perspective.

A *related aim of the study* is to contribute to the identification of data collection methods that are able to provide the necessary inputs to the development of models for freight transport flows. Such inputs are considered important because the present methodologies have shown weaknesses as regards the description of supply and demand for goods transport.

Problem issues:

Hitherto it has been a major and unresolved problem to establish a uniform, consistent and representative description of the goods transport pattern in Denmark. This problem has been further accentuated during recent years, as a number of customs based statistical *data concerning the international goods flows have vanished* due to the introduction of the internal market of the European Union (EU). It is therefore evident that there is a need to evaluate how it would be possible to establish databases, which can be used as a policy-decision support tool and more specifically for the development of transport models.

A second problem is related to the development of behavioural models for goods transport, such as modal choice and route choice models. The definition of the models reflects a tradition for major resemblance between passenger and goods transport models. A number of the actually applied transport models have been based on discrete data, i.e. the so-called 'Revealed Preference' (RP) data and 'Stated Preference' (SP) data. However, it is obvious that the description and modelling of goods transport and passenger transport respectively would require quite different approaches. As a consequence, there is a major need to establish more *appropriate procedures and structures for the collection of behavioural data* concerning goods transport.

Focus of the report:

The present report concentrates on model structures, potential model applications, and the related needs for transport demand data. Thus the report includes a description of the following aspects:

- Model types and their applications
- Data requirements related to modelling
- Description of the needs for goods transport data
- Statistics and data collection methods
- Classification, definitions and concepts related to goods transport, including inter-modal transport, logistic families and commodity groups.

Focus of the report is on the data related to modelling of goods transport flows both on aggregate (macro) and disaggregate (micro) levels. A discussion is carried out as to how the statistics presently collected by the official agencies can be improved for modelling purpose. Also a discussion is carried out concerning the data need for disaggregate modelling, including the methodological difficulties encountered in the definition of experiments aimed at collecting disaggregate data. The latter discussion however is reported in volume two of this project.

Due to the specific aim of this report, the following topics fall outside its limits and scope, although they are of relevance to transport modelling and strategic transport policy formulation as well:

- Socio-economic data of the zones in question. These data are crucial for establishing cause-effect relations between societal variables and transport systems.
- Transport supply data, such as data about transport networks/infrastructure and services, unless they are considered of relevance to the basic definitions and concepts.
- Structural development of the transport sector industry.

One of the points of emphasis in this report is transport demand data in the form of foreign trade data and freight flow data, and how these data could be combined in order to provide data inputs to freight transport modelling. The statistical sources include trade data, data on goods transport flows between zones, and goods traffic data catch in terminals (including ports) and along routes.

Goods transport modelling involves transport demand data for several components of the sequential modelling steps. These steps include the socio-economic potentials of a particular zone, trade patterns between the zones, indicators of goods flows between the zones, modal distribution of the flows, and assignment to the transport network. The latter step also has to consider the conversion of transport flows to traffic flows.

The study has been confined to the modelling of goods transport flows that are completely internal in Denmark or represent import and export flows. The freight transit flows through Denmark constitute a particular statistical problem as they are included in the observed traffic but not covered by the foreign trade statistics. An evaluation of transit traffic would imply that the data analysis should be extended to at least all of the Nordic countries and northern Germany. This is not seen as a realistic possibility within the present study scope. Nevertheless, it is expected that the definitions and concepts will also be applicable to transit traffic through the Danish territory.

2.2 CONTENTS OF THE REPORT

The content of this report is organised in four main chapters.

Chapter 3 gives a brief introduction to goods transport models including important models developed in Denmark and abroad. Chapter 3 also includes a brief discussion of a national goods transport model for Denmark.

Chapter 4 presents some general considerations about the relations between transport modelling and transport data needs. It comments briefly on data categories and relevant EU research activities.

In Chapter 5 and Chapter 6 the focus is on how to improve the data structures.

Chapter 5 presents a short inventory of Danish goods transport statistics and suggestions on how to improve it.

Chapter 6 is based on European experience. The chapter includes definitions in terms of zoning and commodity classification, and missing data elements that are required in order to carry out an up-to-date description of goods flows.

Volume two of the report presents a discussion of the need for strengthening the data collection procedures in relation to Danish experiences on disaggregate ('discrete choice') modelling.

2.3 ABBREVIATIONS

The following table presents a list of abbreviations used throughout the report.

Abbreviation	Description
AKF	**A**mternes og **K**ommunernes **F**orskningscenter (Research institute of the Danish counties and municipalities)
CN	**C**ombined **N**omenclature
CONCERTO	EU Concerted Action Committee on information systems.
CPA	**C**lassification of **P**roducts by **A**ctivity
CTSE	Community **C**lassification for **T**ransport **S**tatistics in **E**urope
ETIS	**E**uropean **T**ransport Policy **I**nformation **S**ystem
EXPEDITE	Expert-system Based Predictions of Demand for Internal Transport in Europe (EU research project)
GEOSYSTRANS	Geographical database on European transport (EU research project)
GISCO	Geographical Information System for the Community (EU project)
ICT	**I**nformation and **C**ommunication **T**echnology
INFOSTAT	Information systems and statistics for strategic transport planning in Europe (EU research project)
LOS	**L**evel **o**f **S**ervice
MESUDEMO	Methodology for establishing general databases on transport flows and infrastructure networks in Europe (EU research project)
MYSTIC	Methodology for Statistical Analyses, Modelling and Data Collection (EU research project)
NST	**N**omenclature uniforme des merchandises pour les **S**tatistiques de **T**ransport
NUTS	**N**omenclature des **U**nités **T**erritoriales **S**tatistiques
OD	**O**rigin-**D**estination

REDEFINE	Relations between demand for freight transport and industrial effects (EU research project)
RP	**R**evealed **P**reference
SCM	**S**upply **C**hain **M**anagement
SIKA	The Swedish Institute for Transport and Communication Analysis
SITC	**S**tandard **I**nternational **T**rade **C**lassification
SMILE	Strategic Model for Integrated Logistics and Evaluation (Dutch research project)
SoftIce	Survey on freight transport including a cost-comparison for Europe (EU research project)
SP	**S**tated **P**reference
STAN	**S**trategic **T**ransportation **A**nalysis **M**odel
STEMM	**S**trategic **E**uropean **M**ultimodal **M**odelling
SUE	**S**tochastic **U**ser **E**quilibrium
TRILOG	Overview of global supply chain management (EU research project)
VTS	**V**essel **T**raffic **S**ystem
TRANSITIE	Dutch pilot study on goods transport data

2.4 GLOSSARY

The following overview indicates some of the terms used in the report. The overview is not exhaustive.

Aggregate models:
Dealing with aggregated transport flows of a county, city, region, country, and other administrative or geographically defined unit.

Disaggregate models:
Dealing with specific consignments of an organisation, a company or a specific project.

Discrete choice models:
A discrete choice model is a probabilistic model where the probability of an individual choosing a given option is a function of his socio-economic characteristics and the relative attractiveness of the option.

Entrepôt:
Temporary warehousing of goods in transit (storage and export).

Entropy maximising algorithm:
The entropy method, when applied to transport studies is concerned with first determining an expression to calculate the number of possible ways for establishing a set of trip matrices, and subsequent finding a trip matrix with a maximum probability of occurrence subject to origin and destination constraints.

EXTRASTAT:
Reporting method for trade with non-EU countries (extra-EU trade statistics).

Inter-modality:
The Commission writes the following definition in COM(97) 243 final: "Inter-modality is a characteristic of a transport system, that allows at least two different modes to be used in an integrated manner in a door-to-door transport chain".

INTRASTAT:
Reporting method for trade with other EU countries (intra-EU trade statistics)

Logistical family:
A logistical family consists of different commodities having similar logistical requirements.

Logit model:
The logit model in its basic form states:

$P_{in} = e^{U_{in}} / \Sigma_j e^{U_{jn}}$

where P_{in} is the probability for individual n makes the choice i. U_{in} is the utility for individual n of the choice i.

Multinominal logit (MNL) model:
These are models for data in which the response is often a set of choices and is therefore measured on a nominal scale. At least some of the independent variables in these models indicate characteristics of the choices (cost, size, attractiveness, etc.) instead of characteristics of the subject or chooser (location, production type, etc.).

Multinominal nested logit model:
A nested model comprises choices in more levels, e.g. land based against sea based transport at level 1, and a choice between different land based modes at level 2.

Probit model:

The probit model is similar to the logit model except the probability is evaluated based on a standard normal distribution in stead of the exponential distribution.

Random utility:

The utility of a certain choice comprises a deterministic utility (Va) and a random element (e_a). The joint utility is referred to as the random utility.

Transport units:

In the ECMT terminology, 'transport unit' comprises both transport equipment and the means of transport or rolling stock (vehicle/lorry, vessel/ship, train etc.). In the present paper, the term “transport unit” is mainly used to indicate the rolling stock. In the cases where transport equipment such as containers, swap bodies, semi-trailers etc. are addressed, the terms 'loading unit' or 'unit loads' are being used.

Utility function:

The utility function V_{in} is the sum of the different characteristics weighted by their respective coefficients. V_{in} is then expressed as:

$$V_{in} = \Sigma_k \beta_{ik} X_{ikn}$$

where X_{ikn} represents the value of the characteristic k for individual n and the choice option i, and β_{ik} is the coefficient associated with this characteristic.

3 MODEL TYPES AND THEIR APPLICATIONS

The *first section* of this chapter gives a short introduction to the long-term and short-term perspectives of modelling, i.e. in the form of strategic models and operational models respectively.

The *second section* discusses the relations between transport policy decisions and the application of models.

The *third section* presents a number of strategic models that are used in practice, and describes the main components of these models.

The *fourth section* gives a short reference to short-term operational models.

The final section outlines the modelling requirements for a national goods transport model.

3.1 MODELS IN A LONG-TERM AND SHORT-TERM PERSPECTIVE

Transport models are developed in order to facilitate decisions about major public and private investments, whether these are aimed at physical infrastructure projects, distribution centres or commercial fleets. The models describe the real systems in a formalised and simplified way. The models should, however, provide a picture in which relevant parameters can be varied and realistic results deducted. If this requirement is adequately met, the models will be able to assist in providing a better answer to questions like:

- How does the goods transport pattern develop overall or within specific sectors?
- How will the improved infrastructure affect the goods transport pattern?
- What are the environmental effects of goods transport in rural and urban areas?
- How is it possible for transport operators to utilise expensive equipment in an optimal way?

Freight transport models can be applied in a long-term context, i.e. the so-called *strategic models*, or in a short- to medium-term context, i.e. the so-called tactical or *operational models*.

The strategic models should be able to depict how national transport policy measures and public investment initiatives will affect the long-term development of transport demand, mobility, traffic pattern and travel behaviour. This is illustrated by Figure 3.1.

The applications of strategic transport models also fit into a wider scope, including the longer-term impact of policy measures and infrastructure investments on the location of economic activities and on regional development.

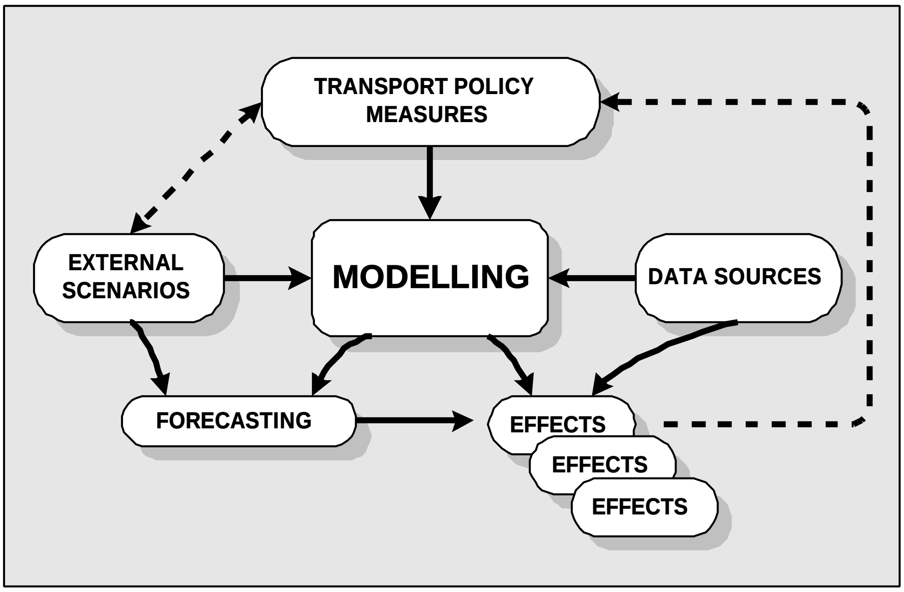


Figure 3.1: Relations between modelling, statistics, external developments and transport policy measures

The operational or tactical planning models can be used for similar analysis in the short term. Typical short-term issues to be analysed are the change of cost structures, change of transport market e.g. the organisation of transport, and weight and dimensions of vehicles.

The modelling requirements for the strategic and operational models respectively are different.

3.2 IDENTIFICATION OF DECISION PROBLEMS

Decision-makers both within the public and the private sector frequently deal with international, national and regional goods transport.

3.2.1 National transport policy questions

In line with EU recommendations, Denmark has adopted a policy of transferring goods from road to rail and sea transport or combined transport.

In relation to the policy formulation concerning increased use of intermodal transport solutions and a change of the modal distribution from road to rail and sea, the Danish government in 1993 presented a transport policy declaration document, 'Trafik 2005'. The declaration is based on the following five main strategies, and goods transport is an element of all these strategies:

- Affecting the volume and distribution of transport and traffic
- Strengthening of alternatives to individual car transport
- Reduction of the environmental problems related to traffic
- Re-orientation of the traffic sector investments
- Improved planning and research in the fields of traffic and transport.

There is an obvious need to monitor the effects and effectiveness of the policy measures. This requires the availability of methodologies, which are able to forecast the effects of the said measures, but there is also a need for databases and information systems, which encompass the relevant transport policy data items to enable an assessment of the actual development.

EU-funded research has led to the recommendation that a future European transport policy information system ('ETIS') should:

- provide a better understanding of the causes and impact of transport, and improve the ability of monitoring the development of various aspects of transport and mobility;
- provide inputs to the formulation of policy scenarios;
- facilitate the development and application of adequate assessment methodologies in relation to policy measures and alternative strategies.

It appears relevant to apply the ETIS principles at the Danish national level as well, and there is a need for transport modelling as an intermediary tool in all of the three aims mentioned above.

3.2.2 Policy issues and modelling requirements

Questions related to the national level require models that are developed for the national level and encompass variables related to the national level. As an example, the national government may want to investigate different policies on truck user costs. Therefore the established models should be able to reflect the impact on prices, on modal split and on route choice for different ways of charging road user costs.

This implies a need for strategic models, which are able to indicate long-term development trends, but which are also able to illustrate the reaction on policy variables of different types.

Policy initiatives concerning how to transfer goods from road to combined road/rail and maritime transport requires a more precise description of the possible potential for such a transfer. There is a major need to analyse more specifically which types of goods that have potentials for being transferred. There is also a need to know more about the factors influencing the choice of transport mode and about the awareness and importance of inter-modal transport.

The goods transport sector, however, is a sector of rapid changes that are triggered by increasing market demands for complex and cost-effective logistical solutions.

Commercial companies do not only consider transport as the immediate costs of moving the goods from one place to another. Rather they are viewing the transport process within a logistic concept where capital requirements related to easy and fast market access might be an even more important economic parameter than the direct transport costs. Companies are also aware of the importance of response times and reliable delivery terms, in order for transport operators to stay in the market.

The importance of flows of goods with a very short lifetime is growing. This is not only the case for e.g. vegetables and fresh fruit. Short lifetime is an increasingly important characteristic for consumer goods, which are either related to the accelerated development in e.g. microelectronics, computers and software products, or related to fast changes of certain commodities e.g. fashion. As a consequence, transport is more and more often being considered as an integrated part of the production process, ensuring that bits and pieces are assembled through the use of logistic chains. There is an accurately calculated trade-off between the value of the goods and the speed by which they are delivered. Therefore freight transport by air is presently growing very quickly, as this mode is able to ensure that high-value commodities reach the consumers before the value is obsolete.

Evidently this logistic requirement is not relevant to all types of goods. Bulk products like crude oil, iron ore and coal are commodities that are all kept in storage for relatively long periods. In this case the direct transport costs are of major importance for the choice of mode and transport route.

Another development trend is that the distribution is being centralised. Major warehouses that are centrally located in Europe are able to serve almost all of the European countries within 24 hours. In Denmark a similar development is taking place, sparked by the opening in 1997 and 1998 of the fixed link across the Great Belt.

Development of the type of logistic management or supply chain management, which serves a widespread need for small and frequent deliveries of high-value goods over short distances, has changed the transport scene during the latest years.

The issues mentioned above are being integrated in the latest freight transport modelling, e.g. the Dutch SMILE project and the SCENES model. However, the lack of data describing the existing situation leads to problems of validating the mathematical models.

Having established the needs in terms of developing logistics, the next point of focus is the transport chains and the possibility of modelling these chains. In the wake of the growing needs to reduce the congestion on particular sections of the European motorway network, an increasing interest is vested in the development of inter-modal transport chains. Hitherto the main effort has been devoted to increasing the competitiveness of the railway system. However, the main problem of inter-modal transport is the necessity to use a number of load transfer points. These transfers are time consuming and rather costly seen in the context of the total transport costs of the consignments. The development of inter-modal transport solutions has also been hampered by high costs of the required feeder and distribution transport that are usually carried out by trucks.

In principle, the composition of inter-modal transport chains could be modelled in mathematical terms. However, the way the chains are composed may be influenced mainly by cost structures and rebates that are not transparent or identifiable. Furthermore, the data requirements in order to describe inter-modal transport chains are quite considerable, and it is doubtful whether the present statistics would be able to give a reasonable indication of the use of inter-modal transport.

Transport chains are also in existence for the road transport mode. The design and function of distribution lorries or vans are quite different in comparison with the types of trucks used for long-distance haulage. Seen in the context of the use of different transport modes this aspect is not significant. However, if the main problem is related to city logistics and freight transport in urban areas, it is important to be able to consider the different ways of managing the transport flows from factory outlet to reception at the retail store. Again, these aspects have not been included in existing transport models, primarily because of lack of data on the issues.

There are a range of modelling issues that relate to different data and information aspects. Transport quality is such an issue. Transport quality cannot be seen in the context of general statistics but should be analysed and subsequently included in modelling based on individual transport data. However, there have been major difficulties in obtaining sufficient data for including such aspects in transport modelling in a meaningful way.

3.2.3 Investments in infrastructure

One of the strategic measures in the “Trafik 2005” has been a reorientation of the infrastructure investments. Several modelling issues are related to infrastructure decision making.

Roads:

The major part of the Danish road network is constructed and maintained by regional or local authorities. The Danish state road network totals about 1600 km only, and it consists mainly of the national motorway system supplemented with a few important roads in Jutland.

Future maintenance of the Danish trunk road network would depend on the volume and route choice of freight transport. There is a specific need to identify the utilisation of trucks, their weights and the related axle loads, and the road sections that are particularly used by heavy freight transport. Furthermore, in order to establish adequate data and indicators about the future freight transport demand, a detailed knowledge is required about the present goods transport flow patterns between regions and zones (e.g. at the county level) in Denmark.

Therefore there is a need for national and regional models, which are able to describe the distribution of truck traffic on the network, the driving pattern and tonnes transported. On this basis the wear or physical deterioration of the road pavement due to the traffic of heavy trucks can be compared with the estimated lifetime of the road. Based on these deterioration models, the need for and type of maintenance can be assessed (ref. pavement management systems) and it is possible to prioritise the maintenance.

Proposals for the extension of the road network are being evaluated in terms of the impact on traffic, impact on environment and impact on safety. Traffic models are developed in order to establish the traffic consequences. However, Denmark has not established a recurrent and long-term infrastructure planning procedure. Therefore modelling needs, related to the development of new physical infrastructure or the upgrading of existing infrastructure, are catered for ad hoc.

Rail:

The development of railway infrastructure has always been a topic for major policy debate. However, as deregulation and privatisation of the national railways are progressing in line with the separation of infrastructure and operation of services, the need for models describing the consequences for rail operations under different assumptions of infrastructure development is imminent. Such models are being developed by the Danish Railway Administration ('Banestyrelsen') under the Ministry of Transport.

The use of railway transport is depending on the availability of combined terminals for both rail and trucks. The optimal location of such terminals can be decided by using location models. The planning and ex-ante evaluation of new terminals requires detailed models that are able to adequately define the use of the terminal as a transfer point. The strategic model 'SAMGODS', which is used in Sweden, enables to a certain extent the analysis of location and use of such transfer points.

The fixed links:

During the recent years, fixed links have been constructed and opened for operations across the Great Belt between the western and eastern part of

Denmark and across the Øresund between Copenhagen and Malmø in Scania, Sweden. The policy background for these two links is comprehensive. However, one of the reasons is that fixed links would provide the goods rail services with a competitive advantage compared with the truck services. As a basis for the decision concerning the construction of the fixed links, some rather coarse and simplified models have been established. These models describe in broad terms the overall consequences of establishing the fixed links, including economic evaluation and environmental impacts. In the wake of the decision of establishing the fixed links, considerably more detailed and complex models have been developed. The main purpose of the latter models has been to estimate the expected income generated by the fixed links. However, the models also serve as valuable tools for the evaluation in more detail of the consequences in terms of the use of adjacent land infrastructure, competing ferry lines and other issues.

The modelling development and application related to the fixed links have been based on state-of-the-art models. This work has included the collection of a range of data describing the overall traffic as well as details of individual shipments carried out by train or truck.

In connection with the ex-ante assessments of a possible fixed link across the Fehmarn Belt, major efforts have been devoted to estimating the future traffic that would use the fixed link. In this context there has been a need for developing methods, which are able to identify the factors determining the volume of transport, the modal choice and the choice of transport route.

3.2.4 Freight Transport Operations

Already for a long time the Danish State Railways (DSB, Since June 2001 Railion) company has been developing and applying strategic planning models for goods transport. The first models were developed about 30 years ago with a focus on important transport segments, like wood from Sweden, steel plates and pipes, machines and mixed goods. This modelling work has been continued, but presently the focus is on unit loads that represent an increasing share of the European rail traffic.

The reason for working with this type of models is the increasing need to justify the relation between volumes of transport produced and income generated to the railway operator. The operational policy is increasingly focused on transport services that generate surplus. Therefore operations are planned by the use of operational analysis. This enables an evaluation of the potential volume throughput at different terminals and the costs of operations by using complex mathematical models, in order to create optimal timetables of operations.

Similar models are applied by the trucking companies in order to optimise the operations, including the collection and distribution of goods, optimal use of drivers and ferry routes, etc.

3.3 STRATEGIC MODELS

A preliminary grouping of the areas of strategic transport model applications could be as follows:

- Effects of national transport policy measures on long-term traffic pattern and modal distribution
- Effects of specific transport infrastructure investment projects on traffic pattern and modal distribution, and the socio-economic and environmental impact of the project
- Assessment of transport network development plans.

The model designers are offered a huge toolbox, which continuously increases in size. Mathematical algorithms, statistics and computation power are under continuous development. The presentation and forecasting of general freight transport flows is an area where various standard software tools are widely used, in order to present the corresponding freight transport flow patterns. This includes the simulation, at the various geographical levels, of different development scenarios. The variables of the scenarios could reflect trade patterns, transport infrastructure supply, costs and service quality levels, and transport policy and regulatory framework etc.

In the EU 4th and 5th Framework Programmes for R&D (FP4 and FP5), goods transport modelling has been included as an important research topic, and both addressing overall and specific problem issues. Some of the experiences, which can be deducted from this research concerning models, will be referred to below.

As mentioned above, the fixed link investments in Denmark have required the development of complex and comprehensive transport models for goods transport. This work has been carried out by internationally known model designers and has led to further knowledge about modelling tools. However, the models and their use have also identified problems both in the context of modelling issues and particularly in the context of data availability.

In Denmark, goods transport models have been developed for specific projects. In Sweden to the contrary, a national goods transport model, SAMGODS, has been developed as part of the Swedish ten year planning procedure ('Inriktningsplanering'). The SAMGODS model is a state-of-the-art modelling system comprising a number of different model types. Irrespective of the advanced nature of the system, some shortcomings are inherent, and these will be mentioned later on in this chapter.

The model designers are dependent on the availability of data for estimation and calibration of the models. It is often needed to establish these data before it is possible to carry out a model design.

Usually the strategic freight transport model will comprise a number of steps or sub-models. The following six steps are considered in the most recent models:

- Assessment of production and consumption per area (zone)
- Distribution of freight flows between production and consumption zones
- Assessment of logistical requirements
- Modal choice for freight flows between zones
- Conversion of freight flows per mode to vehicles
- Assignment of vehicles to networks.

In this report the term *aggregate models* is used to characterise models based on the analysis of aggregated data, that are sets of singular consignments grouped according to geography, goods type, transport mode, etc. Such data are typically available from the statistical offices in the different countries or in EUROSTAT. The aggregated models usually include sub-models for the description of production and consumption, spatial distribution, modal split, and assignment.

The term *disaggregate models* is used to identify models based on the analysis of each singular consignment or shipment.

In the following sub-section, some examples of the application of aggregate and disaggregate models will be presented and discussed.

3.3.1 Examples of model applications

The major models presented here are used as background knowledge and references for the discussions of different data and modelling topics in this report.

Freight transport model for Scandinavian Link, developed by the Danish company Hoff & Overgaard:

The Scandinavian link model is based on fixed demand matrices for Scandinavian goods flows. The matrices are based on an analysis of the year 1984 and 1988 national trade statistics for Denmark, Sweden, Norway and Finland. Modal split is based on a pivot point logit model for seven different commodity groups and three different transport modes. Route choice is carried out for road and rail transport, including ferry transport, for each of the seven commodity groups. The model was used for preliminary assessments of the Øresund and Fehmarn Belt fixed links, and for an assessment of the Scandinavian Triangle project.

Model for the transfer potential of sea transportl, developed by the Danish company Carl Bro for the Danish Ministry of Transport and the Transport Council:

The model for assessing the transfer potential for sea transport is a modal split model based on RP- and SP-interviews (for further details, see the separate report about discrete choice models). Level of service data are described based on networks in the Scandinavian Link model that are updated to 1994 level. The model distinguish between two commodity groups (i.e. low value and high value goods) excluding bulk. The model was applied on data for Danish import and export in 1992.

Transport Quality Model (TQM) for goods transport, developed by the Danish company TetraPlan for the Danish Transport Council:

TQM is a model for assessing transport quality in terms of travel time and delay. It is based on network data and stochastic distributions concerning the expected and actual arrival times for freight shipments by road and rail. Data for verifying/validating the model was not available, and the model has not been used any further. The principles in the model have been considered in connection with improvement of the Swedish SAMGODS model.

Fehmarn Belt Freight Transport model, developed by the Fehmarn Belt Traffic consortium (FTC) for the Danish and German Ministries of Transport:

The model is based on fixed demand matrices of goods flows between Scandinavian countries and countries to the south of the Baltic Sea. The goods flows are separated in ten different commodity groups and are based on trade and transport statistics. The modal split sub-model is a logit model based on RP- and SP-interviews, and with level of service data related to the 1996 networks established for the model. The route choice sub-model is a logit model that distributes the road and rail traffic between the ferry routes crossing the Baltic Sea. Route choice is based on RP- and SP-interviews as well. The model has been used for assessing the potentials of different proposals for fixed links across the Fehmarn Belt.

The NEAC model, developed by the Dutch company NEA:

This model has been used as a traffic forecasting tool on the 'TEN Pan-European Corridors' of Helsinki and in the '2020 European' transport projects. The model includes a production and consumption model in which regional production and consumption (see Section 3.3.2 below) is assessed for all European countries. The Demand for goods transport is distributed between regions based on a gravity type of model including generalised costs between the regions. Modal split is based on a model in which lead times and transport costs related to different transport modes are the main variables for assessing the modal split. Route choice is carried out for road transport.

The German BVWP freight transport model, developed by the German consultant BVU for the German Ministry of Transport:

The BVWP model is based on the assessment of production and consumption in the different regions in Germany in terms of relationship between tonnes produced and economic indicators. The production and consumption expressed in tonnes are distributed between zones based on a gravity type of model in which generalised cost is an important variable. Modal split is carried out for consignments above 2 tonnes and is based on a comprehensive RP- and SP-analysis. A logit model is established that comprises five different types of goods and nine different transport modes. As an example, the following variables were found to be significant for the modal choice of inter-modal transport:

Piggyback/Swap body	**Road -railer**	**Container**
Price	Price	Price
Lead time	Lead time	Lead time
Terminal opening hours	Guarantee of delivery	Terminal opening hours
Guarantee of delivery	Reservation of room in train	Guarantee of delivery
Guarded terminal	Damage risk	Guarded terminal

Route choice is based on advanced route choice algorithms for both road and rail transport. The route choice sub-model combines choice of route and conversion to vehicles/wagons.

SAMGODS, developed by a number of Swedish universities and consultants for a Swedish purchaser group represented by SIKA (The Swedish Institute for Transport and Communication Analysis):

SAMGODS is a comprehensive national goods transport model for Sweden. Transport demand is based on regional-level input-output data for Sweden, and trade data in economic terms distributed between regions in Sweden and regions abroad. The resulting goods flows matrices are based on value/weight relations for 15 different commodity groups and distributed between municipalities in Sweden. The matrices are not sensitive to changes in transport supply. The modal split and route choice is carried out for 12 different product groups by means of the STAN model. The product groups are based on an analysis of more than 1000 different commodities and their possible logistical requirements. The assigned tonnes per link are converted to vehicles/wagons in order to provide the related traffic loads on the road and the rail network.

The STEMM model, developed under the EU research project 'Strategic European Multimodal Modelling' ('STEMM'). The Norwegian NEMO model is part of the STEMM model:

The STEMM model is developed primarily with the aim to investigate modal split and route choice. Demand is therefore assessed based on fixed matrices covering Europe. The matrices are based on a breakdown of the existing statistical data on trade and transport. Modal split and route choice are based on two different models, i.e. the MDST model for analysing cross-Channel and Trans-Alpine goods flows and the STAN model used for analysing the Scandinavian Link and the North Sea corridors. The two models are quite similar and based on the cost and quality assessment of different transport modes or combinations hereof. The STEMM model discusses and redefines the commodity groups according to logistical families.

The SCENES model, developed by the ME&P company (UK) under the EU research project SCENES about European transport scenarios:

This model is the only one of the above mentioned models in which a specific logistical module has been developed. The logistical module assesses different combinations of transport modes, transport routes and distribution centres. The module establishes choice probabilities for the selecting of one combination instead of another combination depending on the logistical requirements for the different commodities. The SCENES model is most likely one of the state-of-the-art models in Europe today. However, the development of modelling is fast and the increase in calculation capacity offered by the computer industry makes it possible to further expand the modelling capabilities.

The SMILE project, developed by TNO-INRO for the Dutch government

The Dutch SMILE project has elaborated on a monitoring system applicable for the Netherlands. This project is being implemented as part of the current Dutch National Transport Plan (National Verkeers- en Vervoersplan, NVVP). The major shift in policy, reflected in this plan, is focusing on relieving the effects of increased mobility and ways of satisfying the mobility in a more sustainable way. The underlying model in the SMILE project is similar to the SCENES model. A great effort has been carried out to establish a well-documented logistical module.

3.3.1 Production and consumption models

In this case the analysis is carried out on total production in tonnes of a certain commodity type in a certain area (country or region). The model describes the total production as a function of one or more independent variables describing the same area. As an example, the German goods transport model describes the production of agricultural produce in an area, expressed in millions of tonnes, as a function of the production of cereals and fruit in billion DM.

Almost in all European models the total production and consumption of goods in different regions are determined based on aggregate models. All the models are based on the analysis of transport data, demographic data, economic data, and production data by region or country.

The SAMGODS model system describes the production and consumption based on regional-level input-output tables for Sweden. It further includes a model for employment development within the different sectors. The production and consumption model is based solely on economic indicators, and does not contain any goods flows at all. The model has been developed mainly to produce projections of Swedish economic development in a medium-term perspective (5-10 years).

The SAMGODS production and consumption model serves as input for:

- A model for interregional transport demand within Sweden

- A model for regional forecasting and a regional breakdown of Swedish foreign trade.

The mathematical equations used in the production and consumption models are based on different types of regression analysis. However, the background data will usually show some major variations, which cannot be reflected in the model. These aggregate models are, nevertheless, widely used as a tool to define the general framework for total production in the goods transport models.

As indicated above, the independent variables in the production and consumption models are in general economic variables. The resulting production and consumption could be presented in tonne figures as in the case of the German model, or it could be economic figures as in the case of the Swedish model. Irrespective of the unit applied, a translation from economic terms to volume (weight) terms has to take place. In the German model the transformation is done endogenously in the model system. In the Swedish model a specific weight/value model has been developed in order to transform the economic flows per commodity group to volume flows indicated in tonnes.

A major distinction should be made between 1) bulk of predominantly low-value and 2) other types of cargo such as general cargo and cargo organised in 'loading units' or as 'unit loads. Bulk is often served by separate transport systems. This is particularly the case for railway and maritime transport. Here bulk and unit loads respectively are served by different types of rolling stock and vessels, and by special and separate terminals. No functional or physical inter-linkages exist between the two systems, and as a consequence, it would not make any sense to pool together in a common model the two systems and the two types of goods.

3.3.2 Distribution models

These models describe the distribution of freight transport flows between areas. Distribution models distribute goods production from one region to all other regions based on the consumption of each area and the distance in terms of costs or time between the areas. The use of distribution models should be considered carefully. As an example, consumer goods would most likely be distributed based on consumption expressed in terms of population and distance as specified above. On the other hand, transport of coal between a coal quarry and an electricity plant would be modelled in a wrong way if a similar distribution model were applied. The model may neglect the fact that the electricity plant has been located close to the mine because of the coal resources available. In such a situation it should be considered to apply a *heuristic model*, describing the de facto situation.

A major consideration is the amount of data required for such a description, which increases fast when the zoning system and the commodity classification become more detailed.

In Sweden a foreign trade model system has been developed, which is able to forecast trade (export and import) between the 286 Swedish municipal areas and about 180 foreign trade regions divided into commodity aggregates. The model system has two main subsystems, namely a bilateral trade model subsystem and a subsystem modelling lower-level regional trade flows. The model system comprises some 50 bilateral trade models for suitable trade areas (countries or groups of countries) and commodity aggregates. These models are of the gravitation type using interalia GNP, population, and distance as independent variables.

The SAMGODS model distributes the interregional transport demand within Sweden. As input data the model uses the output from the production and consumption model, and output from the foreign trade model system transformed to tonnes using the value/weight model. The model utilises an *entropy algorithm* to estimate forecast demand matrices for the relevant commodity groups. Estimates of the domestic interregional transport flow matrices for 1998 are based on available empirical information as well as the corresponding foreign trade matrices. The latter matrices are established based on available surveys of road, rail, sea and air transport, as well as on sources describing the import to and export from Sweden.

3.3.3 Modelling of logistical requirements

The development of transport solutions related to the general development of production has lead to an increased demand for modelling, which is able to include logistical systems. Examples on this particular issue are the EU research projects REDEFINE and SCENES, and the Dutch SMILE project. Definition of commodity groups in relation to logistical families or clusters are also used in the SAMGODS and STEMM models.

It is evident that different types of commodities require different logistical solutions. As an example, fresh fruit will demand a different type of transport solution compared with canned fruit, and the primeur wine from Bordeaux is transported by air and taxi in order to reach the customers the same night it has been released, whereas usual wine is transported in tanks.

The aim of REDEFINE (Relations between demand for freight transport and industrial effects, see Chapter 5) was to study truck transport and logistics in Europe. The project did not produce any assignment, but discussed the logistical requirements. The commodity classification was based on the same 542 commodity groups as used in the Dutch project SMILE (Strategic Model for Integrated Logistics and Evaluation). These were aggregated in 50 so-called *logistical families.* The definition of the logistical families were based on the following seven characteristics:

1. Bulk or general cargo
2. Density (kg/m3)

3. Packing density (packs/unit/m3)
4. Value (NLG/kg)
5. Use of distribution centre (yes or no)
6. Consignment size (small, medium, large)
7. Value of time (lead time) (low, medium, high).

In the REDEFINE project the logistical families have been aggregated to logistical clusters as indicated below:

Logistical cluster	Type of goods	Density (ton/m3)	Value (euro/m3)
1	Bulk	High	Low
2	Bulk	Low	Low
3	General cargo	High	High
4	General cargo	High	Low
5	General cargo	Low	High
6	General cargo	Low	Low

High density = >1tonne/m3 High Value = >5 euro/m3
Low density = <1 tonne/m3 Low value = < 5 euro/m3

Table 3.1: Logistical clusters and their characteristics

The REDEFINE project argues that the above mentioned characteristics, i.e. Nos. 1, 2 and 4, are sufficient to describe also the other characteristics i.e. Nos. 3, 5, 6 and 7 as indicated above, because the latter are seen as depending on the type of goods, density and value.

3.3.4 Modal split

For a number of years modal split for goods transport has been modelled by analysing how the total volumes of goods by road, rail and ship have been developing related to the supply characteristics of the transport modes. During the late 1970s and the 1980s, models were constructed based on singular observations. When a sufficient number of observations had been collected, the modal choice was related to supply characteristics, and a simple modal choice model was constructed. In 1984, NEA and Inrets (a French transport research organisation) developed the first transport model covering the whole of Europe and based on this type of analysis. For each of 16 different commodity groups, competition between the transport modes were established based on a function comprising distance and consignment size. Small consignments were mainly related to the use of truck, whereas large consignments primarily were related to rail and ship.

Modal choice models are increasingly being based on the analysis of choices related to a number of individual transport activities. Each observation concerns a specific consignment and describes for example the transport type, commodity group, transport mode and supply characteristics (price, transport time, possible other supply variables), i.e. revealed preference (RP) data. Based on a statistical analysis of the observations it is possible to estimate model parameters in a modal choice model based on the principles in utility theory, e.g. a *multinominal logit model* or similar types of models. The established model is able to forecast how each observation will change under different assumptions concerning the supply characteristics or some of the other independent variables. This type of model is most often used in the context of mode and route choice. It has been widely used both in the major Danish transport models and similar European models.

The observations of present transport patterns and characteristics are often supplemented with interviews of buyers of transport services ('shippers') and transport companies (transport service providers or 'carriers'). In these interviews they are requested to respond to different hypothetical supply characteristics and the related choices of route or mode. The results are stated preference (SP) data. In this way it is possible to obtain other choices than the observed, and it is possible to establish independent relations between variables. As an example, the comparison of quantitative data concerning transport time and transport price will often show a high degree of correlation. This needs not be the case for data based on the hypothetical questions of a SP-analysis. The models developed in Denmark and in a number of other countries in Europe are based on application of the 'ALOGIT' software that is a programme for estimating discrete choice models. ALOGIT has been developed by the Dutch company Hague Consulting Group. Other companies have developed similar software packages for estimating discrete choice models. The particular aspects of discrete choice models are further elaborated in the separate report on the methodology and data of discrete choice models.

In recently developed transport models, modal split and route choice are often interrelated. The reason is that goods transport over the last two decades has developed a number of characteristics, which makes the use of single mode transport less frequent, and the use of transport chains including a number of modal shifts quite common. The transport chains are particularly relevant for unit loads like containers, swap bodies and trailers.

The Swedish SAMGODS model uses the STAN model software as the mode and route choice model. The STAN model is able to allocate a certain type of commodity flow (product) to different transport modes and different routes according to the costs of the mode and route. The cost expressions have been evaluated based on the analysis of cost structures and based on calibrations. Cost functions are available for different types of nodes and links and for different commodity types.

The commodity types used in the present STAN application (STAN99) are to a certain extent similar to the logistical clusters mentioned for the REDEFINE project (ref. Table 3.1). A conversion is carried out between the commodity groups used in the distribution model and the commodity groups. The cost functions have been developed for the logistical clusters that are based in principle on RP and SP data.

3.3.5 Vehicle models

The vehicle models convert goods transports by road or rail expressed in tonnes to number of vehicles or wagons. The conversion of tonnes to vehicles may take place at different positions in the modelling set-up, and some of the freight transport models are developed based on vehicles.

The conversion of goods transport in tonnes to vehicles assumes knowledge about the percentage of empty vehicles and the capacity utilisation. In the German goods transport model and in the Fehmarn Belt model the number of empty vehicles is established as a function of the goods balance between two regions. If the volume of goods to be transported in one direction is equal to the volume of goods to be transported in the other direction, the number of empty vehicles is at the lowest but it is not zero because of imperfections in the market.

3.3.6 Route choice and assignment models

Assignment is one of the mathematical methods, which is currently being developed due to improved theory of mathematical algorithm and increase in computational power. For a long time the assignment principle of 'all or nothing along the best route' has been applied in goods transport. However, the theory has developed, which means that today assignment is based on:

- Stochastic choices, i.e. a route choice in which a stochastic element is included. The SUE algorithm developed by the Technical University of Denmark is an example of an efficient stochastic algorithm, which has been used for describing the lorry traffic in the Copenhagen metropolitan region. Assignment is based on the travel time and cost, and evaluation of travel time is based on the application of RP- and SP-analysis.
- Multiple routes, which means assignment of traffic to a number of routes of almost equal importance. An example is the STAN software package, which is widely used in goods transport models in Sweden, Norway and Finland. Assignment is based on travel time, cost and quality of transport. Evaluation of travel time and cost of quality is established based on the application of RP- and SP-analysis.

Models for the assignment of long-distance international trips between Scandinavia and the Continent should consider factors like driving and resting time, ferry routes and the possibility of unaccompanied trailers. The STAN model takes these factors into consideration through generalised costs, which also makes STAN able to shift between different transport modes in the assignment. STAN assigns tonnes to the different legs in the networks, and thus carry out both modal split and assignment in the same procedure. After assignment, STAN converts the assigned tonnes to vehicles or wagons.

As indicated there is a broad experience on the use of the STAN model in Scandinavian countries. It has been used in several EU research projects under the 4th Framework Programme (FP4) with the participation of Finnish, Swedish and Norwegian partners.

As another example, a consultant consortium of TFK-Hamburg and Temaplan ('Transnord') has applied STAN as a tool for simulating the road transport flows between the Stockholm region and Eastern Europe via alternative routes across the Baltic Sea and through Southern Sweden. The consortium has developed an extensive database that includes international trade and freight transport flows, networks for trucks, railways, ship and air transport; and transfer nodes between different modes (through terminals and ports).

The STAN model software appears to be a potentially useful tool for the modelling and presentation of freight transport flows at the regional and interregional level. Its limitations may occur mainly in urban areas. The assignment algorithm is not as far advanced as the assignment algorithms used in Danish transport models like the Copenhagen Transport Model and the Copenhagen-Ringsted model.

The main drawback of the application of STAN in the SAMGODS model complex is related to the multitude of cost functions used for calibration of the model. Cost functions consist of two elements, i.e. operational costs and costs of quality. Costs of quality are composed of value of time, costs of delays and frequency costs. Costs of quality are determined for logistical clusters (products) that are very similar to the clusters indicated in Table 3.1. It has proven difficult to determine the costs of quality, even based on sophisticated value of time studies.

The abundant number of cost functions in SAMGODS illustrates the schism between a very accurate description of the existing situation concerning all goods transport aspects as opposed to an efficient tool for analysing specific planning questions.

It should be mentioned that under certain conditions the STAN software apparently provides some odd results. The software developer is aware of this and is currently developing the software in order to resolve the issue.

The assignment of rail transport is based on other determining factors than economic optimisation of route choice. It is a question of timetables, and possibilities for deviations from these timetables if the volumes of goods become too large to being handled within the existing rail infrastructure capacity. These types of models have been developed in Germany. The model is approaching the assignment in a stepwise manner, as follows:

1. The model calculates the traffic load between areas.
2. For each major terminal area the model forms trains in the direction towards the other terminals.
3. The formed trains are assigned to the rail network.
4. Bottlenecks in terms of constrained capacity are identified.
5. Route and timetable are adjusted at locations with constraints.

6. New assignments are carried out, and the process is repeated.

The model provides, as a result, the time lost in relation to the optimal time consumption.

In relation to the Great Belt and Øresund fixed links, assignment models have played a subordinate role. In many cases only a limited number of routes have been included, and the main analysis has concentrated on selecting routes across the Belt or Sound, but not looking at the land-based route choice.

A Norwegian network model:

An example belonging to these application areas is the network model for freight transport to and from Norway that has been recently developed by the Institute of Transport Economics ('Model for transports in Norwegian foreign trade', 2000). The model is based on an existing Norwegian model for freight transport (NEMO) and the STEMM model for international freight transport.

The data used are foreign trade data from 'Statistics Norway' combined with other relevant data, in order to construct the OD-matrices for freight flows.

The Norwegian network model is restricted to the analysis of general cargo. The network of the model is constituted by road links, railway links, sea links, terminals etc. It includes a set of general cost functions that enables it to find solutions that minimise the total costs of a system, given a fixed demand for transport between the zones of the model. In principle the model could analyse the impacts of changes in transport infrastructure and supply, changes in transport demand as well as in transport costs and tariffs.

The model was used to estimate the changes in modal distribution, costs and emissions for six 'scenarios'. The variables used to define the scenarios were the frequency of liner services in Norwegian ports, the degree of concentration of services in a few important ports, capacity restrictions on the continental road network, conveying times and precision in freight transport, and relative increase of costs for road transport. The Norwegian study report stresses that the model is a first version, and that there is still a high degree of uncertainty attached to the model calculations including the cost functions. However, the report also concludes that, despite these uncertainties, network models of this type could be suitable tools to compare different scenarios and roughly evaluating their effects.

3.3.7 Other strategic models

The *growth models* describe the development in terms of a dependent variable, e.g. goods transport between two regions, as a function of one or more independent variables, e.g. growth in production or economic development. Growth models are established based on a statistical analysis of time series.

Growth models are often used as models for strategic decisions, and are often components in more complex model designs.

The Danish Road Directorate ('Vejdirektoratet') has developed a goods transport model that is an example of a simple growth model. In this model, the national economic development by production sector has been used as an explanatory variable. The development of goods transport work is the dependent variable that is indicated in tonne-km by commodity group. The model provides a breakdown into the following commodity groups:

1. Agricultural produce including fertiliser
2. Food stuff and fodder
3. Fuel, liquid and solid, and chemical products
4. Manufactured products, semi-manufactured products etc.
5. Minerals and building materials.

Figure 3.1 shows the actual development and back-casting carried out by the model for the relevant years.

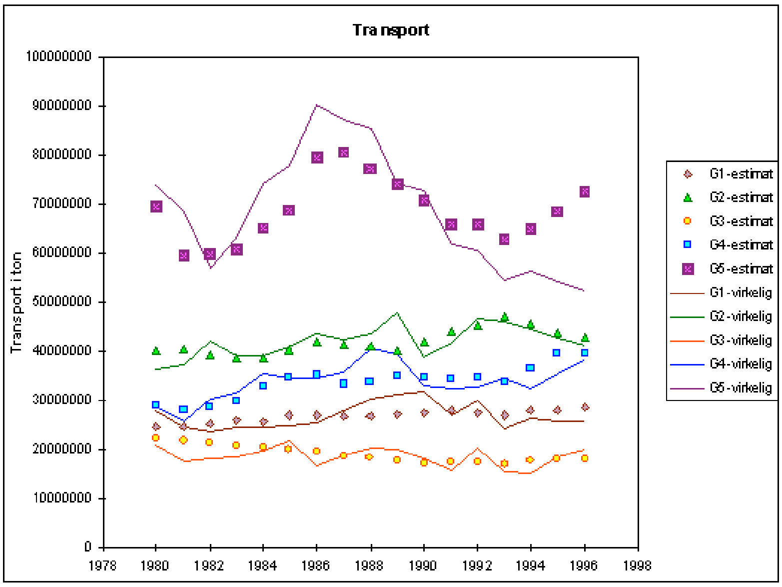


Figure 3.1: Comparison of observations and model results of a simple growth factor model for goods transport by lorry in Denmark
(Source: Vejdirektoratet, Denmark 1998)

Growth models are simple and their application is limited to overall development trends. Development of a growth model requires the availability of time series for relevant data.

3.4 SHORT-TERM OPERATIONAL MODELS

A multitude of data is potentially available for modelling at the operational level. Information can be extracted from forwarding notes and consignment letters. This implies that information is available concerning operational costs for the different transport modes. The availability of data means that models could be developed at the operational level to reflect the specific questions of interest.

Operational modelling is potentially useful for the planning and decision-making by individual shippers and transport operators (carriers). Operational models should be based on operational research as well as on disaggregate behavioural models. Operational modelling may also include urban logistics. In this case there is a particular need to develop new modelling tools and algorithms.

Research has been carried out at the operational level in order for the transport operators to establish routes for the collection and distribution of goods. An aspect, that has become increasingly important, is the current use of the vehicles during operations by means of mobile phones and satellite positioning. Consignments can be picked up during an empty haul and delivered with a minimum of additional driving. Furthermore, instant (online) route planning has become a way to improve customer service and increase the utilisation of the fleet of vehicles.

3.5 CONCEPTUAL FRAMEWORK FOR A NEW GOODS MODEL

3.5.1 Purpose and definitions

The purpose of establishing a goods transport model is to provide a tool that can be used for the analysis of questions with relevance to the goods transport sector. Among these are:

- Effects of changes in production and consumption patterns
- Effects of new physical infrastructure
- Effects on maintenance of the infrastructure
- Effects on changes of the transport equipment and transport regulations including changes of costs.
- Establishment of indicators for goods transport.

Zone system:

The proposed zone system will be based on a subdivision of counties in Denmark, and NUTS 2 zones in the nearest neighbouring countries. The subdivision of counties will be based on municipalities or groups of municipalities. Only the largest municipalities like Copenhagen and Århus are supposed to be subdivided.

Commodity groups:

Commodity groups will have to be defined taking into account economic development, logistical requirements and mode and route choice. At the same time it is a requirement that the commodity groups can be easily formed based on available statistics.

3.5.2 Model framework

Production/consumption model:

The production model should be developed from the available models for economic development, production by sector and import, and the available transport content in each sector.

The EMMA model, which is a sub-model to the ADAM economic input/output model for Denmark (see Section 6.1) with particular emphasis on transport, subdivides the input/output figures for the two ADAM transport sectors to 10 EMMA transport sectors. This provides a more accurate and detailed description of the transport content in DKK in each of the production sectors. The EMMA model has been developed in order to establish the transport energy content in each of the different production sectors. Thus no effort has been carried out in order to assess the transport content in tonnes or tonne-km in each of the sectors.

A model has developed in which the production by sector in DKK is transformed to tonnes based on a value/weight relation (Kveiborg, 2000). The production in tonnes is further elaborated to traffic work by using a handle factor as an important parameter for converting the transport sector output in tonnes to actual traffic work. McKinnon and Woodburn first defined the handle factor in an empirical study of the transport related to the food production sector in UK.

It is important to note that the Kveiborg study does not represent the goods transport in a geographical context. Thus a number of the above-mentioned purposes cannot be fulfilled with the proposed model. Therefore the study can be used only as an overall determinant for relationships between production and consumption sectors, including final consumption, but the study would have to be fitted and adjusted to the economic indicators by sector for each of the counties.

The recently published EU white book on transport sets the focus on the detachment of economic development and goods transport development. It is therefore important to be able to distinguish the transport needs of the different sectors in order to produce some sound relationships between transport and production. The Kveiborg study would be of use in this respect.

Another source for inspiration on this topic is the SCENES project, in which a Regional Economic Model (REM) has been developed based on input-output tables (IOT) for 15 EU-countries. The IOT does not distribute import and export by country. Therefore part of the SCENES project was the development of a methodology to distribute the import and export in monetary terms between countries. The important aspect has been to develop a methodology, which is able to split the import by production factor and final use in terms of national consumption and export.

Another source of inspiration is the Swedish SAMGODS model. This model includes an economic model based on regional-level input-output tables for Sweden and a specific model for import and export by commodity type and regions in Europe.

Distribution of monetary flows between regions:

The national economic flows by commodity group based on an aggregation of Economic Activity categories is broken down by county. The regional breakdown could be accomplished by using regional-level economic key figures, as they are produced e.g. by the Research Institute of counties and municipalities in Denmark (AKF).

Conversion of monetary flows to goods flows in tonnes:

The regional-level goods flows by employment category are transformed to flows in tonnes using value/weight relations. Development of value/weight relations can be based on import/export data. A specific consideration should be given to the development of value/weight relations over time. Attention should also be paid to how goods flow changes when the organisation of production and costs of transport change, and how such changes can be dealt with in the model.

A model for Danish import and export by county:

Import and export data are available from the EXTRASTAT and the INTRASTAT. As indicated in Chapter 6, geographical information concerning the county of origin and destination in Denmark is not available, and would therefore have to be assessed based on modelling. The international end of the trip should be assessed based on other modelling work, e.g. the SCENES model, SAMGODS and the German BVWP model.

A methodology for assessing the international end of goods flows has also been proposed in the modelling work on the potential of sea transport carried out for the Danish Ministry of Transport and the Transport Council. The methodology that is proposed in this study is, however, only applicable to EU countries.

A specific problem is the transit transport passing through Denmark, or with a possibility to pass through Denmark. It is therefore necessary to establish international goods flows between countries surrounding Denmark. Possible sources could be the models mentioned above, and statistics from the different national statistical offices.

Subdivision of flows to the zone level:

The flows based on county to county matrices for the national goods transport in tonnes and for county to international zones for the international goods transport are subdivided to zone level based on a distribution model. The distribution for different commodity types should be based on population and/or number of workplaces within particular employment categories.

Development of logistical module:

The flows have been viewed in light of assumptions made in the SCENES model. Among the factors described, in order to take into account logistic chains, are the configuration of freight flows into logistic families and the definition of flows in relation to their handling requirements.

The SCENES model defines thirteen different freight flows and corresponding handling, as indicated below.

Freight flow	NST/R 2-digits	Handling category
Cereals and agricultural products	00 01 04 05 06 09 17 18	General cargo
Consumer food	01 11 12 13 16	Unitised
Conditioned food	03 14	Unitised
Solid fuels and ore	21 22 23 41 45 46	Solid bulk
Petroleum products	31 33 34	Liquid bulk
Metal products	51 52 53 54 55 56	General cargo
Cement and manufact. build. mat.	64 69	Unitised
Crude building materials	61 62 63 65	Solid bulk
Basic chemicals	81 83	Solid bulk
Fertiliser, plastic and other chem.	71 72 82 84 89	General cargo
Large machinery	91 92 939	General cargo
Small machinery	931	Unitised
Misc. manufactured articles	94 95 96 97 99	Unitised

The SCENES model apply a Spatial Logistics Appended Module (SLAM), which transforms trade flows into transport flows taking into account the logistic costs and bundling possibilities of freight flows. SLAM does not specify the modal choice in a chain, but identifies typical distribution structures for chains based on regional characteristics, products and network.

The developed model redistributes traffic via distribution centres. The model consists of three steps:

- A module for identifying the attractiveness of region for location of a distribution centre.

- A module for calculating the probability of using a specific alternative transport chain based on location score and logistical costs of a specific chain. The costs include transport costs, inventory costs and warehousing costs.
- A module re-assigning volumes of total commodity flow for an OD-relation to chain types.

The outcome of the model would include a more heavy traffic load in the regions where important distribution centres are located. In Denmark this would most likely be the case for the Taastrup Area near Copenhagen and the Triangle area in Jutland.

Data need to be collected in order to establish the probability functions for choosing one type of logistic chain to another. In the Dutch SMILE project more than 500 different commodity types were analysed in terms of logistical requirements.

The development of modal choice and route choice model:

Several models are available for modal choice and route choice. One type of model could be similar to the SCENES model in which main transport modes are defined with reference to the handling categories. Transport flows are distributed between different modes and combinations of different journey stages based on stochastic choice models.

Modal choice in the SCENES model is carried out based on a *multinomial nested logit model* with three choice levels. The first choice is between land modes and other modes, the second choice is between land modes (rail and truck), and the final choice is between heavy goods vehicles and light goods vehicles.

Assignment is carried out in tonnes for all transport modes except truck. Trucks are assigned to the road network as vehicles.

The mode and route choice models resemble the models developed in the Fehmarn Belt Transport Study.

Another possibility is the introduction of the STAN model. STAN is integrated in the national goods models in Norway, Sweden and Finland. The STAN model was further applied in the STEMM project, where data for Denmark was integrated in the STAN model.

Presently the STAN model does not allow for the inclusion of a logistic module of the SLAM type. However, it has been evaluated by STAN users in the Nordic countries that such a module would be required in order to describe in a better way the use of distribution centres. It could be foreseen that the developer of STAN, i.e. INRO in Canada, would develop the model in such a way.

3.5.3 A national model for road transport only

Road transport is the dominant mode of transport in the national freight transport. Many needs are therefore related to a thorough description of freight transport by truck, the distribution of truck sizes, the loading of trucks and the resulting maintenance and strengthening of the road surfaces and strengthening of bridges.

Therefore it could be considered to limit the modelling scope in Denmark to freight transport by road. This would in many ways simplify the modelling task.

4 DATA REQUIREMENTS FOR MODELLING

The *first section* of this chapter presents an overview table of relations between transport policy issues and the type of data required.

The *second section* provides a description of the main types of data required for the modelling of transport systems and their impact.

The *third section* points out some constraints and shortcomings concerning data availability. Data validity is also discussed with reference to some of the modelling development work recently carried out.

4.1 TRANSPORT DATA AND POLICY DECISIONS

EU research has recently been discussing the requirements for data concerning transport policy decisions (MESUDEMO, 2000). In broad terms, the following data items have been identified for different data levels and different policy levels.

Type of data	Requirements for the development of a transport policy database			Identified policy issue
	Causes	Performance	Impacts	
Macroeconomic issues	Explanatory factors of demand	Origin/destination flows	Employment Value added Externalities Spatial impact	Sustainability Quality of transport
Market issues	Industrial/commercial structure Organisation of logistics Value of transport	Vehicle use	Impacts on companies, consumers etc.	Inter-modality Market integration
Transport network issues	Infrastructure investment Time and costs	Traffic on the network	Costs Environmental externalities (emissions, accidents, noise etc.)	Environmental requirements Transport quality Infrastructure policy Traffic safety Interoperability

Figure 4.1: Relations between data types and policy issues

The MESUDEMO research project (see Chapter 5) points out that the challenge is to find a reasonable balance between the needs for detailed data and the cost constrains on producing and maintaining the statistics.

Similar data requirements have been identified in Denmark. However, presently the statistical system does not support a description to the shown level of detail.

4.2 MAIN TYPES OF DATA ABOUT TRANSPORT SYSTEMS

The required data for transport policy decisions may be broken down in specific data categories. The EU research projects 'INFOSTAT' and 'MESUDEMO' suggest a distinction between the following categories of data:

A. Spatial and land-use characteristics of the area of analysis
B. Population, economy and other socio-economic characteristics
C. Transport demand indicators (passengers and goods)
D. Transport network characteristics (nodes and links)
E. Transport services indicators (passengers and goods)
F. Transport impact indicators, including public health, traffic safety, environmental impact, land-use and socio-economic impacts.

With a view to transport modelling, these categories of data can be grouped into:

1) data on external factors influencing transport demand and mobility, i.e. the categories A and B above;
2) data about the transport system, i.e. the categories C, D and E above; and
3) data on the impact of transport, i.e. category F.

These categories and groupings are illustrated in *Figure 4.1*.

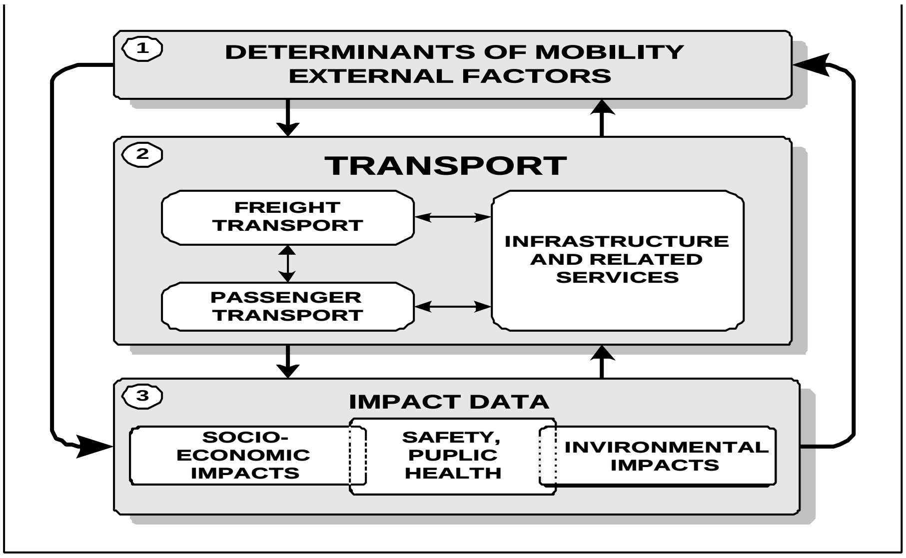


Figure 4.1: Main structure of transport related data

The figure also indicates that transport systems data could be divided into *transport demand data* i.e. data about freight transport flows and passenger transport flows, and *transport supply data.* The latter include data about physical infrastructure and related services, and data about commercial transport services.

4.2.1 Data on external factors

Data on external factors mainly comprises the determinants for mobility. In a number of different studies the main determinants have been assessed as being:

- Economic development by nation and by region
- Development within main economic sectors
- Development in trade
- Demographic development by region or zone
- Land-use by zone.

A number of other factors have an impact on the demand for goods transport, i.e. production processes, logistic chains including consolidation and ways of organising the transports. It is still to be resolved how to organise data collection and databases in a way that makes it possible to collect, analyse and apply these data for modelling.

4.2.2 Data about the transport system

Transport demand:

Transport demand data comprises overall flow and volume data as well as data on individual consignments.

Overall flow data are based on available statistical sources like traffic counts, ferry statistics, trade and transport statistics, etc. There are a number of possible uncertainties that are related to these data. As an example, a peculiarity of the present Danish transport statistics is that the weight of containers is included in the tonne-figures of surface transport but not included in sea transport.

The EU studies carried out about the issue (e.g. INFOSTAT and MESUDEMO) unanimously conclude that the lack of more detailed origin/destination (O/D) information is one of the major shortcomings of available transport statistics.

Strategic modelling of freight transport involves transport demand data for several components of the sequential modelling steps. These data relate to the socio-economic potentials of a particular zone, trade patterns between the zones, indicators of goods flows between the zones, modal distribution of the flows, and assignment to the transport network. The latter component also has to consider the conversion of transport flows to traffic flows.

Some particular problems relate to the establishment of an OD-matrix during the first two sequences, to the modal choice including the consideration of inter-modal chains and transhipments, and to the assignment sub-models.

OD-matrices (e.g. one for each mode) could be established to describe an existing situation, provided the rather ideal situation that exhaustive data are available to describe the relevant transport flows. Such OD-matrices of the existing situation could form the basis for a study of cause-effect relations between socio-economic variables and transport flow variables. Forecasts could be based on combining such relations with socio-economic scenarios that describe alternative future situations.

As a basis for establishing an OD-matrix of freight transport flows, the trade flows between the zones in question constitute the core information.

Figures are not regularly available for domestic trade between regions in Denmark. However, in recent years a lot of research work has been carried out, among others by the AKF and the Institute for Transport Studies in Denmark, to establish a set of hypothetical data describing trade between the Danish counties. The Danish statistical office ('Statistics Denmark') has recently started publicising a regional distribution, at county level, of the national account (in terms of production and income indicators). This statistics could form part of the basis for regional trade flow modelling. A similar system has been in operation in Sweden for a number of years. The latter system forms the basic data for the Swedish SAMGODS production sub-model.

Individual observations on modal choice and route choice are collected based on interviews and observations. The number of interviews to carry out should be determined taking into account the purpose of the model, including its level of detail and on the other hand the available resources for carrying out interviews. Usually, the resources are a considerable constraint as to the level of obtainable detail.

Design of the analyses is an important aspect in obtaining the correct information that can be used for estimating the models. Questions should be clear and unambiguous, and experiences have shown that it is difficult to fulfil this requirement in all aspects. Misunderstood questions and answers may influence the result considerable. However, as pointed out in the separate report on discrete choice models, considerations should also be given to the way individual interviews are carried out and the methods used for analysing the answers.

As illustrated by the above-mentioned examples, it is necessary to continuously keeping an eye on the statistical data that are made available. Are they able to fulfil our needs? Are new types of data required? Are the data sufficiently accurate and consistent to be included in future models with the aim to depict some of the questions raised above?

A major aspect is the overall development of the goods transport sector, and the use of aggregated and strategic models for the analysis of this issue. The data collection procedures still in use have been developed in the past. Thus they reflect some business and logistical conditions that differ completely from present and expected future conditions. In addition, the establishment of the EU internal market has considerably diminished the possibilities of retaining any adequate data on the relations between trade and transport flows. This accentuates the necessity of identifying new concepts and approaches in terms of data collection.

Transport network:

Data about the physical transport infrastructure belong to the transport supply information, and comprise a description of the different types of infrastructure and their use. In transport models this description is simplified in order to focus on the aspects that are of significant importance for the particular model. Therefore in most cases the description and data will not include all elements of the infrastructure. Besides, a number of elements represent average figures (e.g. costs) that cannot be observed in the real world.

Distances in a road and rail network can be defined rather precisely. However, travel speeds will vary according to the different drivers and their way of driving, traffic loads, speed limits and adherence to these, as well as to the organisation of transport units. This implies that the observed individual travel times could deviate considerably as compared to the travel time calculated by the model.

Transport services:

Data on transport services comprise timetables, opening hours of transport centres or terminals, and other services related to and using the infrastructure. These data can be picked up fairly easy, but they constitute massive amounts of data. Moreover, it is not evident how the data should be combined. This issue could relate to a timetable, which is based on a change of transport mode at a certain time. If a delay has occurred it is not clear how such a situation should be treated in the model.

4.2.3 Impact data

Monitoring of impact is an important aspect concerning the improvement of quality of life, reduction of emissions, improvement of safety, and reduction of green house gasses.

Impact data relate to data describing the effects of freight transport. Statistics on the impact of freight transport have to be collected and elaborated in order to establish coherence between freight transport and its impact.

Impact monitoring could be facilitated using indicators for the development of freight transport and the related impact. Such indicators are important because they serve as a basis for decisions about freight transport issues. EU and its Member States have been considering the establishment of general indicators for traffic development and impact monitoring. In Denmark, general indicators concerning the impact of freight transport is not made available, whereas indicators concerning overall freight development are published.

Indicators concerning freight transport development will be discussed in the following section.

4.3 DATA CONSTRAINTS AND SHORTCOMINGS

The major constraints to data availability include the general unavailability or unreliability of data, the differences between national statistical systems, the increasing commercialisation of data sources, and different weight definitions.

4.3.1 Unavailability of data

The necessary basis for establishing an aggregate or disaggregate transport model is detailed knowledge about the actual transport pattern. However, this knowledge is not precise as the data required are either unavailable or uncertain. The unavailability is also related to the commercialisation of data, see below.

A massive amount of operational information is potentially available with the commercial operators. The information of relevance for strategic models is generally more limited, and often the accessibility to these data is restricted by the organisations or operators who control them referring to 'commercial property rights'. Therefore it is necessary to base the strategic models on aggregated data and a limited amount of data concerning individual shipments. It is a question whether the strategic models could be improved if access was granted to operational data.

As an example of the lack of precise or reliable data for strategic modelling, the goods transport between Denmark and Germany on the ferry route between Roedby and Puttgarden is not stated in tonnes. However, the number of trucks transferred by the ferries is known with a degree of statistical uncertainty. The number of trucks varies a great deal depending on the statistical information supplied by Germany and Denmark respectively. No information is available on the origin or destination of the trucks. Therefore this type of information should be determined by interviewing the truck drivers. Such surveys could also determine the share of transit traffic.

On the international ferry routes connecting Denmark with Germany and Sweden, a number of unaccompanied trailers are transferred on the direct routes. Therefore there will be no persons to question about the origin and destination of the trailer. The trailer could contain a number of different consignments to be distributed at different points in Sweden or Germany. However, this information is almost impossible to establish.

Uncertainty or unreliability regarding the basic data has major consequences for the modelling work. It will be problematic to use sensitivity analysis for these data, since they are the basic data for all estimations and calculations. Thus it is important that these data are as representative and reliable as possible.

4.3.2 Different national statistical systems

A number of data items included in the modelling of international transport have to be described based on statistical systems of different countries. These systems may vary considerably from country to country, and hence they produce data with quite different content and preconditions. As an example, international goods transport data in Germany are based on transport statistics, whilst similar figures in Denmark until 1997 were based on trade statistics. Since the two ways of measuring are quite different, it is difficult or almost impossible to compare the data of the two systems.

A problem, which relates to the national statistics, is the change of statistical procedures. Changes are obvious over time, but each change creates a break in the time series collected over a number of years. Furthermore, the EU system imposes other statistical standards on the national statistical systems. This has been the case for example in relation to the introduction of commodity classifications designed for the particular needs of EUROSTAT. Moreover, the change of trade relations has led to similar changes in statistical procedures, reflected most notably with the introduction of the internal market and the abolishment of customs procedures for goods transactions between the EU countries. This has led to a severe break in transport statistics related to intra-EU trade.

4.3.3 Commercialisation of data

Another constraint is the increasing commercialisation of transport data sources, and the lack of common requirements and guidelines for data reporting in Europe. This tendency is particularly reflected by the national railway operators, who are in the transition process of being organisationally and commercially separated from the providers of infrastructure services and being exposed to competitive market conditions.

As a result and for commercial reasons, national railway operators (e.g. the DSB) and private operators (e.g. the UK based 'Arriva' company) would be reluctant to report data about the operations on the Danish rail tracks except at a high level of aggregation. Some general guidelines should be elaborated at the national and EU levels, in order to ensure access to more detailed information and data than the presently publicised statistics, and without jeopardising commercial confidentiality.

Furthermore, there has not been any direct public access to the results of the ad hoc OD-surveys that were carried out in connection with the ex-ante analyses of large-scale infrastructure projects in Denmark such as the fixed links across the Great Belt and the Øresund. Access to these surveys is, however, now possible via the Ministry of Transport.

4.3.4 Net weights and gross weights

A statistical problem, which has been known for a long period of time, is the use of net weight in trade statistics and gross weights in transport statistics. The gross weight may include the weight of container or swap body in truck transport.

Comparison of weight data between Germany and Denmark in the Fehmarn Belt Transport Study disclosed some inexplicable differences related to road, rail and sea transport. As an example, the 1997 goods traffic in Rødby was estimated to 5.4 million tonnes, whilst the same traffic in Puttgarden was estimated to 6.1 million tonnes. No explanations exist to this significant difference except that possibly the general average load per vehicle or wagon is differently estimated in Germany and Denmark.

The Dutch TRANSITIE pilot study has disclosed some similar findings. The study points out that for transit and entrepôt transport, only gross weight figures are available to some extent.

For some types of goods in the EU international trade, weight is no longer recorded. Instead a measuring unit (e.g. litres, number of items, pairs etc.) is defined per individual goods type. This problem has to be overcome by estimating some ratios of weight/value or weight/measurement unit by goods type.

5 TRANSPORT DEMAND DATA STRUCTURES

Based on EU experiences and recommendations, this chapter addresses the data requirements for goods transport modelling.

The *first section* discusses how these data could be grouped and categorised. Furthermore, some potential application areas are pointed out in relation to the use of modelling.

The *second section* provides a short overview of EU research activities of particular relevance to goods transport data and modelling.

The *third section* presents some of the main findings and recommendations derived from the recent EU research projects.

The *fourth section* suggests some new data concepts to be introduced in order to catch the industrial and logistical trends.

5.1 CATEGORIES AND CLASSIFICATION OF DATA

The first part of this section provides some considerations about categories, data sources and the zoning system to be used for statistics and modelling. The second part gives a presentation of the EU classification of trade statistics.

The general categories of information include:

- Trade statistics in value and volume/weight terms;
- Transport flow quantities (mainly net or gross weight indicators);
- Distribution on bulk and various loading unit types (container, swap body, semi-trailer etc.);
- Transhipment and transport chain information;
- Traffic volumes and composition on routes and through nodes and terminals.

The major data source groups are:

- The foreign trade statistics reported through INTRASTAT and EXTRASTAT;
- Data catch of information concerning goods flows, transport units and traffic in terminals, railway stations and sea ports;
- Counting of traffic flows on the road network etc.;

- Periodic surveys, e.g. to provide O/D information about specific transport and traffic flows;
- Operational data extracts.

Any upgrading of the intra-EU trade statistics concerning data on modal distribution has to be resolved at the EU level. However, a solution is not foreseen in the medium term. In the meantime, available input data for modelling has to be supplemented by ad hoc OD-surveys, and by a more systematised collection and sorting of the present statistics on freight traffic and transport flows. In this context, a main focus should be on developing a better linking of transport flow data and foreign trade statistics.

5.1.1 Geographical representation and zoning system (NUTS)

The spatial O/D dimension has to apply a geographical unit concept such as 'area', 'region', 'territory' or 'zone'. Transport modelling uses the 'zone' concept. The co-ordinates of the zone's 'centroid' define the location of the zone. Several types of statistical information can be attached to the centroid of a zone, e.g. figures of population, production of goods, consumption of goods etc. O/D information about transport flow volumes could be linked to explaining variables that are attached to the centroids of the zones.

In principle, the geographical division into the zoning system of a transport model should be based on functional criteria. However, for practical reasons it is feasible to base the zoning on the administrative boundaries and areas of each country, as the national statistics follows these administrative units.

The NUTS zoning system:

The EUROSTAT (EC's statistical office) applies a standardised zoning system, i.e. the NUTS zoning, which is a six-level hierarchical classification. In most cases the NUTS zones are delimited by the existing administrative boundaries applied by each Member State. The NUTS system is specified for several levels of territorial aggregation. At the NUTS 0 level, the zone is equal to the territory of the national state.

For some small countries such as Denmark, the NUTS 1 level zones are not required. As for Denmark, NUTS 2 level zones are not defined although it would be convenient to constitute East and West Denmark as two separate zones at that level. The Danish and Swedish counties are defined as NUTS 3 level zones and the municipalities as NUTS 5 level zones. Sweden is represented also by eight NUTS 2 level zones that are geographical aggregates of 2-5 counties each.

Figure 5.1 shows the zoning at the NUTS 1 and NUTS 3 levels.

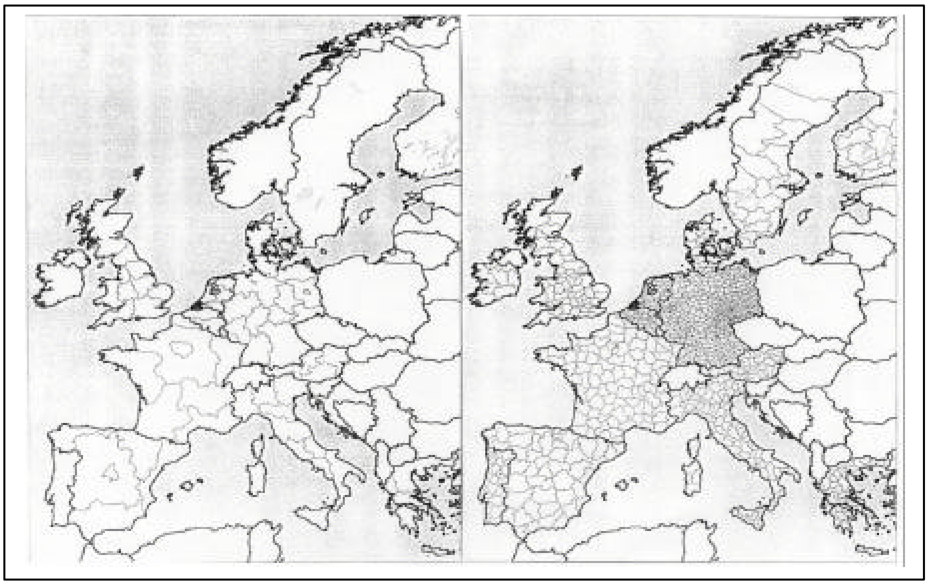


Figure 5.1: The NUTS 1 and NUTS 3 zoning levels
(source: EUROSTAT/NEA)

Due to the differences between the administrative classifications of the various countries in Europe, the size of zones at each level varies considerably from country to country. As an example, the largest NUTS 2 level zone in Spain is larger than the whole of the Netherlands.

In the case of Germany, each zone at the NUTS 1 level covers the territory of one of the 16 states ('Länder'), and the NUTS 2 level zones consist of 40 'Regierungsbezirke'. The NUTS 3 level zones of Germany ('Kreise') and of the Benelux are much smaller than the similar zones of neighbouring EU countries e.g. de Danish counties.

NUTS 3 provides a sufficient level of accuracy and detail in most cases, and several countries have chosen to develop their national transport planning models at the NUTS 3 level. In practice NUTS 2 is considered the most realistic zoning level for statistics and models involving more countries, given the present state of affairs for transport statistics (MESUDEMO, 2000). However, at present complete O/D information is not even available for European goods transport at the NUTS 0 level.

Goods transport statistics compiled by Statistics Denmark are available at least at two geographical levels, i.e. the country as a whole (NUTS 0) and the county level (NUTS 3). Particular data on goods transport are related to a more specific geographical location, e.g. port statistics by each port. In general the national transport statistics can be broken down into the county level, whereas statistics on international transport in most cases cannot be broken down at any geographical level below the country level unless modelling is applied. In the Fehmarn Belt Transport Model transport flows were distributed to municipalities, by using distribution models for each commodity group.

The communication from the European Commission (EC) concerning the collection of road goods transport statistics, COM(97) 443, suggests that future statistical data about national and international goods transport by road will be based on the NUTS 3 geographical zoning. This implies that the Danish counties would be the geographical basis for future transport statistics. Such a zoning level is appropriate and consistent with the socio-economic data presentation that follows the administrative boundaries of the counties.

To the extent that demographic and macroeconomic variables and other external indicators are provided at the Danish county level, this degree of detail suffices for aggregated models at a relatively high geographical level. It is, however, evident that this level of zoning is insufficient for studies concerning particular projects in Denmark. Even the Great Belt transport model, which comprised the whole of Denmark, a zone system had to be fitted to the particular needs of this model. The zone system comprised elements from the NUTS 3 and 5 levels.

As for the zoning of neighbouring EU countries, the NUTS 2 level appears to be sufficient in most cases. However, the request of the EC to carry out data collections at the NUTS 3 level makes it possible to aggregate geographical zones to the geographical detail required in the particular modelling case. The NUTS 3 level zones for Germany constitute, as mentioned previously, much smaller geographical units than the Scandinavian counties that are also defined as NUTS 3 zones, and the former should be appropriately aggregated.

Data collection based on NUTS 3 level fits the EU purpose of establishing goods flows at a level adequate for EU decisions, although as mentioned above the NUTS 3 level seems too detailed in Germany and Benelux. For national projects, the NUTS 3 level would not be sufficient in general. However, the availability of data at this level would provide a considerable better basis compared to what is presently available, particularly in relation to neighbouring countries.

5.1.2 CTSE and NST/R nomenclature for commodities

The establishment of models for goods traffic requires a definition of the commodities included in the modelling work. Commodities traded nationally are almost unknown, but commodities traded with international partners are recorded in detail. This recording is carried out using the so-called Combined Nomenclature (CN) to an 8-digit level, and comprising approx. 10,000 commodities. For all practical purposes the CN 4-digit classification is ample (CN4). The CN4 includes 1,200 commodity types. Most of the usually applied classifications can be formed from the CN4.

For overall planning purpose it is evidently not required to work with the CN4. Other much more aggregated commodity groups form the basis for transport modelling work carried out in Denmark and abroad. The most commonly used classifications are SITC, CTSE and NST/R.

In the recent years a new type of classification based on the logistical requirements of the commodities has been introduced in transport modelling. This classification cuts across all other aggregate classifications used so far.

The 'Commodity Classification for Transport Statistics in Europe' (CTSE) is a common nomenclature established by the United Nations' Economic Commission for Europe (UN-ECE) situated in Geneva. CTSE is based on 20 Categories, 52 Divisions and 169 Positions. These are related to the 'Standard International Trade Classification' (SITC) used by the UN.

The CTSE classification has formed the basis for transport statistics elaborated by Statistics Denmark for a long period of time. Therefore the use of CTSE has been widespread in Denmark. Also in Norway and Sweden the CTSE classification with minor modifications has been used for describing goods transport flows based on trade and transport data.

The "Nomenclature Uniforme des Marchandises pour les Statistiques de Transport" (NST) is a commodity classification for transport statistics used by the EU. Experts from the Member States and the European Commission established the nomenclature in 1961. Since a revision in 1967, it has been indicated as NST/R. The purpose was to set up a uniform goods classification system applicable throughout the EEC, and the nomenclature is based on the CTSE classification, but as indicated below there are some differences.

The Danish national transport statistics based on trade data have been using the CTSE and the SITC nomenclatures. The NST/R classification has also been registered for each dispatch in the statistics, but the NST/R classification was not established with the same precision as the CTSE and SITC goods classification. However, the NST/R classification is used throughout the EU and by EUROSTAT, and the Danish and Swedish systems have been harmonised with this EU system. Thus the NST/R classification is now used on a more widespread scale.

The NST/R classification is described by a 1-, 2-, 3- or 4-digit code. In the 3-digit code it consists of 10 main Chapters, 52 Groups and 176 Positions. The first digit refers to the main chapter, the second digit refers to the group and the third digit refers to the position. On the position level there is a one-to-one relation between the NST/R and the CTSE. Six positions in the CTSE were, however, subdivided in the NST/R.

The CTSE classification and its conversion to the NST/R 2-digit codes are shown in Figure 5.1.

It is seen that the NST/R chapters can be formed from the CTSE categories, except for Nos. 0 and 8 which each contains part of CTSE 10. In the Swedish classification system, CTSE 10 was subdivided along the same lines as indicated by NST/R. The reason is that paper pulp is an important commodity in Sweden.

CTSE	NST/R 2-digits	Description
1	01	Cereals
2	02,03	Potatoes, other fresh or frozen fruits and vegetables
	00,06	Live animals, sugar beet
3	11,12,13 14,16,17	Foodstuff and animal fodder
4	18	Oil seeds and oleaginous fruits and fats
5	05	Wood and cork
6	71,72	Natural and chemical fertilisers
7	64,69	Cement, lime, manufactured building materials
	61,62,63 65	Crude and manufactured minerals
8	41,46	Iron ore, iron and steel waste and blast furnace dust
9	45	Non-ferrous ores and waste
10	04,09	Textiles, textile articles and man-made fibres, other raw animal and vegetable materials
	84	Paper pulp and waste paper
11	21,22,23	Solid mineral fuels
12	31	Crude petroleum
	32,33,34	Petroleum products
13	83	Coal chemicals, tar
14	81,82,89	Chemicals other than coal chemicals and tar
15	95	Glass, glassware, ceramic products
	96,97	Leather, textile, clothing, other manufactured articles
16	51,52,53 54,55,56	Metal products
17	94	Manufactures of metal
18	91,92,93	Transport equipment, machinery, apparatus, engines whether or not assembled and parts thereof
19	99	Miscellaneous articles
20		?

Figure 5.1: The CTSE classification and its conversion to NST/R 2-digit codes

Based on the NST/R-classification, EUROSTAT has developed the '24 Groups of Goods' applied in statistical publications. These Groups of Goods are usually referred to as NST/R-24. The NST/R-24 positions grouping is the basis for statistics that is published by EUROSTAT concerning road, rail and inland waterway transport.

The relationship between the NST/R 2-digits classification and the NST/R-24 is shown in Figure 5.2. Conversion between CTSE and NST/R-24 requires the subdivision of 4 CTSE categories, namely 7, 10, 12 and 15.

EUROSTAT Groups of Goods		
Groups of Goods NST/R-24	**NST/R 2-digits**	**Description**
1	01	Cereals
2	02,03	Potatoes, other fresh or frozen fruits and vegetables
3	00,06	Live animals, sugar beets
4	05	Wood and cork
5	04,09	Textiles, textile articles and man-made fibres, other raw animal and vegetable materials
6	11,12,13,14,16, 17	Foodstuff and animal fodder
7	18	Oil seeds and oleaginous fruits and fats
8	21,22,23	Solid mineral fuels
9	31	Crude petroleum
10	32,33,34	Petroleum products
11	41,46	Iron ore, iron, steel waste and blast furnace dust
12	45	Non-ferrous ores and waste
13	51,52,53,54,55, 56	Metal products
14	64,69	Cement lime, manufactured building materials
15	61,62,63,65	Crude and manufactured minerals
16	71,72	Natural and chemical fertilisers
17	83	Coal chemicals, tar
18	81,82,89	Chemicals other than coal chemicals and tar
19	84	Paper pulp and waste paper
20	91,92,93	Transport equipment, machinery, apparatus, engines - whether or not assembled - and parts thereof
21	94	Manufactures of metal
22	95	Glass, glassware and ceramic products
23	96,97	Leather, textile, clothing, other manufactured articles
24	99	Miscellaneous articles

Figure 5.2: Relations between NST/R-24 and NST/R 2-digit codes

Both CTSE and NST/R are classifications used basically for transport. Since CTSE is based on trade statistics reported by the companies, usually all goods is allocated to the available commodity groups. The group 'miscellaneous articles' covers types of consolidated dispatches. The NST/R has been developed directly for transport purpose. This means that unless the NST/R is aggregated from trade data, the recorded goods would include weight of packaging. Furthermore, consolidated goods will be classified as goods in containers, swap bodies or trailers thus making a translation between NST/R and CTSE very difficult. A number of the major deviations observed in statistics based on CTSE and NST/R are most likely caused by this difference.

In databases for sea transport, a drawback is that containers, RO-RO units and barge transport are practically always categorised by the NST/R

4-digits code 9990 that is included in the NST/R 2-digits code 99: 'miscellaneous articles' (ref. MESUDEMO, 2000).

5.1.3 The NST-2000 outline proposal

Currently the NST nomenclature is being discussed with the purpose of designing a new NST-2000 nomenclature, which better fits the present needs for transport statistics. There are a number of other reasons for reconsidering the nomenclature.

The NST nomenclature was designed with reference to the trade nomenclature of NIMEXE. However, the latter was replaced by the CN (Combined Nomenclature) also used in Denmark. There are no direct links between the CN and NST, but a link can be established as seen in the SAMGODS model in Sweden.

The NST groups are not considered as adequate, and the number of groups in the NST/R-24 is too large for practical purposes.

The EC has decided that all goods classification systems should be linked with the economic activity expressed in the 'Classification of Products by Activity' (CPA) nomenclature. Thus the NST-2000 proposal is a completely new commodity classification system with strong links to the trade flows following the CPA classification.

A preliminary proposal for the NST-2000 classification is presented in Figure 5.1. The composition of some of the groups is still under discussion, and further modifications are expected.

NST-2000 Group	Short description	CPA-code description and/or other description
01	Agricultural products	A: Products of agriculture, hunting and forestry B: Fish and other fishing products; services incidental to fishing
02	Coal, crude petroleum	C: Products from mining and quarrying
03	Metal ores	
04	Food, beverages, tobacco	DA: Food products, beverages and tobacco
05	Manufactured products: Textiles, leather, and related products	DB: Textiles and textiles products DC: Leather products
06	Wood and products of wood and cork	DD: Wood and products of wood and cork DE: Pulp, Paper and paper products, recorded media; printing services
07	Refined petroleum products	DF: Coke, refined petroleum products and nuclear fuel
08	Chemicals, chemical products, rubber and plastic products	DG: Chemicals, chemical products and man-made fibres DH: Rubber and plastic products
09	Other non-metallic mineral products	DI: Other non-metallic mineral products

10	Basic metals, fabricated metal products	DJ: Basic metals, fabricated metal products
11	Transport equipment	DM: Transport equipment
12	Furniture, other manufactured goods N.E.C.	36: Furniture, other manufactured goods N.E.C.
13	Manufactured products: Machinery and electronic products	DK: Machinery and equipment N.E.C. DL: Electrical and optical equipment
14	Secondary raw materials, municipal wastes and other wastes not specified in the CPA	37: Secondary raw materials Municipal wastes and other wastes not specified in the CPA
15	Mail, small parcels	Normally used for goods transported by postal administration and specialised courier services
16	Equipment and material utilised in the transport of goods	Such as: empty containers, palettes, boxes, crates, roll cages
17	Grouped goods	
18	Unidentifiable goods	
19	Other goods	

NB! The groups 17-19 are still under discussion

Figure 5.1: The NST-2000 outline proposal (December 2000)

The NST-2000 is simpler than its predecessor is. It has been important to define a limited number of groups, which were adequate in size in order to avoid problems of representation. NST-2000 consists of only 15-19 groups and with no division into subgroups. The intention is to use these groups both for the exchange of statistics between various EU bodies, to include the Accession Countries in this exchange, and to use the groups for the general dissemination of statistics.

Remarks concerning the NST-2000 classification:

The problem of discontinuity with regard to the existing statistics is evident. However, the NST-2000 proposal indicates that it is possible to establish reasonably accurate methods in order to convert existing statistics to the new classification system. This conversion will be necessary in order to avoid the loss of an immense data source depicting the last 20 years of development and trends in goods transport.

In principle the linkage between the NST-2000 and the CPA is very appropriate for modelling, as it might facilitate OD-zoning inputs concerning production potentials that are consistent with the goods transport flows.

As mentioned above, an important criterion for the design of the NST-2000 is that it should always be possible to determine to which group a certain commodity belongs based on the activities that produces it. However, some special groups are included in the NST-2000 proposal as well. These deal with specific types of goods, which are either difficult to classify according to activity (e.g. loads organised in containers), or do not appear as a trade commodity (e.g. transport of solid waste), or are not linked with a production activity (e.g. transport of empty containers for repositioning).

It should be clarified to which extent a conversion can be made between the SITC and the NST-2000. There is also a need to look into how the NST-2000 classification groups correspond with the logistic family concept (see Section 5.3).

As experienced by the Dutch TRANSITIE pilot study, a large share of containers, RO-RO units and barge transport in export is presently nearly always indicated in the national sea transport database under the NST/R 9990 code.

A similar experience was pointed out by the REDEFINE project. In its list of recommendations is raised one point related to the NST/R 99 goods classification, i.e. a suggestion to split the 'miscellaneous articles' into sub-classes, as this class currently accounts for approx. 25% of all road freight activity. The mixed and aggregated characteristics of this class make it difficult to identify the logistic trends embedded in it.

Containerisation is envisaged to increase and to include a growing share of goods from some of the other NST groups. In principle the process sequences are 1) the manufactured goods (or provided services) to be consigned, 2) the choice of transport loading unit (e.g. bulk, liquid, tanker, container, swap body, trailer, RO-RO etc.), and finally 3) the choice of mode and/or chain of transport. There are even examples where breakbulk is being containerised.

These reasons are contributing to the still pending definition of the groups 17-19 under NST-2000. It will probably take a few years time, before EUROSTAT will be able to implement the new NST-2000 classification as described above.

5.2 EU RESEARCH ACTIVITIES

The lack of solid transport data has been dealt with in a range of different research projects.

The specific research programme for transport within the EU 4th Framework Programme for Research and Technological Development (FP4) and other initiatives have launched a range of project and activities that are relevant for freight transport statistics and modelling. Among those particularly focusing on transport statistics are:

CONCERTO	Concerted Action Committee on information systems.
GEOSYSTRANS	Geographical database on European transport.
INFOSTAT	Information systems and statistics for strategic transport planning in Europe.
MESUDEMO	Methodology for establishing general databases on transport flows and databases on infrastructure networks.
MYSTIC	Methodology for Statistical Analyses, Modelling and Data Collection

REDEFINE	Relationship between demand for freight transport and industrial effects
TRILOG	Overview of global supply chain management.

The findings of these projects of relevance to goods transport data are briefly commented upon below.

CONCERTO:

The concerted action is an ongoing activity initiated by the EC and with the participation of EUROSTAT. It consists of regular meetings with representatives of the Member States, in order to compile and co-ordinate the research activities related with European transport information. CONCERTO is served by a secretariat.

GEOSYSTRANS:

This project has concentrated on developing a methodology for the establishment of a main framework regarding demographic and socio-economic data in a European spatial context. Thus the project mainly relates to the group of data influencing transport demand and not directly to transport flow data.

INFOSTAT:

The aim of the INFOSTAT project was to identify the ideal requirements of a new European Transport Policy Information System (ETIS) considered as complementary to national transport statistical systems. The project was finally reported in 1998.

Concertation with national administrations is found to be essential for the success of ETIS. INFOSTAT concludes that '*ETIS definitions and architecture should minimise the needs for the Member States to change definitions used by their national information systems*'. INFOSTAT further advises that '*As data needs and requirements at the EU transport policy decision-making level will change continuously over time, a different approach to data collection and processing may be considered, e.g. based to a larger degree on ad hoc surveys*'.

Further recommendations related to INFOSTAT are presented in Section 5.3.

MESUDEMO:

The MESUDEMO project was under completion in year 2000. The aim was to follow up on the conceptual definitions that were elaborated by the INFOSTAT project. MESUDEMO has focused on providing practical and realistic proposals for the operational contents of the concepts, and with a view to implementation.

The background for the INFOSTAT and MESUDEMO projects was that current statistical practices of most countries lack the European dimension. Thus there is a need for national efforts to extend the geographical scope of data collection, in order to cover border-crossing flows and produce O/D information for such flows.

The need for a future ETIS is supported by the following quotation from the MESUDEMO reporting:

'*All over Europe central and local governments, institutions, businesses and citizens every day make transport and transport motivated decisions: where to go and how to get there, what to ship and whereto, which transport modes to use, and where to locate facilities and make investments. Transportation constantly responds to external and internal forces such as individual and collective preferences, market conditions, population patterns, safety concerns, weather conditions, and energy and environmental constraints. Good decisions require having the right information in the right format at the right time. Relevant information about goods transport, passenger travels and transport mode use is basic for understanding the transport demand*' (MESUDEMO, 2000).

The requirements of a transport policy information system are that the concepts used are appropriate, that there is a match between concepts, operational definitions and measurements, that information is at the appropriate level of detail, and that the statistical information is timely.

The objectives of the part of the project addressing data on goods and passenger transport flows to be included in ETIS were formulated as follows:

- Definition of the indicators and the variables relevant for European transport flows.
- Identification of the harmonisation procedures, which will be needed in each country, in order to ensure that the indicators and the corresponding data can be brought together in a system leading to one database.
- The methodology recommended should lay the foundation for a consistent transport chain database.

Based on a pilot study (TRANSITIE) in the Netherlands it is concluded that:

'*After the disappearance of the European inner borders the quality of import, export and transhipment statistics has deteriorated extensively….If no action is being undertaken the quality of the statistics will get worse rather than get better*' (MESUDEMO, 2000).

A general conclusion is that the lack of more detailed O/D information is one of the major shortcomings of the existing transport statistics. Traditional statistical data on European transport flows are mostly restricted to uni-modal registration at an annual basis of tonnes and tonne-kilometres that are broken down by commodity group. The geographical specification is confined to the country of origin and country of destination (NUTS 0 level).

The aim is to build a harmonised database that gives sufficient insight into goods transport system performance. In this relation, a main conclusion from the establishment and harmonisation of the ATIS database

(Alpine case study) was that it is more efficient to aggregate databases from different sources than to rely on one single source/database.

MYSTIC:

Originally this FP4 project concerning a methodology and evaluation framework for modelling included a component on goods transport, but the main focus has been on passenger transport.

REDEFINE and TRILOG:

These two EU supported research projects were finally reported in 1999. They do not contribute with any specific conclusions on modelling structures and data collection. However, the two studies identify some policy issues and trends in freight logistics that should be considered by the future freight transport databases and modelling.

The REDEFINE research consortium consisted of partners from the Netherlands, Sweden, UK, France and Germany. The main focus of the REDEFINE study, which included the mentioned countries, was to analyse the relationship between economic activity and road freight traffic. This involves the question about how much of the growth of lorry traffic is a function of economic growth and how much is attributable to logistic changes.

The specific objectives of REDEFINE was:

- to model the factors affecting the increased demand for road freight, and relationships between these factors and changes in industrial and logistic structure;
- to develop strategies to manage and improve road freight transport and logistics in order to reduce or arrest the negative externalities caused by transport;
- to forecast the effectiveness of alternative policies, and where such policies might fit into the comprehensive transport strategy of governments.

The results of the analysis (REDEFINE, 1999 indicated that

- an increase in the average length of haul is the single most important contributor to increased road freight transport demand;
- the increased average length of haul has led to a significant growth in vehicle km over and above the growth in production.

The major factors contributing to this trend are the spatial concentration of production and inventory, wider geographical supply and distribution, and the concentration of international trade on hub ports. The REDEFINE study analyses the logistic trends in more detail, with a focus on supply chains, and discusses a number of policy measures and their likely effectiveness to reduce the externalities of road freight transport.

The TRILOG research consortium was represented by the same countries as REDEFINE and with a couple of overlaps of specific partners. Focus of the study was on the Supply Chain Management (SCM) concept, and the related prospects of inter-modal transport in Europe. The summary report (TRILOG, 1999) points out some interesting development trends in logistics and freight transport that is of particular interest for the development of the transport industry.

The particular conclusions of TRILOG on inter-modal transport are that the market share in the EU is limited to 7.5%, but has shown a considerable growth; and that in some EU regions the geographical and infrastructure constraints determine the modal choice.

The TRILOG report finally advises that indicators and data in order to monitor the developments (on logistics and SCM) are lacking, and that the collection of data at the 'meso' (branch) level may be facilitated by the current ICT developments.

5.3 MAIN FINDINGS FROM EU RESEARCH

This section summarises the main findings and recommendations of the INFOSTAT and MESUDEMO projects, and presents some conclusions about the data levels.

The general approach of modelling in relation to data requirements should be pragmatic and practical, assuming that in the very short term the quality of available statistics will not improve.

5.3.1 Recommendations by INFOSTAT

The situation and relevance of existing data has been assessed by INFOSTAT. The project points out that satisfactory data sources already exist for all EU countries to adequately describe the following fields:

- Topographic and land-use data. These are normally available at the NUTS 2 level with GISCO as the appropriate source. However, main demographic and economic indicators are available mostly at country level (NUTS 0) only.
- Detailed physical and operational transport network characteristics. These transport supply data are available for most countries both in national and international databases.

As for global indicators of goods transport demand, such as tonnes and tonne-kilometres broken down by mode and commodity group (according to the NST classification), it is possible only to get data of reasonable quality at the NUTS 0 level.

The INFOSTAT review identified the following main gaps and deficiencies concerning data related to goods transport:

- No coherent data exist on OD flows at the European level. INFOSTAT concludes that: *'Here the absence of transport chain information is a very important gap in policy decisions. Other deficiencies are the lack of transit flows, and the non-existence of appropriate data on inter-modal and dangerous goods transport in Europe*' (INFOSTAT, 1998).

- INFOSTAT states that data on combined transport infrastructure (including necessary mobile equipment) are missing.
- No database on transport prices has been identified at the multinational level (except for air transport), and no systematic collection of reliability indicators has been established.

The INFOSTAT reports conclude that a future ETIS cannot be fed more or less directly from existing databases, and it is recommended that additional efforts will be necessary in three different fields (INFOSTAT, 1998):

- Harmonisation procedures for building ETIS indicators on the basis of existing data
- Co-ordination of ETIS standards and Statistical Directives at the EU level
- Complementary data collection or changes in the data collection to be initiated in the long term.

The availability of indicators, in the sense of harmonised and coherent data for all EU Member States, are usually restricted to the country level (NUTS 0 level), as EU policy was initially geared to the supra-national level. The transport statistical system has not kept pace with the requirements of broader policy approaches including cross-border regions of Europe and Trans-European networks.

Although the concept of inter-modality has been given a high policy priority in recent years, a common definition of inter-modal transport has not yet been accepted. The INFOSTAT project concludes that:

'The current philosophy of statistics is still mode-based, neglecting the empirically and politically desired integration of single modes into intermodal networks, and more generally a complete view of the chain of transport from the region of production to the final destination of the goods. Special attention has to be devoted to the transfer points and links between modal networks' (INFOSTAT, 1998).

Following the introduction of the Single Market in January 1993, the traditional customs based statistics for modal distribution of internal EU trade was abandoned and the available database smouldered. At the same time relevant data have become more urgent. This is due to the growing policy needs, transport modelling needs, and the needs for monitoring of the drastically changed trade flow pattern between Western and Eastern Europe following the commercial integration since 1989. As a result, there is an acute need for both interregional and international OD-matrices.

INFOSTAT recommends that immediate actions should be taken in order to provide:

- OD transport chain matrices for the various transport modes and commodity groups at defined NUTS levels;
- insight into causal mechanisms relating to goods transport;

- linkage between goods transport and goods traffic;
- revitalised data reporting procedures.

INFOSTAT also suggests some immediate actions to be taken concerning inter-modal transport. This includes that the EC (Directorate General for Transport and Energy, DG-TREN) and EUROSTAT should study the Statistical Directives with a view to inter-modal transport in the following directions:

- the scope for compiling inter-modal transport data from railway companies and shipping companies should be looked into;
- a European-wide region-to-region transport chain database, including chains of up to two transhipment points, should be developed;
- the development of telematics (e.g. within tracing and tracking) together with traffic counts could have a potential for supplying information about inter-modal transport.

As already mentioned, a future ETIS is considered complementary only to national statistical systems. However, for this purpose INFOSTAT advises that it might be necessary to ensure a certain degree of standardisation and harmonisation between national databases.

5.3.2 Indicators suggested by MESUDEMO

The INFOSTAT project suggested that the grouping of statistical indicators should include a group of transport demand indicators.

The MESUDEMO has specified six goods transport demand indicators, out of which four are considered fundamental. The indicators, which have to be regarded as the desired end-result of a transport policy database, are presented in the overview table of Figure 5.1:

No.	Variable label	Unit of Measurement	Observational Unit	Priority Level
1	Total annual goods transport flow between the zones - by commodity group, mode (or combination of modes) and type of transport chain	tonnes/year	O/D pair	fundamental
2	Average distance between origin and destination of the transport unit - by mode	km	O/D pair	fundamental
3	Average distance between origin and destination of the goods – broken down as for indicator 1	km	O/D pair	fundamental
4	Loading factor (ratio between volume and capacity) per type of transport unit - by mode	percentage	O/D pair (main leg)	desirable
5	Annual total number of tonnes transported - broken down by: - size of load/consignment (weight) - value of load/consignment (euro) - trip distance (km)	tonnes/year	O/D pair (main leg)	fundamental

	- containerisation (yes or no) - type of transport unit			
6	Average number of hours of use of the transport unit	hours/year	zone	desirable

1) Multiplying the tonnes/year and the distance between O/D pairs can derive tonne-kilometre figures.
2) It might be considered to present the value flows in the same spatial and modal details as volume (in tonnes) flows

Figure 5.1: Goods transport demand indicators

The projects recommend a policy oriented and comprehensive database, keeping the following aspect in mind when aggregating the linked transport information:

- A harmonised classification should be introduced to be able to ensure that the most discriminating information is included, e.g. the classes to be used for indicating the weight or value of the consignment
- Individual items should be counted in new variables, e.g. number of trips, number of consignments, etc.
- Some additional variables should be calculated, e.g. weight multiplied by distance of shipment in order to derive at tonne-km figures, and capacity multiplied by (loaded and/or empty) trip distance, etc.
- Consistency with the classification used in trade databases for building transport chains
- Consistency with the classification used in other parts of a future European transport information system (socio-economic data, passenger flows, infrastructure links and nodes, impacts).

Specification of indicators:

The MESUDEMO project presents the following specification of each indicator for practical applications and modelling purposes.

Indicator 1 (goods transport flow volumes between zones)

The 'original' origin and the 'final' destination of a transport chain should indicate some original and final activity respectively, e.g. production processing. The transport chain specification is discussed under indicator 3.

Indicator 2 (average distance of transport unit)

The origin and destination of the transport unit are not always the same as the place of loading and unloading of the goods during the trip.

The transport performance figures in tonne-kilometres only include the loaded kilometres of a trip, whilst vehicle-kilometres include both loaded and empty vehicles.

A complete presentation of this type of indicators in the database record structure could be as follows:

1. Zone of loading of the transport unit
2. Zone of unloading of the transport unit

3. Transport mode
4. Transport unit
5. Weight (in tonnes) transported
6. Transport performance in tonne-kilometres
7. Average distance in kilometres (= 6/5).

Indicator 3 (average distance between origin and destination)

As an example, for a transport chain involving one transhipment only, the average distance could be represented as follows in the database record structure:

1. Zone of original origin of the goods flow
2. Zone of final destination of the goods flows
3. Zone of transhipment
4. Mode used from original origin to node of transhipment
5. Mode used from node of transhipment to final destination
6. Weight (in tonnes) transported
7. Transport performance in tonne-kilometres
8. Average distance in kilometres (= 7/6).

Indicator 4 (loading factor)

The capacity of a transport unit can be specified in various ways, for example as weight carrying capacity, volume carrying capacity, and capacity in number of units.

A database record for the loading factor could be specified as follows:

1. Zone of loading of transport unit
2. Zone of unloading of transport unit
3. Mode of transport
4. Transport unit
5. Weight (in tonnes) transported
6. Carrying capacity (in tonnes) moved
7. Loading factor.

The MESUDEMO report is not clear as to whether a 'transport unit' in this particular context is used according to the ECMT terminology that implies both including vehicles used and equipment, e.g. containers. In this context concerning capacity and loading factors it probably makes most sense to replace 'transport uni' by 'means of transport unit' or 'rolling stock', e.g. truck, train, or vessel. However, the number of empty containers being transported is also a relevant figure for the database, as they represent unused capacity often in connection with unbalances between import and export flows.

Indicator 5 (annual total number of tonnes transported)

This indicator provides information about the *main leg* of a transport chain. It is related to the load (consignment) carried by a specific type of transport unit. However, in more complex transport chains it might be arbitrary to choose a certain leg as the main leg.

The indicator is measured in total number of tonnes but could include the total value that has been transported as well. Thus a database record could be represented with the following specifications:

1. Zone of loading of the transport unit
2. Zone of unloading of the transport unit
3. Transport mode
4. Transport unit
5. Load weight class
6. Load value class
7. Trip distance
8. Cargo type
9. Weight (in tonnes) transported
10. Value (in euro) transported.

The specification in classes - of the load weight, of the value of the load, and the trip distance and cargo type - intends to make a link to the modal choice. It offers the possibility to differentiate e.g. between bulk and general cargo, between liquid and dry bulk, and between loading units such as containers, swap bodies etc.

The use of EDI, tracing and tracking etc. was originally suggested by the INFOSTAT project as a separate indicator. MESUDEMO recommends it to be aggregated to an O/D relation under indicator 5, with perhaps further specification of the type of consignment.

Indicator 6 (average annual number of hours of use)

This indicator intends to describe the availability of the transport unit used. It is considered more adequate to measure it in effective hours rather than in days per year.

The database record structure could be as follows:

1. Zone of registration
2. Transport mode
3. Transport unit
4. Carrying capacity class
5. Number of transport units
6. Average weight carrying capacity
7. Average number of hours of use per year.

The reason for including more information about the carrying capacity than the number of hours actually used is in order to make the indicator more supply-oriented.

5.3.3 Levels of data in relation to transport models

In a second phase, the TRANSITIE pilot study, which was carried out under the MESUDEMO umbrella, aimed at considering four levels of information to be included in the database structure:

A. transport mode and unit used;
B. trip by the transport unit;

C. load carried
D. transport chain information.

As for levels A-C, information could probably be acquired at the individual level through transport statistics, particularly in the cases where registration is based on the vehicle or transport organisation, such as the Danish national truck surveys (see Chapter 6). These levels of information could be compared with the actual situation of a logistics operator:

'In planning the consignments they get from shippers, each of these orders should be assigned to one of their transport units to get it from the original origin to the final destination. Big operators running a global (multimodal) network create in fact their own transport chain in choosing the optimal route through their network' (MESUDEMO, 2000).

In principle, the national statistical organisations could obtain this type of data by means of EDI. However, a conclusion of the MYSTIC project on this topic is that there is still a long way to go before representative statistical figures can be obtained from EDI-equipped companies.

As for the practical way to collect the information, the TRANSITIE study suggests in the short term to provide the information through shippers' surveys, and in the medium or longer term through the EDI systems of shippers and major transport integrators. Through the same process, it is possible also to obtain information about the economic and logistical determinants of transport.

5.4 NEW DATA CONCEPTS

The data collection procedures in use have been developed in the past, and thus they reflect some business and logistic conditions (dominated by economies of scale) that differ from present and expected future conditions (increasingly reflected by economies of scope). In addition, the establishment of the EU internal market has considerably diminished the possibilities of retaining any adequate data about the modal distribution of trade and transport flows. This accentuates the necessity of identifying some new concepts and approaches in terms of data collection.

The following new statistical concepts have been suggested and briefly explained by the INFOSTAT and MESUDEMO projects:

- The transport chain
- Inter-modality
- Logistic families
- Transport quality
- Infrastructure and environmental bottlenecks
- The transport corridor
- Advanced accessibility.

The transport chain:

The INFOSTAT project defines a transport chain as:

'*a sequence of transport modes used to carry a certain quantity of goods from its origin to its final destination. Along the chain, one or more transhipments may take p*lace'.

As indicated above MESUDEMO suggests a database record structure for a transport chain with one transhipment node. It is being realised that more comprehensive types of transport chains will lead to a very complex data structure. This problem is related to several of the indicators presented in Figure 5.1 above.

Inter-modality:

The TRILOG project presents the EC launched definition of inter-modal freight transport ('freight inter-modalism') as follows:

'*The optimal integration of different transport modes, enabling an efficient and cost-effective use of the transport system through seamless, customer-oriented door-to-door services, whilst favouring competition between transport operators*' (TRILOG, 1999).

Inter-modality is considered to be a subset of the transport chain concept, as the latter consists of different modes.

Logistic families:

A logistic family is a grouping or classification of commodities, with a view to how they are being handled and transported from a logistical point of view. Thus they are featuring similar logistical requirements, that are more closely related to manufacturing and commercial behaviour, ref. economies of scope. A classification according to such a principle is still at the conceptual stage only. As indicated in Chapter 3, the REDEFINE project had specified some logistical clusters taking into account the characteristics of the transport, but neglecting the usual commodity classification. The application of the logistical clusters had been carried out in Sweden. This required a major conversion ('translation') work between the NST/R classification applied in the transport surveys and in the distribution model, and the logistical clusters used in the STAN99-model. The conversion is based on the Combined Nomenclature (CN), which is used as classification for import and export expressed in tonnes and value. NST/R and CN are linked, and a translation has been produced between logistical clusters and the 4-digit CN containing 1200 commodity groups.

It has been seen as a disadvantage that the logistical clusters are completely detached from the usual commodity classification applied in transport statistics. Therefore another EU research project has aimed at creating a linkage between the logistical clusters and the usual commodity classification. This project, STEMM (Strategic European Multimodal Modelling), is based on commodity groups that are as homogeneous as possible in relation to following characteristics:

1. Price (low, medium, high, very high) in EURO/tonne

2. Delivery size (small, medium, large)
3. Density (low, medium, high) in tonne/m3
4. Type of goods (bulk, chemicals, parcelled)
5. Temperature control (yes, no)
6. Risk of damage (low, medium, high)
7. Level of service (low, medium, high).

The STEMM project proposes the use of 12 different commodity groups based on the SITC classification.

Commodity	**Price**	**Delivery size**	**Density**	**Type**	**Tem-perature control**	**Risk of damage**	**Level of service**
1. Food and live animals							
	Medium	Medium	Medium	Parcelled	Yes	Medium	High
2. Beverage and tobacco							
	Medium	Medium	Medium	Parcelled	Yes	Medium	High
3. Crude materials, inedible, except fuels							
	Low	Large	Medium			Low	Medium
4. Mineral fuels							
	Low	Large	Medium			Low	Medium
5. Animal and vegetable oils, fats and waxes							
	Medium	Medium	Medium	Parcelled	(Yes)	Medium	Medium
6. Chemicals and related products							
	Medium	Medium	Medium	Chemi-cals	(Yes)	Medium	Medium
7. Paper, paperboard and articles of paper, pulp							
	Medium	Large	High	Parcelled		Medium	High
8. Metal products							
	Medium	Large	High	Parcelled		Medium	Medium
9. Manufactured goods							
	Medium	Medium	Medium	Parcelled		Medium	Medium
10. Machinery							
	High	Medium	Medium	Parcelled		High	High
11. Miscellaneous manufactured articles							
	High	Small	Medium	Parcelled		High	High
12. Valuable machinery and manufactured articles							
	Very high	Small	Medium	Parcelled		High	High

Table 5.1: Commodity groups used in the STEMM project based on logistical characteristics

The 'SoftIce' project has also looked into the relationship between models and statistics. The main relevant factors for identifying homogeneous commodity groups in terms of choice of transport solutions were found to be:

- Direct transport costs
- Reliability
- Transport time/speed
- Ability to use inter-modal chains.

Transport quality:

This concept is described by means of a set of performance indicators. A Danish research project (the Danish Transport Council, 1997) identified seven quality performance indicators of particular importance, i.e. transport time, reliability, transport frequency, damage risk, access to information system, flexibility of services, and customs services. Other indicators could be added, such as safety and security etc. The INFOSTAT project suggests including time indicators, reliability, flexibility, qualification, accessibility, control, safety and security.

Infrastructure and environmental bottlenecks:

The MESUDEMO project suggest all national congestion and traffic load data to be processed in a standardised form, in order to reflect the variety of situations at critical periods of the day or year.

Environmental bottlenecks occur where a given link suffers from both traffic and environmental constraints.

Transport corridors:

A transport corridor can be defined as:
'*a geographical sequence of strongly interconnected regions*' (MESUDEMO, 2000).

This concept bridges the gap between single transport links and whole networks. MESUDEMO advises that it is crucial to identify the *economic* corridors that constitute the sources (origins and destinations) of the traffic that flows through the *transport* corridors.

Advanced accessibility:

This indicator ranks how well a region, a city or another area is connected to the infrastructure network in terms of centrality and population density etc. As an example, the total road costs required for a European region to cover a market of say 10 million people could be calculated.

Common accessibility indicators for cities and regions should be defined and applied.

6 DANISH TRANSPORT DATA SOURCES

This chapter provides an overview of Danish statistical publications and sources of potential relevance for the description of freight transport patterns and flows with origin and/or destination within the Danish territory, and for transit flows through Denmark.

The *first section* briefly points out some sources of socio-economic statistics.

The *second section* discusses foreign trade statistics based on the INTRASTAT and EXTRASTAT reporting.

The *third section* presents and discusses the issue of providing data about interregional trade flows that are not published in the national foreign trade statistics.

The *fourth section* presents an overview of Danish mode-specific statistical sources and recommends how to improve them.

Finally the issue of transit flow data is being briefly addressed in *the fifth section*.

Neither at the national Danish level nor at the EU level is there any agreed set of guidelines or rules concerning the harmonisation and standardisation of transport statistics, except for the zoning and trade commodity classification systems described in Chapter 5.

6.1 SOCIO-ECONOMIC STATISTICS

Besides the transport statistical sources provided by Statistics Denmark, some general macroeconomic variables, such as GDP, income levels and employment figures by geographical zones, may substitute the lack of direct production figures in order to estimate the generation and attraction by each particular zone.

The ADAM databank:

This databank includes the macroeconomic key figures at the country level and as far back as from 1947. A breakdown is provided into 19 industrial branches and 10-15 foreign trade components. Figures are provided both in current and fixed prices. The databank is available as the PC-AXIS programme on a CD-ROM.

Databases on various socio-economic variables:

Databases on demographic figures, housing, employment, incomes and commercial companies, etc. are available at the county level (NUTS 3 level) and municipal level (NUTS 5 level).

More detailed data extracts can be provided as ordered statistical services.

Regional macroeconomic accounts:

This statistical source includes some indicators of production and income. It is based on a new national accounting system introduced in 1995, and was first publicised in January 1999. It includes, whenever possible, a regional distribution at the county level of the Danish national account.

The source is compatible with international statistics. It is publicised annually, and data are included in EUROSTAT's REGIO CD-ROM.

6.2 FOREIGN TRADE STATISTICS

The trade statistics is a major source of data about freight flows between Denmark and other countries. Thus it is also a primary source for freight transport modelling.

The foreign trade statistics is based on the data collection via INTRASTAT and EXTRASTAT, see below. Both methods are described in the annual publication 'Danmarks varer. Import og eksport'.

At the national level this statistics is quite detailed and comprehensive as concerns the breakdown into commodity groups. It includes Danish export and import, and a breakdown on approx. 10,500 different commodities. For each commodity is provided the geographical distribution on all of the trade partner countries.

The standard publications that cover the foreign trade statistics present some aggregated data as well. The totals are publicised in SITC categories, but at a more aggregated geographical level than the trade partner countries.

Foreign trade figures are presented both in value and in net-weight, whilst most sources on transport statistics indicate gross-weight figures. A comparison between net-weight and gross-weight figures is not possible without adding or subtracting the weights of packaging etc.

Detailed and tailor-made extracts from the database are possible as ordered statistical services. Such extracts would be necessary for carrying out in-depth analyses and for specific modelling purposes.

A drawback as concerns OD-information is that the foreign trade statistics does not provide a geographical breakdown into regions of the trade partner countries. As a consequence, trade flows between neighbouring regions separated by a national border are not available in the foreign trade statistics, like for instance between the Copenhagen area and Southern Sweden.

EXTRASTAT, trade with non-EU countries:

EXTRASTAT covers trade with all of the non-EU countries. In this case, the sources are the customs documents. The breakdown of figures into commodities within the EXTRASTAT trade statistics with non-EU countries is the same as for trade figures based on the INTRASTAT reporting, see below.

As mentioned previously, the statistical figures do not provide any geographical breakdown into regions of the trade partner countries. As to the geographical origin or destination in Denmark, it is possible to a certain extent to distribute the EXTRASTAT trade figures according to the geographical location of the customs office and the border passage location. However, such detailed data extracts require ordered statistical services.

The modal distribution is available from the EXTRASTAT trade statistics with non-EU countries. These figures are aggregated at the national (NUTS 0) level, but they could be specified at the border passage as well.

It should be mentioned that transport mode refers to the transport mode at the border. This means that no indication exists on inter-modal transport. Moreover, transport mode information was changed in 1993 following the introduction of INTRASTAT. Transport modes like truck on ferry and rail on ferry were abandoned and replaced by sea transport. Therefore it is not possible to separate ferry transport and sea transport in the EXTRASTAT database.

INTRASTAT, trade with EU countries*:*

This reporting method was introduced as from 1 January 1993, following the establishment of the Single Market. Simultaneously, the registration of trade transactions between EU countries, which was previously based on customs documents, was replaced by the INTRASTAT reporting. For Sweden, Finland and Austria this change entered into force two years later, i.e. 1 January 1995.

INTRASTAT is now the common data registration method for internal trade movements between all EU Member States. A breakdown is made on all of the EU partner countries.

The national source is based on reporting from around 10,500 Danish companies. The INTRASTAT declarations are directly transmitted from the enterprises to the national statistical bureau/administration by using monthly recapitulative declarations (MESUDEMO, 2000). A certain number of enterprises representing low import/export quantities are exempted from the obligation to report. The total share of trade of these companies is estimated not to exceed 5% of the total intra-EU trade value for each Member State.

Considering the sample size and composition, the INTRASTAT reporting provides reasonably representative figures about the foreign trade flows and their composition, to and from other EU Member States at the country (NUTS 0) level.

This is, however, not the case when it comes to information about the distribution on transport modes, see the following description.

The basis for collection of transport data by the national statistical offices of the EU Member States was drastically changed by the introduction of INTRASTAT. A drawback of the method is that it is not able to provide any reliable data on the modal distribution of trade flows. As a result, adequate information on the modal distribution of Danish international trade flows that cover EU Member States have vanished since the start of 1993. Thus it is difficult to produce any consistent time series of the distribution on transport modes that includes the period before and after the introduction of INTRASTAT.

Until 1996, 'Statistics Denmark' continued publishing international transport statistics with a breakdown on import/export, commodity groups and transport modes. However, the figures for modal distribution of trade with other EU Member States became increasingly doubtful. A contributing reason is that in many cases the specific transhipment is passing through a transport chain and it is not adequate to categorise it by a single mode. As an example, INTRASTAT categorised goods transport by ferries as transport by ship, and not according to the mode of transport (e.g. lorries/road haulage) using the ferry service. As a consequence, since 1997 'Statistics Denmark' has ceased to publicise modal information in relation with INTRASTAT statistics, and other EU countries are following suit.

A potential advantage of the INTRASTAT reporting is that it contains rather detailed geographical information about the location of the companies involved in export and import. This information could be more efficiently exploited than hitherto. On the basis of the information about the addresses of the approx. 10,500 companies, it should be possible to allocate the trade flows to a more detailed geographical level of regions or counties in Denmark as the source country. It is not possible to provide a geographical breakdown of these trade flows into regions of the trade partner countries. However, information derived from the INTRASTAT reporting for each trade partner country could provide a similar geographical breakdown as to the origin and destination of that particular country.

6.3 DATA ON INTERREGIONAL TRADE FLOWS

In order to facilitate freight transport modelling, there is a need to break down the trade flows into smaller geographical units than the country level (NUTS 0 level). Figure 6.1 illustrates the types of interregional flows in relation to two neighbouring countries A and B.

The *type I* flows are covered by the national trade statistics (INTRASTAT and EXTRASTAT).

As for trade flows between regions located in different countries, e.g. the *type III* trade flows between regions separated by the borders of two neighbouring countries, the data problem has not yet been solved. One of the major statistical challenges in relation to freight transport modelling is to provide some reasonably reliable estimates of the type III flows.

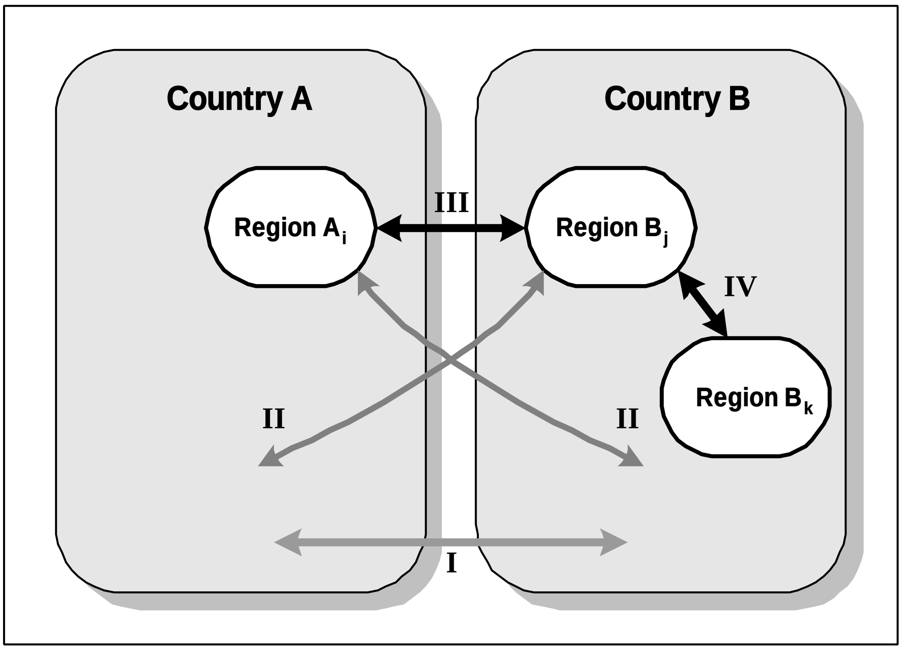


Figure 6.1: Types of interregional trade flows

Statistics for the *type IV* trade flows, i.e. between administrative regions within a country, has been developed in some countries. Figures of domestic trade between regions in Denmark are not regularly available. However, various research studies have been carried out, among others by the Danish institutes 'AKF' and 'Institut for Transportstudier', to establish a methodology in order to estimate trade flow figures between the Danish counties. A fully transparent methodology has not yet been published, but reasonable estimates of regional trade flows within the country could be elaborated.

Within the Interreg IIC Programme ('North Sea Commission') and the NTN ('Nordic Transport policy Network') project, a research study has addressed the question about establishing OD-matrices for goods of the *type II* trade flows. The study included Danish, Norwegian and Swedish counties or regions located around the Skagerrak and Kattegat Seas. It was reported mid-2000. Trade flow figures were presented for the year 1997, including a breakdown on commodities and modes. A fully developed common methodology for cross-border trade flow between regions in Jutland, Southern Norway and Western Sweden was not presented as a result of the study.

The foreign trade statistics for Norway is still fully based on EXTRASTAT, and also enables a geographical breakdown on each county ('fylke'). Thus it was possible for the study report to present, for each Norwegian county, some estimates of the import/export trade flows with all of the neighbouring countries and the rest of the World. These estimates are based on studies carried out by two Norwegian research bodies, i.e. Institute of Transport Economics and Agder Research Foundation.

In the case of Sweden, it was not possible to provide a breakdown into counties ('län'). Trade flow figures with a few selected European countries were presented for Western Sweden ('Västre Götalandsregionen'). These figures were estimates provided by the EU research project SCANDINET. The Swedish SAMGODS model contains estimates of trade flows in tonnes between municipalities in Sweden and regions in other countries. Sea transport, however, is only included as transport from import and export port to the relevant municipalities. For rail and road transport, the complete transport from origin to destination is included and used in the STAN mode and route choice model.

As for Denmark, the NTN study presented some foreign trade flow figures for Jutland divided into two zones (Jutland North and Jutland South), i.e. a higher geographical aggregation level than the county ('amt'). These estimates were provided by the 'Institut for Transportstudier', based on a study carried out about the commercial and industrial impact of a fixed Fehmarn Belt link.

As mentioned above, the 'Statistics Denmark' has recently started publicising a regional distribution, at county level, of the national account (in terms of production and income indicators). This statistics could form part of the basis for regional trade flow modelling at the aggregated level and both involving domestic and foreign trade.

6.4 MODE-SPECIFIC STATISTICS FOR DENMARK

The selection of data sources in this section is made with a view to the use of statistical variables for the modelling of flows in major corridors or along main routes.

The existing and most relevant sources of mode-specified Danish transport statistics are reported quarterly and/or annually. They cover the following areas:

- Ferry traffic
- Maritime traffic
- Railway transport
- Goods transport by Danish trucks.

The sources are commented briefly upon below. The sources are based on reporting from operators or on regular/annual sampling. They are published in 'Statistiske Efterretninger' under the heading of 'Transport' and in the Statistical Yearbook.

Freight transport by air is not addressed by the present report, as its share is still insignificant compared to other modes in terms of weight/volume. However, the share calculated in value figures is increasing rapidly.

Piped transport is increasing but the transport volumes are not reported.

Passenger ship and ferry traffic via Danish ports:

The publication on passenger ships and ferry traffic is distributed quarterly and provides statistical figures based on monthly reporting by the commercial ferry operators. The statistics is divided into traffic on international routes and traffic on domestic routes. The publication provides traffic figures, i.e. number of ferry trips, number of vehicles served including the distribution on single lorries, trucks with trailers, trucks with semi-trailers, and unaccompanied semi-trailers. However, for some of the ferry routes (e.g. between Hirtshals/Hanstholm and Norway) a breakdown into these categories of trucks is not available.

A breakdown on direction is not presented in the publication. However, this information is available in the basis material, and it is of particular importance in order to help separating import flows from export flows.

It should be noted that the figures include transit traffic through Denmark.

Maritime traffic via Danish ports:

This annual publication on maritime traffic provides a very informative and useful overview of the maritime traffic for each port. It is based on reporting from the ports and the relevant ferry operators.

The publication presents the annual freight transport figures in tonnes on each ferry route, including a breakdown on lorries, railway wagons and miscellaneous. However, it should be noted that the tonne figures are calculated data that are based on the traffic figures and estimates of average load figures for the trucks.

The present form of statistics was introduced in 1997. It includes, as main variables, the number of ships served (per month; per year only for smaller ports), tonnes of incoming/outgoing freight, and category of commodity.

For each port, a breakdown is provided on major ship types and sizes, on domestic and export/import goods in tonnes, and between in- and outgoing flows in tonnes. Separate total figures are presented for liquid bulk and for ferry freight.

For each port, a geographical breakdown is presented of domestic goods distributed on each Danish county (NUTS 3 level). Export/import figures are broken down into most of the European countries, and into continental regions as for the rest of the world.

The publication specifies the total annual throughput for each port into 21 commodity categories. Five of these can be grouped as liquid bulk, and nine categories as solid and/or break bulk. The remaining seven categories are wood products, steel and metal products, other general cargo, goods in containers, non-licensed motor vehicles, ferry goods, and other RO-RO goods. In 1998, around 30 Danish ports obtained a total annual throughput exceeding 0.5 million tonnes. However, the 15 largest ports contributed to more than 75 pct. of the overall throughput.

A special overview of annual throughputs, in tonnes and units of containers and RO-RO, is presented for the few Danish ports servicing this type of traffic. The two dominant ports are Aarhus (as for containers) and Esbjerg (as for RO-RO). These are followed by Copenhagen as the second largest container port, and by the Aalborg port area that is mainly serving bulk except for the container traffic with Greenland.

As for ferry traffic, a distinction is made between goods by lorries and goods by railways. The DSB ferry service between Nyborg and Korsoer ceased to operate in 1997 when the fixed railway link across the Great Belt commenced its operations. The opening of the fixed link across the Great Belt has also reduced the role of the Roedby-Puttgarden route as regards rail ferry freight. Since 1999, the servicing by Stena Line of railway wagons on the Frederikshavn-Gothenburg route has been insignificant. Only a few railway wagons are transported by Color Line on the Hirtshals-Norway ferry routes. As from year 2000, rail ferry freight on the Øresund routes has been affected by the opening of the fixed link between Copenhagen and Malmoe.

Traffic through Danish Waters:

Lloyds keeps a register of ship movements based on port calls. Each single ship in the world merchant fleet is monitored, and in this sense it is possible to estimate the calls made by different ships, including frequent routes applied by ships not moving in liner traffic. Furthermore, this makes it possible to evaluate the nationality of the ships in operation, the average size of the ships, and the number of calls i.e. frequency for each ship.

The Lloyds Voyage Records have been used in a Danish study comprising the effect of a deepening of the strait of Drogden in Øresund. The Voyage Records (VR) alone are not able to provide a complete picture of the traffic. The VR does not include ships apart from those in the merchant fleets of the different countries. Ships in scheduled traffic e.g. ferries or RORO-ships are not included. Therefore, the VR must be complemented with scheduled traffic. Secondly, even if the VR are quite accurate, there are still uncertainties about some of the data items, which has to be considered.

In order to distribute the ship flows between the Øresund route and the Great Belt route, a calibration of the VR-data was performed based on the observations made by the Danish VTS monitoring of the traffic sailing through these main routes. It is therefore possible to estimate the number of ships per year in the Danish waters by ship type and size.

The databases do not render any indication of goods transport by ship. An identification of goods transport through the Danish waters will have to be established by means of a model combining ship types, ship sizes and commodity groups as indicated by trade data. The Swedish SAMGODS model may be able to establish such data.

Railway transport:

The annual statistical publication on railway transport is based on reporting from the Danish Railway Administration ('Banestyrelsen'), The main rail goods operating company in Denmark, Railion (since June 2001), and the regional railway services providers ('Privatbanerne').

The main part of the publication is occupied by statistics about the rolling stock. As for goods transport, the information provided is limited. Total annual production figures since 1990 are presented in tonnes and tonne-kilometres, and for national, international and transit respectively. For each of these figures, a breakdown is made on general cargo, combined transport cargo (containers, swap bodies, trailers) and full wagonloads (block trains and miscellaneous).

An OD-matrix is provided of the annual figures (in tonnes and tonne-kilometres) for domestic transport between the counties (NUTS 3 level). It should be noted that the domestic freight transport quantities produced by rail services are relatively modest compared with international rail transport (export/import and transit) and compared with domestic transport by road and sea.

The publication also presents some figures of the daily average number of passenger and freight trains on particular subsections of the railway network. However, this only gives a rough idea of the overall geographical flow distribution.

The publication does not provide a clear pattern of the geographical distribution of the import and export flow figures. The inlets and outlets would notably be the routes via Øresund, Roedby-Puttgarden and Padborg. However, since the fixed railway link across the Great Belt was opened in 1997, the freight traffic via Roedby-Puttgarten has been rerouted to pass via Padborg. It would also be relevant to present a breakdown of goods flows on major sections of the trunk network, including flows via the Great Belt and Øresund links.

As a conclusion, the present railway statistics is insufficient as regards key information about the geographical distribution pattern of the international and transit freight transport flows. However, Railion (DSB) may be able to provide such aggregated geographical information, at least since 1997 when its databases were computerised. The access for more detailed statistics is probably restricted by Railion for commercial reasons.

Goods transport by Danish trucks:

This statistics is published quarterly and as a more comprehensive annual publication. The statistics is based on periodic sample questionnaires addressing goods carried by Danish trucks and by means of the driver's logbook (see below).

A drawback as concerns international goods transport by road is that the sample does not include freight transport by foreign trucks. Thus the picture provided is only partly representative for the overall goods flow pattern by road in Denmark.

Similar sample surveys are also carried out in other European countries. The survey covers a certain proportion of national trucks above 6 tonnes gross-weight that is active in international haulage. Each record in the logbook describes the origin and destination, the amount and type of goods transported, or if the truck was empty during the trip.

Based on these data it is possible to relate the development in road transport of different commodity types with the economic development in different business sectors. However, the sample is relatively small, and therefore the survey is not able to establish very detailed pictures of segments of the goods transport market.

National goods transport by Danish trucks:

The figures on domestic haulage are based on a quarterly questionnaire sampling. Out of the approx. 43.000 licensed lorries (above 6 tonnes gross weight), a sample of 900 lorries is questioned about the activities during one week.

The publication provides a distribution of the figures (in tonnes and tonne-kilometres) on several categories of commodities, which is at the same level of detail as the NST/R-24 groups of commodities. These include four categories of bulk and general cargo as one category. A breakdown of the tonnes and tonne-kilometres figures is also presented on two groups of loading units i.e. containers/swap bodies and 'others'. Furthermore, the presentation includes a specification on dangerous goods types and a breakdown into 12 industrial branches.

Based on the sample questionnaires, this statistical publication also presents an OD-matrix of the annual freight transport figures (in tonnes and tonne-kilometres) between all of the Danish counties. Estimated total figures are presented of the number of trips to/from each county, and the total number of trips and million tonnes between East and West Denmark.

International goods transport by Danish trucks:

Before 1993 these figures were based on the reporting of foreign trade statistics. As from 1993, the figures presented for international haulage are based on a quarterly sample of 700 road haulier companies out of approx. 1600 companies that frequently perform international freight haulage. The sampled companies report on the haulage activities they have started in a specific week or half-week.

For the international haulage a breakdown is provided that presents figures for import, export, cross trade, and cabotage haulage.

As it was the case for the domestic haulage statistics, the presentation of international haulage figures provides a breakdown of the tonnes and tonne-kilometres figures on two groups of loading units i.e. containers/swap bodies and 'others', on several categories of commodities, and a specification on dangerous goods types.

The publication presents a geographical distribution on each country of origin or destination, and a further breakdown of Germany into three regions. These figures are publicised in more aggregated groups, but include a breakdown on 15 European countries, and the total figures for EU countries and 'other countries'.

Based on similar road haulage statistics from other European countries, it should be possible to obtain some information as well about the quantities of transport by foreign truck to/from and through Denmark. EUROSTAT may consider publishing OD-matrices on international freight transport by trucks.

Manual counting of lorry traffic by the Danish Road Directorate:

These traffic figures are provided periodically but not always annually. As an example, during 1997-99 there have been no counts of lorry traffic crossing the border between Denmark and Schleswig-Holstein. Manual counts have been resumed in year 2000, in co-operation between the Danish Ministry of Transport and the Ministry of Taxation. The counts from 2000 provide a breakdown into three national groups of trucks, i.e. Danish, German and other countries.

Other sources including neighbouring countries:

There is a good possibility to extract computerised and detailed data from the ports and terminals, notably the container and RO-RO terminals in the ports of Aarhus, Esbjerg and Copenhagen, and the railway and combined transport terminals. Such data extracts could provide more specific and precise information about the geographical and modal distribution of freight flows, including transit flows.

As concerns ferry transport, the 'Cruise Ferry Info' publication presents aggregated figures for traffic on all international ferry routes, including annual figures of 'trailers'. This source is based on the reporting of ferry operators to the national statistical offices. It does not provide any data about the quantities of freight flows, e.g. in tonne-figures, or any other goods transport information in addition to the national statistics. However, the publication offers a useful international compilation covering all countries.

The Lloyds database on voyage records provides a description of vessel traffic in the Danish waters. However, a comprehensive analysis of the database is required in order to extract these data.

The trade and traffic statistics of Norway, Sweden, Germany and Poland are of particular interest to Denmark. A problem is that the statistical systems vary considerably from country to country. The data are seldom directly comparable, even when it comes to the same international ferry route or cross-border inlet/outlet.

6.5 TRANSIT FREIGHT FLOWS

Figure 6.1 above illustrates the trade flows that are embedded in the INTRASTAT and EXTRASTAT statistics. Transit flows are not indicated.

The national foreign trade statistics does not include any registration of transit trade through a particular EU country, either this transit consists of trade between two other EU countries, or third countries are involved in the transit trade. This shortcoming also applies to Denmark as an important transit country for freight between other Nordic countries and the Continent.

The TRANSITIE pilot study (see Chapter 5) was carried out under the MESUDEMO project and in co-operation with the Dutch Ministry of Transport. The study advises that the problem of the lack of data on intra-EU transit cannot be sufficiently solved in the short term. The TRANSITIE study states that the quality of information about transit through the Netherlands involving third countries has deteriorated during the last few years, especially in the case of transport by sea and rail. 'Disguised transit' is transit freight that is customs cleared in the particular country. The opportunity to distinguish this type from real import/export is envisaged to disappear in the longer term.

The TRANSITIE estimate of transit was based on a sample survey, i.e. during 1997 one month of every quarter was processed together, and the observed transport flows were multiplied by the factor 3 to obtain yearly volumes.

Some examples of the difficulties of estimating countries of origin and transit flows were pointed out in the TRANSITIE study. Transit could already have taken place in another country before being imported by sea, e.g. crude oil from Saudi-Arabia is probably transferred in Egypt and oil from Russia is transferred in Latvia. In the case when a commodity from country X is transhipped to a ship in country Y, in the trade database country X will be registered as origin whilst in the transport database country Y will be registered as origin.

In principle, the data about transit flows should be determined as the item that remains from the linking of transport data and national foreign trade data. Such an analysis was carried out in 1990-92 for ferry freight between Northern Jutland and Sweden, at a time when the INTRASTAT had not yet been introduced. A similar analysis has not been attempted for ferry freight across the Baltic Sea between Denmark/Sweden and Northern Germany.

Sea transport statistics and foreign trade data, that are provided to the EUROSTAT from other EU Member States, might be considered as sources for further analyses on this topic.

REFERENCES

Cruise Ferry Info, various issues

Danmarks Statistik, 1999: Vejviser i statistikken (*Guide to the statistics. Statistics Denmark*. Copenhagen September 1999). This guide provides an excellent overview and short presentations of the various statistical sources included in the national Danish statistics.

Danmarks Statistik (*Statistics Denmark*): 'Statistiske Efterretninger. Transport':
- Jernbanenet og transport med tog (*Railway transport*)
- Vejnet og –trafik (*Road network and traffic*)
- Godstransport med danske lastbiler (*Goods transport by Danish trucks*)
- Skibsfarten på danske havne (*Maritime traffic via Danish ports*)
- Passager- og færgefart på danske havne (*Passenger ship and ferry traffic via Danish ports*)

EUROSTAT, 1999: Methods and Nomenclatures (1999 Edition. EC, Luxembourg 1999)

EUROSTAT, 2000: NST-2000. An outline proposal from EUROSTAT for the replacement of NST/R (EUROSTAT C.2 version 1.0, dated 14/01/2000)

Forskningsministeriet, 2000: National delstrategi for dansk transportforskning (*National Strategy for Transport Research*. The Danish Ministry of Research, Copenhagen May 2000)

FTC (Fehmarn Belt Traffic Consortium), 1996: Fehmarn Belt Traffic Demand Study (Progress Report No. 4, Surveys, September 1996)

Lillehammer, Giske and Hovi, Inger Beate, 2000: Freight Transport in the Expedite National models (Institute of Transport Economics. Oslo July 2000)

Henriques, M, 1998: Transportkæders miljø og omkostninger (*Environmental impact and costs of transport chains*. Trafikdage på Aalborg Universitet, August 1998)

Hensher, David A. and Button, Kenneth J., 2000: Handbook of Transport Modelling (Pergamon, Elsevier Science Ltd., UK 2000)

Holm, Jan, 1997: Godsmodeller – en oversigt (Vejdirektoratet, Copenhagen March 1997)

INFOSTAT, 1997: Final report – Task 1. Basic concepts and standards for a European transport information system (NEA, the Netherlands October 1997)

INFOSTAT, 1998: Final report for publication (INFOSTAT Consortium/NEA, the Netherlands October 1998)

Kveiborg, O., 2000: En model for godstransportens udvikling (includes an English summary: *A model describing the development of freight transport*. National Environmental Research Institute, Denmark 2000)

Københavns Havn, 2001: Drogden feasibility studie. Aktivitet 3.7 og 3.8: Trafikprognose (internt notat, Copenhagen Oktober 2001)

Langeland, Anders, 2000: NTN – Delprosjekt 2 Hovedrapport, Anneks. Regionale transportdata for planlegging - Godsstrømme fordelt på vareslag, transportmåte og OD (*Regional transport planning data. Goods flows and their distribution on commodity groups, transport modes and OD*. Agder Research Foundation and Nordic Transport Policy Network, reporting of sub-project No. 2 June 2000)

Madslien, Anne et al, 2000: Modellverktøy for transporter i norsk utenrikshandel (includes an English Summary: *Model for transports in Norvegian foreign trade*. TØI rapport 480/2000, Oslo May 2000)

McKinnon, Alan et al.: Logistical restructuring and road freight traffic growth: An empirical assessment, Transportation, vol. 23 no 2, May 1996.

MESUDEMO, 2000: – WPs 2 and 3. Identifying data on goods and passenger flows (The MESUDEMO Consortium, July 2000)

NEA, 1998: Spatial Patterns of Transportation. Atlas of Freight Transport in Europe (Ministry of Housing, Spatial Planning and the Environment, the Netherlands 1998)

Ortuzar, J. and Willumsen, L.G., 1994: Modelling Transport (Second Edition. John Wiley & Sons, 1994)

Ortuzar, J. and Willumsen, L.G., 2001: Modelling Transport (Third Edition. John Wiley & Sons, 2001)

Petersen, M. S., 1995: Potentialet for Søtransport (*Potential of sea transport*. Transportrådet og Trafikministeriet, Copenhagen May 1995)

REDEFINE, 1999: REDEFINE Summary Report (REDEFINE Consortium, 15 February 1999. CORDIS web site download)

SAMGODS, 1999: Efterfrågemodeller, Estimering av varuvärden, (*Demand models - estimation of goods flow values*. Temaplan, Sweden 1999)

SAMGODS, 2001: Efterfrågematriser (*Demand matrices*. Underlagsrapport til SAMPLAN 2001:1)

SAMPLAN, 2001: The Swedish National Model System for Goods Transport SAMGODS – a brief introductory overview (SAMPLAN Report 2001:1)

SCANDINET, 1999: Summary Deliverable D8 (Internet download)

ShipPax, 2000: Statistics 00. The Yearbook for passenger Shipping Traffic Figures

SIKA, 2001: Stråkanalyser för godstransporter (*Corridor analyses for goods transport*. SIKA rapport 2001: 1)

SoftIce, 1998: Survey on Freight Transport Including Cost Comparison for Europe, State of the art and freight cost database (November 1998, Internet download)

Statistics Denmark, 1999: Key figures for transport (Danish Ministry of Transport, Copenhagen 1999)

Trafikministeriet, 2000: Godstransporterhvervet i international konkurrence (*International competitiveness of the goods transport industry*. Danish Ministry of Transport, Copenhagen October 2000)

TRILOG, 1999: TRILOG-Europe Summary Report (TNO Inro, the Netherlands December 1999)

PREFACE

This report is volume two from a study of data requirements and methodologies for goods transport modelling. The study is intended to provide the basis for a general discussion about the application of goods transport models in Denmark. Volume one provides an overview of different types of models and data availability. Volume two discusses in detail the problems associated with data collection for discrete choice models.

The writers are responsible for the conclusions and recommendations in this publication, which are not necessarily shared by the Danish Transport Council.

The Danish Transport Council, February 2002.

Table of contents

1 SUMMARY AND CONCLUSION

1.1 DISCRETE CHOICE MODELS

Discrete choice models have been in use since the beginning of the 1980s. They are probabilistic models where the probability of an individual person choosing a given option is a function of his/here socio-economic characteristics and the relative attractiveness of the option (Ortuzar, J. and L.G. Willumsen, 1994/2001).

Some useful properties of discrete choice models are indicated below.

Discrete choice models are based on theories of individual behaviour. Thus it is likely that these models remain stable in time and space.

Discrete choice models are estimated by using individual (freight company) data. This has got the following two implications:

- The available information in discrete models is more accurate compared to aggregate models, because aggregate models apply average information for hundreds of individuals.
- Discrete models are less likely to suffer from biases due to correlation between aggregate units. When aggregating information, the individual behaviour may be hidden by unidentified characteristics associated to the zones, and this could result in biases being introduced in the aggregate models.

Discrete choice models are probabilistic, i.e. they yield the probability of choosing an alternative from the choice set and do not indicate a priori which one that is selected. This is a more accurate approach to modelling choice behaviour than the approach assuming complete consistency in the way people perceive and express their preferences.

The explanatory variables that are included in discrete choice models are multiplied by their specific coefficients in the utility functions. The number of variables can be fairly large in these models, including some very specific policy variables. The coefficients have a direct marginal utility interpretation, e.g. in discrete choice models the value of travel time is calculated as a ratio of time and cost coefficients.

Discrete choice models have improved the travel demand modelling due to the following innovations:

- The importance of individual behaviour in travel demand modelling has been recognised.
- Compared to the aggregate models, the disaggregate models is able to capture more information from the data. Furthermore, disaggregated data is not a subject to significant statistical biases, which may occur in aggregated data.
- New theories for travel demand modelling, such as the theory of *random utility*, have been developed.
- The application of new types of data, such as Stated Preference (SP) data, has been developed.

The theoretical background of the discrete choice models seems to be sound. Establishing the functional relationship between the respondent's preferences and the attributes of model travel alternatives (e.g. travel cost and travel time) in the form of a *utility function* has found no critics in the literature. The *logit models*, which are the mostly applied types of discrete choice models, seem to cover the whole variety of needs from these types of models.

1.2 DATA PROBLEMS RELATED TO DISCRETE CHOICE MODELS

Discrete choice models are data dependent. The more segments that are included in a model, the more observations are needed. Introduction of the SP data has decreased the magnitude of this issue. The introduction of SP data improves two more aspects in discrete choice models. These are 1) the accuracy of the achieved estimations, as there is a problem of correlation of variables among Revealed Preference (RP) data; and 2) the possibility of a better estimation of the attractiveness of new modes/routes, i.e. the estimation of alternative specific constants.

The hardware and software technology has successfully followed the development of the discrete choice models since the beginning of the 1980s, allowing practitioners to build more and more sophisticated models with time.

Some of the problems related to the development of the discrete choice goods models are the following:

- The quality and the amount of available data play an important role in discrete choice models. For a number of reasons, many important observations are difficult to obtain from respondents. This applies first of all to transport costs, which might be a company's commercial secret or might not be known (in the cases where transport is part of the production process of the company in question). The uncertainty of goods forecasts is therefore connected to the uncertainty of measuring the importance of the cost parameter.

- It is shown that rating and ranking exercises (RP-exercises) produce different results compared to the SP-experiments. Particularly the importance of transport costs seems to be strongly increased in the SP-experiments compared to ratings and rankings. One reason is that the SP-experiments do not reflect, in a true manner, the way that the companies comprehend some of the important parameters in ratings and rankings. Therefore, there is a major need to reconsider the way SP-experiments are carried out, particularly concerning parameters with a content that is not straight-forward.

- It has also been shown that the measurement of 'soft variables' (e.g. 'information system' and 'customer service') is rather difficult in discrete choice models. Such variables are of no use unless they are decomposed and included in great detail in the SP-experiments.

- In many cases the interviewed companies have never applied more than one transport mode, despite that the type of shipment they work with could be an inter-modal transport. The presented mode or route alternative in a SP-experiment needs to be based on the calculated values for transport costs etc. for the particular shipment and the zone-pair. Goods assignment models are usually more rigid than passenger assignment models, producing therefore less accurate files of level of service (LOS). Therefore there is a need to improve the available files of LOS. Moreover, traffic planning needs to take into account the regulations (e.g. regulations for working and resting time), empty trips and possible discounts that the respondent might obtain for example from ferry operators. Only if the hypothetical situations in SP-experiments are realistic, they could lead to a trade-off between the offered alternatives, which is a necessary requirement for a successful survey.

- Goods transport matrices might change fast because goods transport is a function of the demand for different commodities in the market, companies' degree of competitiveness, changes in technology, and price changes. Therefore there is a need for a permanent updating of goods OD-matrices.

- There is an obvious need for more goods data of the disaggregated form. This is particularly related to the revealed preference (RP) data. In the ideal situation where a large number of goods RP data is available in the form of a goods flow analysis, e.g. the continuous analysis on goods flows carried out by the Swedish Statistical bureau, it would certainly help the forecasting of the discrete choice models. It should be noted that the RP data constitute the core for model forecasts. The problem of the required amount of data is also reflected in the issue of segmenting the goods models. Depending on the specific purpose of the project, it is proven desirable to build models for two or more segments according to e.g. commodity groups. In the case of lack of RP/SP data, the uncertainty of the model's accuracy (i.e. the quality of the obtained estimates) could be so high that the forecasts achieved are unreliable.

1.3 APPLICATION OF DISCRETE CHOICE MODELS IN DENMARK

Three Danish projects have been examined concerning the application of discrete RP/SP goods models. The overall conclusions of these studies are:

- Evaluation of different quality parameters per transport mode and goods operator (transport producer or transport buyer). Applying different techniques (rating, ranking and choice exercises) could give a good overall picture of the quality parameters that are most worth investing in, in order for the companies to obtain the highest future profits.
- Single mode choice goods models, where no new alternatives are included: The potentials of a certain transport mode at the general level can be successfully examined by the application of discrete choice models. Based on this examination, a calculation can be made - within mode and across mode - of the elasticity or sensitiveness of the different parameters (transport cost and time).
- Modal choice component of goods models, where new alternatives are included: Apart from the modal split and route choice, discrete goods models could be applied in order to determine the potential use of new fixed links (e.g. across the Fehmarn Belt and Øresund).

The possible users of discrete choice goods models are both planning bodies (e.g. transport ministries, directorates), transport producers (carriers) and transport buyers (shippers). It is therefore an imperative to establish a platform where data can be obtained at a broad level, in order to facilitate the future goods transport planning in Denmark.

From a theoretical point of view, there is a need for a permanent theoretical and, more importantly, methodological improvement of the discrete choice goods models. The 'probit' type of models is still seldom applied in practice due to the lack of appropriate software packages. Improvements in the surveying methods, improvements in the SP-experiments, and improvements in the files of level of service (i.e. better assignment models) are to be included from this point of view.

2 INTRODUCTION

This report is prepared as part of the study on Goods Transport Modelling supported by The Danish Transport Council through the research programme Transforsk99. Two reports have been elaborated in the project, and the present report discusses in detail problems related to data collection for discrete choice models. The main conclusions to be drawn from this report are presented in the summary and overview in the other report from the study: Goods Transport Modelling I; Data and Methodologies.

Goods transport demand models have been a truthful shadow of passenger travel demand models for decades. The theoretical development of travel models has been till now focused mostly on peoples' need to reach the place of destination, where a certain activity is to be completed. These models were then tested on the data related to personal trips (predominantly on personal car trips) in order to measure how good the models describe the reality. We state here that most goods transport models rely largely on the modelling methodologies of the passenger travel demand models ignoring therefore many substantial differences between the two model types.

Discrete choice models or the second generation models have been in use since the Third World Conference in Travel Behaviour in Australia in 1977. These models are based on individual data (e.g., a producer of certain types of goods or a producer of freight services) which are related to the background information (so called socio-economic data) and a specific trip. In goods transport planning socio-economic data relates to the company's turnover, number of employees, location, type of production or service and so on. A trip can be described through travel costs, travel times of different nature (e.g., in-vehicle time, access-egress time, delay), frequency of travelling (or headway, or hidden waiting time) and the quality of travelling related to a possibility of shipment's damage.

This note aims to point out the strong and weak sides of the freight discrete mode choice models. It will be fundamentally discussed if the discrete choice modelling can efficiently be applied in the goods models, and if yes what the most important points to take care of are. The note will focus on the long distance freight models, i.e. the Danish international freight transport. The reason for that is that on the short distances lorry transport is overwhelmingly dominating over the other transport modes. Therefore, for short distances the modelling effort primarily relates to the establishment of number of lorry trips, their trip-end distribution and assignment in the road network.

In Appendix 1 the basic theoretical facts about the discrete choice models are described. The remainder of this chapter focuses on desirable and weak sides of these types of models.

3 DATA NEEDS FOR DISCRETE CHOICE MODELS FOR GOODS TRANSPORT

3.1 HETEROGENEITY OF FREIGHT MARKET

All freight trips are made for the sake of transporting the particular shipment to the place of consumption. The utility functions of the mode/route alternatives in the goods discrete choice models are therefore derived or indirect utilities of these alternatives. The discrete choice models work under the assumption that the respondent has a, more or less, perfect knowledge of what the available options are when travelling between two points, and that he chooses the one with the highest utility value.

Goods can obviously not decide what alternative should be applied for the trip between two points. A person or a group of persons decide on goods transport. The first doubt one should have about the application of discrete choice models in goods transport is therefore, if the decision maker has a perfect knowledge of what the available alternatives for a particular transport are, and whether he is able to choose the alternative with the highest utility-value.

Most realistically, the above assumption cannot hold for all cases. The explanation to that should be found in a large heterogeneity of the goods transport market. This heterogeneity in goods transport is related to the shipment size, commodity groups and the distance between the points of origin and destination. Another possibility is that the shipper has a perfect knowledge about the best choice, but that this information is not revealed to the interviewer/transport planner, because of insufficient description of the selected transport or failing precision in the interview design. If this is the case the development of the model will be erroneous.

Shipment size ranges between a small package and a full-loaded oil tanker. The size can therefore relate to the weight, volume and length. There are also shipments in gas or liquid form.

A number of commodity classifications are in use nowadays. Applying the commodity classification in a goods discrete choice models would imply that VOT are, for instance, different for each commodity group or between a number of commodities grouped together. However, one commodity group represents usually a conglomerate of sub-commodities, which again can differ in nature, as for instance:

- Price per kg,
- Perishability,
- Time sensitivity,
- Dangerous goods,
- Sensitivity to damage.

Each time a commodity group has characteristics which separates it from other commodity group characteristics it should be considered to build a separate model-segment for it. (Otherwise, the methodological advantage of the discrete choice models over the aggregate models would diminish). That is a problem in itself, because the goods data is seldom available at such disaggregate level. We will come back to the problem of the available data in discrete choice models later in the chapter.

Finally, length of transport can range between a couple of kilometres and a couple of thousands of km, the latter regarding the European international transport. It has been broadly accepted that the minimum threshold in the length, for which a mode choice will occur, is 300 km. On shorter distances the lorry transport is dominating.

Regarding the shipment size and commodity groups there will be situations where realistically only one mode can be applied. The appropriate discrete models must take into consideration these situations implying that for a specific zone to zone combination only selected alternatives are available.

3.2 WHO TO INTERVIEW?

A multitude of potentially active decision-makers is involved at different levels of goods transport planning. These are shippers, freight forwarders, carriers and receivers. In a broader perspective manufacturers and consumers can also decide upon planning the goods transport. It is possible to imagine that Danish consumers of ecological products will not accept that Spanish cucumbers are transported by air and they therefore have an impact on mode choice.

Only the transported shipment itself cannot be involved in planning of the transport. This trivial fact makes actually the greatest difference between the goods transport and the passenger transport.

Taking previous into consideration it has to be decided before hand which person(s) should be interviewed in the name of the shipment. That is obviously not straightforward because all the previously listed decision-makers will have their own view of what is to be prioritised in the choice situation, i.e. a bias is likely to be introduced in the sample. It is believed that no single person possess all necessary knowledge on the transportation chain.

In general it could be considered that the shipper wants a safe transport which reaches its destination at the prescribed time at the lowest possible cost. Therefore the shipper is not particularly interested in choice of mode or route, rather in some characteristics of the transport which can be fulfilled by one or more transport means or combination of transport means. The forwarder is interesting in finding the transport mode, which fulfils the shipper's demand. Therefore, the forwarder is actually making a choice of mode. The forwarder may be an internal function in a major company, or it may be an independent company servicing a number of different shippers, and possibly collaborating with a transport centre thus making it possible to consolidate goods flows to obtain lower prices at certain legs of the transport.

The transport company who is actually carrying out the transport will select the most appropriate route taking into account time tables, fleet utilisation, resting time for drivers, etc. The bigger the company, the more ways will be available for carrying out the transport in the most efficient manner.

3.3 AVAILABILITY OF GOODS DATA

Discrete choice models are data dependent. This dependency relates both to the quality and the amount of data. The freight market is a competitive market where companies make internal agreements across different levels of co-operation. One of the most important variables in discrete choice models is the cost of transport. Larger producers usually negotiate short and long term agreements with certain carriers. Further, carriers might have their own agreements with say ferry operators. Thus, the freight rates and tariffs are in general not publicly known. In the case of goods transport for own account, where the transport of goods is a part of the company's production costs, the transport costs can more easily be assessed based on the different cost components in the transport chain.

Transport time is a concept with different meanings in relation to different interview persons. A shipper may look at transport time as the lead time, i.e. the unbroken time from a piece of goods leave the factory to the same goods is received at the point of destination. Usually the delivery time is an important parameter. A carrier may look at transport time as being composed of a number of time elements related to different transport modes or different incidents along the route, e.g. border crossing or waiting time for ferries. All time elements are of a stochastic nature and therefore the precise time related to the components may not being known to the respondent.

'Soft' variables, e.g. the level of information system and customer service, are often difficult to achieve in the interviews. One reason is that the description of information system or customer service needs to be broken down in a number of elements which can be interpreted in terms of continuous or nearly continuous functions. Such a break down has not been performed in any of the studies carried out in the Scandinavian countries.

Some respondents are so used to apply the same type of transport (e.g., mode and route) that they have no knowledge of alternative types of transport in terms of transport costs and travel time. Discrete choice models are commonly based, as it will be presented later in the note, on choice situations between two alternatives. If one of them is poorly presented in the choice context one risks to obtaining poor model results.

A very general problem is that discrete choice models for goods transport have to be based on interviews with companies dispatching goods. Usually there will be only a limited number of companies available for interviewing. These companies may on the other hand dispatch big amounts of goods. In order to keep interviewing time at a reasonable level it is not possible to carry out a number of different interviews with one company about a number of different dispatches. Usually only one transport will be described in detail, therefore limiting the statistical base for estimating models.

The missing information, which are caused by the missing knowledge or by the non availability of information, causes problems in designing a good questionnaire and therefore completion of reliable data. Obtaining too few data causes problems in the estimation phase.

3.4 TRIP MATRICES

Goods transport is a production input factor of which the companies may demand more or less depending on the changes in the market. Goods transport is a function of the demand for different commodities in the market, companies' degree of competitiveness, changes in technology and price changes. The commodity flows and origin-destination matrices are therefore more sensitive to possible changes over time than passenger trip matrices. Therefore, the goods trip forecasts have a higher degree of uncertainty than passenger trip demand forecasts.

3.5 DEFINITION OF THE SAMPLE IN GOODS MODELS

Samples in goods transport analysis should be designed carefully. Care should be taken of both the nature and the obtainable size of sample in the analysis. It should be clarified whether the sample should be based on commodities in tonnes, commodities in shipments, tons irrespective of commodities, tons per km, number of transports, the nature of companies (producers, shippers, carriers, receivers), modes of transport or some combination of the above.

Definition of samples requires a clarification of the purpose of the goods transport analysis. There is a major difference concerning specification of a transport if this is done by a shipper, a forwarder or a carrier. Therefore, selection of samples must be carried out in the light of which questions to analyse, and whether this analysis is best accommodated by use of interviews with shippers, forwarders or carriers.

Experience has shown that comparable transports in terms of origin and destination, type of good, etc. show major variations in terms of transport time and transport costs. A certain variation is expectable due to quantity, conditions of transport and conditions of delivery. But if these aspects have a major influence they should obviously be part of the analysis in order to define sensible samples.

A common problem caused by the poorly defined samples in the goods models is that it is often difficult to complete enough interviews per model segment. That causes problems in the estimation of the discrete choice models where often not significant estimates are achieved per segment. A usual procedure in these cases is to merge (join) a number of segments with poor estimates to one segment. That is critical because the very reason for building discrete models, relative to aggregate models, is to establish different forecasts for different goods segments.

4 METHODOLOGICAL CONSIDERATIONS WITH SP DATA

This chapter will focus on the theoretical and practical considerations in work with SP data in the goods discrete choice models.

4.1 IMPORTANCE OF SP DATA IN DISCRETE CHOICE MODELS

Discrete choice models are based on a combination of RP and SP data. In multinomial logit (MNL) models and nested logit (NL) models SP data is usually applied due to the following reasons:

- Survey costs in discrete choice models would be far too big if only RP data are to be completed,
- SP data improves the accuracy of the estimates, and
- SP data helps the estimation of the alternative specific constants of the presently non-existing alternatives, as for example the *'lorry alternative over the fixed-link'* in the Fehmarn belt project.
- SP data is very useful when analysing the importance of new policies such as road pricing.

The RP data finds its role primarily in the model calibration and forecasts.

4.2 SP QUESTIONNAIRE

Respondents in Stated Preference surveys are companies' logistic managers or persons in charge of transport planning. Respondents are interviewed with the help of personal computers, where the questionnaire including variables is usually pre-defined. The German consultant, BVU, has developed a questionnaire in which the companies select the most appropriate variables in the context of the transport priorities of the companies. Costs are compulsory. Further, there are 16 different qualitative and quantitative variables of which 5 can be picked for inclusion in the interviews. This way of conducting the interviews assumes a big amount of interviews. 800 interviews were carried out using this methodology for collecting data for estimation of modal split models for the German Infrastructure Planning (BVWP) Model.

The strongest side of the SP data is that it originates from an interview that is entirely controlled by the researcher. How that is done is described below.

SP respondents in goods models are selected in advance as the companies/respondents possess desirable characteristic(s) for the project purposes. The questionnaire is designed for the interview to being completed in phases. In the first phase of the interview the broad characteristics of the company in question are revealed. This information relates to the:

- Name and size of the company,
- Type of company: producer of goods, freight forwarder, carrier or receiver,
- Company's turnover,
- Possession of own car-park,
- Selection and description of the commodities the company works with,
- Company's infrastructure possibility (e.g., existence of industrial track, own port, access to transport terminals), and
- Company's policy towards application of different modes of transport.

In the second phase of the interview a quality analysis is carried out. The quality analysis is divided between the rating and ranking exercises. In the rating exercises the respondent is asked to rate the importance of selected variables in a certain scale (usually the scale is defined in 5 or 10 points, ranging from *'not important'* to *'extremely important'*) for the most commonly applied mode of transport. The selected variables are commonly transport costs, transport time, reliability, frequency of transport, risk of shipment being damaged at the destination, customer service, information system and flexibility, but others could be included, as mentioned above. BVU worked with ratings of 17 different variables.

In the ranking exercise the respondents are asked to rank the chosen variables in order of importance based on their own perception of the most general transport for their companies.

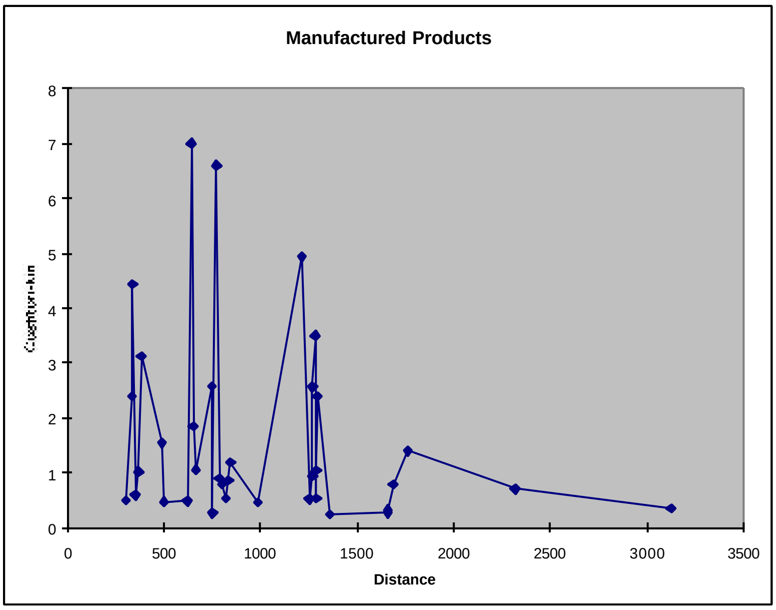
In the third phase of the interview the respondent is asked to describe a recently completed trip that is common for his company. The place of origin and destination is stated together with the mode of transport and values for travel costs, travel time and other variables, which are decided to be included in the SP experiments.

If one or more alternative modes (or routes) could be applied for the trip in question, it is highly recommended to complete a set of similar values for that transport mode(s) from the respondent. Alternatively, this information can be sourced from the background data (so called files of level of service (LOS)). The experience shows, however, that many respondents have troubles in accepting the computed values either because the files of LOS are poor or because their experience is rather different from the computed values. This underlines the need to be able to producing credible LOS data. A discussion of this aspect is quite important, because LOS data are not only used in the interviews but are also an important part of the developed mode choice and route choice models. If therefore the LOS data used for the interviews are not credible in the view of the interviewee the model based on the LOS will not be credible either.

Definition of LOS relies on definition of a proper network for the different transport modes. Based on the network data it should be possible to calculate the required variables (travel time, travel costs, other quantitative variables). LOS for qualitative variables are by definition not computable based on network data.

Figure 4.1 shows how cost per ton-km varies with distance for manufactured products. The results have been established in a survey carried out in Denmark 1994. The results stress that it would not be possible to establish a simple relationship between transport distance and transport cost for manufactured products. Other variables should be included as well, e.g. whether the goods has a high or low volume/weight ratio, whether goods is shipped as part container load or full container load or whether the goods has specific features which require temperature control. This also stresses the need for establishing not only very detailed network data, but also other types of data to be included in LOS calculation formulas.

Figure 4.1: Relationship between cost per tonkm and distance for manufactured products as established in a Danish survey.



In the Fehmarn Belt analysis discrepancies were found between the LOS as calculated by the network models and the time and costs as stated by the companies. It was found that computed costs by road were 20% higher than observed for haulier transport, and 80% higher for own account. For rail the computed costs were 60% higher than the observed. Computed time was found to be 25% - 30% lower than observed. These results underline the difficulties in establishing relevant and precise LOS-data.

In the final phase, the respondents are involved in the SP experiments that are described in details in the following section.

Applying computer-based interviews achieve the following advantages over the traditional interviews:

- The automatic presentation of the questions relevant to the respondent,
- The possibility of computing the background data (data of level of services) along the interview,
- The application of the pre-defined SP design that ensures the independent variation of the variables in the SP experiments (i.e., orthogonality), and
- Automatic coding, storage and back-up of response data.

When each interview is finished all the answers (data) are stored in an output file that has a unique identification code. Joining all the obtained data from the sample takes usually several minutes. In that way human errors connected to coding of data are eliminated.

4.3 SP EXPERIMENTS

In the SP experiments the respondents are involved in binary choices between alternatives that can represent the same alternative (e.g., lorry alternative 1 vs. lorry alternative 2) or that can represent two different alternatives (e.g., lorry alternative vs. train alternative). The first type of experiments is called *'with-in mode experiments'* while the other type is called *'across-mode experiments'*.

The alternatives are described through a number of variables, which vary in values because the SP experiments are hypothetical experiments. Both with-in mode and across mode experiments need to be contextual. That means that the alternatives are related to the originally described alternative and that they are based on the information sourced from the respondent himself, at least regarding the original alternative. Contextual SP experiments ensure that the respondent can relate himself to the presented alternatives, i.e. the SP experiments are credible.

With-in mode SP experiments are easier to understand than the across-mode SP experiments. It is therefore recommended to begin this phase of the interview with with-in mode SP experiments. In the modelling context the data from with-in mode SP experiments are used for estimating statistically more accurate estimates of the presented variables than is possible based on RP data only, i.e. the correct values of time (VOT).

In across-mode experiments the original mode of transport is presented against the alternative mode. The alternative mode can be an existing one or a presently not-existing mode (e.g., train alternative over the Fehmarn belt fixed-link). In the modelling context the data from across-mode SP experiments are used for assessing the statistically more accurate estimates of the presented variables and for estimating the attractiveness of the new modes (i.e., the estimation of the alternative specific constants for the new modes). The biggest problem related to the across-mode SP experiments is the problem of trade-off. If the alternatives in the across-mode experiments are not presented realistically (especially the alternative mode) the respondent will tend to stick to his original choice throughout the experiment. If across the whole experiment only one alternative has been chosen no new information are added in the data file and no improvements in the models estimation are achieved. It is therefore critical to obtain realistic starting values for the alternative mode as discussed above concerning the LOS.

It is necessary to make a balance between the number of SP experiments per interview, the number of SP questions per experiment and the number of interviews (i.e., the sample size). The sample size should be related to the model segmentation in the way that for each segment minimum 35 to 50 respondents are available in the sample. The number of SP experiments presented to the respondent in the interview can range between 1 and 3. It is not recommended to involve the respondent in more than 3 experiments because the quality of obtained data decrease with the increased interviewing time.

One SP experiment consists of a number of SP choice situations (questions). They are pre-defined by the researcher, which ensures the independent variation of the applied variables across the experiment. It is recommended that the number of questions in an experiment vary between 6 and 9.

4.3.1 Variables in SP experiments

The variables to be presented in SP experiments are usually decided beforehand. It is however, possible to build SP experiments where user-chosen variables are implemented in the SP experiments.

It is necessary to consider two aspects in the design of questionnaires regarding this point. These are the choice of variables and the number of variables in the experiment. If the number of variables exceed four it is a good idea to define two SP experiments combining the variables in such a way that they can be linked by a common variable (that can be, for instance, transport costs). It can be argued that the variables that can be presented in a number of ways, as for example, *'customer service'*, *'information system'* and *'flexibility'* should be omitted from the choice experiments. These variable are often called 'soft variables', because it has been proved through a number of analyses in the past that each respondent will tend to have his own definition of these variables. This causes a large spread of the importance that respondents in the sample attach to these variables. Under such circumstances poor estimates are achieved.

In some SP experiment the interviewee will tend to simplify the choices focusing on one or a few variables and omit those that are less important to him. This problem is known as a *'lexicographic problem'* because the respondent goes through the SP experiment without taking into consideration the whole complexity of the SP experiments.

There is a number of ways to deal with this problem:

- The SP experiments should be customised, which means that the starting values for variables included in the experiments are given by the respondent,
- SP design should be of that kind, where each new choice situation is an improved situation from previous choice situations. That is called an 'adoptive SP design', and
- The order of appearance of variables changes from question to question.

4.3.2 Levels of variables in SP experiments

In order to obtain orthogonality in the SP surveys it is necessary to define beforehand what the levels of variation for each chosen variable are. Orthogonality ensures that variables vary independently of one another in SP experiments. Usually three levels per variable are applied in SP designs, e.g. the observed value (100%), +10% and –20% around the observed value. The total number of alternatives in an experiment equals the number of levels to the power of number of variables. For example, in an experiment with 2 variables that vary in 3 levels each and 2 variables that vary in 4 levels the total number of alternatives equal 3^2 x 4^2 = 144. This total number of alternatives is called 'full factorial design'. Full factorial design includes all combinations of the variables and their levels and it includes therefore all the levels of interaction between the alternatives.

It is practically impossible to present a SP experiment with 144 alternatives to the same respondent. Five solutions are found to this problem:

1. To apply what is called 'fractional factorial design'. This approach rests on the assumption that some or all of the interaction effects between variables in the SP design are negligible. In that case the opportunity then exists to modify the experimental design such that the number of alternatives are reduced. The fractional factorial design consists only of 9 orthogonal alternatives for a design with 4 variables into 3 levels each (the full factorial design consists of 3^4 = 81 alternatives).
2. To remove those alternatives (i.e., options) that will dominate or be dominated by all other alternatives in the SP design.
3. To separate the alternatives into blocks, so that the full factorial design is completed by groups of respondents, each responding to a different sub-set of alternatives.
4. To carry out a series of experiments with each respondent, offering different variables, but at least one variable in common (say transport costs) to all respondents, to enable comparisons.
5. To define variables in terms of differences between the alternatives (i.e., transport cost of combined road-rail transport equals 80% of lorry costs).

In most cases fractional factorial designs are applied.

4.4 PILOT SURVEYS

A successful SP survey relies heavily on pilot surveys. The purpose of their existence is to pinpoint the critical parts of the questionnaire. This relates to:

- The length of the questionnaire,
- Wording of questions,
- Branching throughout the questionnaire,
- Check of the SP experiments through the number of experiments, number of variables and levels, and
- Check the validity of the data of level of services.

Pilot sample consists of up to 10% of the number of interviews in the main survey. 5% is the bottom border. Pilot interviews may be completed in one or more phases depending on the difficulties that are faced in planning the questionnaire.

Data from the pilot survey should always be processed estimating a logit model. It is expected to achieve parameter estimates that, at least, have the correct sign. Level of correlation (between say time and cost variables) should be carefully checked in the output data.

It is desirable to carry out pilot surveys in the same geographical and organisational setting as would apply to the main survey. Pilot surveys that are based on these requirements help to create the necessary experience concerning organising and conducting the main SP survey. It is of great importance that the same persons are involved in both creating the questionnaire and in data analysis.

4.5 MAIN SURVEYS

Organising a main SP survey is a fairly complex task and it covers two main topics i.e. sampling and survey administration.

In freight surveys the sample represents a collection of freight companies (transport producers or transport buyers) that has been carefully selected in order to represent the population due to possession of certain attributes that are valuable for the project. The most commonly applied sampling methods are based on forms of stratified and random sampling.

Most freight SP surveys are nowadays completed in face-to-face interviews applying personal computers. The interviewers in face-to-face interviews help to organise the interview, guide the respondents through the interview (if necessary) and write the specific characteristics and comments concerning each respondent. Computer-based interviews are superior over paper-based interviews.

5 USE OF RP/SP DATA ILLUSTRATED THROUGH EXAMPLES

This chapter presents three different ways of applying the individual data collected through RP/SP surveys. The three examples are drawn from the experience in Denmark concerning RP/SP-surveys and a number of comments is provided concerning the possible improvements of the survey technique and further research in disaggregate modelling.

5.1 PROJECT 'GODSTRANSPORT OG KVALITET'

In Transportrådet's project 'Godstransport og kvalitet' (note 97-02) 131 SP interviews were completed with transport producers and shippers. This project is an example of how discrete choice models are applied in order to evaluate the importance of a number of different parameters describing the transport.

45 companies in the project were located on Zealand, 15 on Funen and the remaining 71 companies were located in Jutland. The investigated parameters in the project were transport time and transport costs, reliability, risk of shipment being damaged at the time of arrival, travel frequency (for rail and sea modes), flexibility, information system and customer service.

Two types of information were collected from the respondents through the computer-based interviews:

- The RP information included a general indication of importance of a number of quality parameters in the companies' present transport policies. For this purpose respondents were asked to *rate* every parameter separately in a scale of five points of impedance and also to *rank* the parameters as first, second and so on, when viewed in a group. Analysis of the RP information was separated for buyers and producers of transport services for road, rail and sea transport modes.
- As input to the SP experiments, each respondent described the most recently completed transport in terms of the above listed variables. SP type of information was collected from within-mode SP experiments. Depending on transport mode applied in the described freight journey respondents were involved in a number of SP questions where two transport alternatives of the same mode-type are presented through

values of the chosen quality variables. Three models were built in the project based on the completed SP data i.e., models for road, rail and sea transport modes. The models gave estimates for each of the investigated parameters.

5.1.1 Results of the analysis of the RP observations

Each respondent ranked the chosen variables in a scale of *'not important at all'* to *'extremely important'*, in a 5-point scale. In the sample of road transport users, 42 companies were understood as transport buyers and 24 companies were understood as producers of transport services. In Table 5.1 is shown the importance of the chosen quality parameters for these two segments respectively.

Table 5.1 – Importance of the quality parameters for road transport producers and buyers; rating exercise

Variable	Transport producers	Transport buyers
Transport time	Very important	very important
Reliability	Extremely / very important	extremely / very important
Risk of damage	Very important	very important
Transport cost	Very important	very important
Transport frequency	Very important	very important / important
Information system	Very important	very important
Flexibility	Very important	very important / important
Customer service	Extremely /very important	very important

Eight quality parameters were rated from both transport buyers and transport producers to be at least *'important'* parameters in their present transport policies. In most of the cases the parameters were rated as *'very important'*. It can also be noted that in most of the cases there is no difference in importance of the quality parameters seen from the transport buyers and producers point of view.

In the second part of the questionnaire the respondents were asked to rank the eight quality parameters in order of importance based on their own perception for the most general transport for their company. The results of this exercise for road transport producers and buyers are presented in Table 5.2. Parameters are grouped in 3 major groups: most important, secondly most important and thirdly most important parameters.

Table 5.2 Importance of the quality parameters for road transport producers and buyers; ranking exercise

	Transport producers	Transport buyers
Most important parameters	Reliability, Transport time, Transport costs	Reliability, Transport time, Transport costs
Secondly most important parameters	Customer service, Flexibility	Customer service, Flexibility, Risk of damage
Thirdly most important parameters	Transport frequency, Information system, Risk of damage	Transport frequency, Information system

The obtained data from the ranking exercise shows a very nice overlap between the transport producers and transport buyers. The ranking exercise shows that among eight quality parameters (that are all rated to be important in the company's present transport policy) three parameters have the highest importance. These are reliability, transport time and transport costs. Therefore, being more reliable, delivering goods faster and/or decreasing transport costs would most likely please the buyers of road transport most. On the other side, it is probably not worth improving much the present level of information system and transport frequency among lorry operators.

5.1.2 Discrete choice models based on SP observations

Each respondent in the sample described the most recently completed goods transport, which was longer than 300 km. Table 5.1 summarises the obtained data.

Table 5.1 Average values for the chosen variables from the RP data

Variable	Road Transport	Rail Transport	Sea Transport
Travel Distance (km)	840	1070	4350
Shipments Value (DKK/kg)	39	19	5
Shipments Weight (tons)	12	21	179
Travel Time (hours)	33	59	228
Driving Time (hours)	16	30	228
Average Driving Speed (km/hr)	58	36	19
Travel Cost (DKK)	6600	8300	33600
Risk of Damage (per mille)	7	5	5
Risk of Delay (per cent)	3	5	4

In the SP experiments respondents were involved in within mode experiments where values of the quality parameters varied around the observed values, which were given by the respondents. The obtained data were afterwards applied in the model estimation. One model was estimated per transport mode. The model estimates are presented in Table 5.2 in form of the monetary values for travel time, risk of damage, reliability and frequency.

Table 5.2 Monetary values of transport time, reliability, risk of damage and frequency, in DKK (1996 prices)

Parameters	Units	Model for road transport	Model for rail transport	Model for sea transport
Transport time	DKK/hr	66	26	18
Risk of damage	DKK/per mille of damage	16	208	483
Reliability	DKK/per cent of delay	41	78	535
Frequency	DKK/per weekly departure	---	105	546

Value of travel time is highest for road transport (66 DKK/hr). Rail transport's value of time is about 3 times lower than value of time for lorry transport while sea transport has the lowest value of transport time (18 DKK/hr). This is in accordance with the knowledge that more expensive goods (i.e., high value commodities, small shipments on shorter distances, etc.) are usually transported on roads while cheaper goods (i.e., low value commodities, large shipments on longer distances, etc.) are usually transported on rail and sea. In the model sample the smallest shipments (12 tons) with the highest value (39 DKK/kg) are transported on roads at the shortest distances (840 km). Opposite to that, the largest shipments (178 tons) with the lowest value (5 DKK/kg) are transported on sea at the longest distances (4356 km).

The value of risk of damage for road is lowest among the three modes. Opposite to that, the value of risk of damage for sea transport is highest. Average reported values for all the modes are very close (between 5 and 7 per mille). This indicates that travel distances and shipments' weights (see Table 5.1) are closely related to the calculated values of risk of damage.

Reported values of risk of delay (reliability) in the RP data are low and similar between the modes (between 3 and 5 per cent). However, the longer the travel distance the more prepared users are to pay for improvement of this parameter. This is why rail users are prepared to pay about 80% more for a decrease of 1%, in average, of risk of delay in rail transport compared to road transport. Average sea transport distance in the sample is more than 4 times longer than average rail transport distance and is more than 5 times longer than average road transport distance. In return risk of delay in sea transport is valued about 7 times more than in rail transport and more than 10 times than in road transport.

Values of transport frequency are calculated for rail and sea transport. Users of sea transport are willing to pay more for improvements of transport frequency based on the fact that fewer weekly departures are observed in the sample for sea transport than for rail transport.

5.1.3 Discussion of the project results

Two things are worth pointing out from the 'Godstransport og kvalitet' project in the present discussion. Firstly, there is a very good agreement between the producers and buyers of transport services regarding the valuation of importance of the eight quality parameters. This was proved to be correct both for ranking and rating exercises.

Secondly, the obtained values of transport time, risk of damage, reliability and transport frequency across three transport modes correspond well with the observed values for these variables.

The obtained discrete choice models in this project are built in order to evaluate the importance of the chosen set of variables. Nothing can be therefore said, based on these models, what the potential of say rail transport is, if one or more of the rail transport parameters improve in the future. For that reason additional set of RP/SP data needs to be completed and included in the model. These are data that describe how respondent would react in mode choice situations where transport service changes for one or more transport modes.

One of the observations concerning the result of the SP-experiments is a very different evaluation of parameters in the SP-experiments compared to the rating and rankings of parameters.

Based on the SP-experiments transport costs accounts for about 60-70% of the general transport costs, whereas reliability accounts for about 4%. This contradicts the ranking where reliability is considered the most important parameter.

This result indicates that in the SP-experiments cost and time most likely are defined in a way, which is comprehended of the interviewees. This is not the case with reliability. There are several explanations for that.

Due to difficulties in defining an expression for reliability, this term is replaced with chance of delay. The chance of delay is assessed based on company statements of observed delays, say 1% of shipments are delayed compared to agreed delivery times. But companies may have a different assessment of delays than of reliability.

Due to the structure of the SP-experiment it was not possible to trade a zero percent delay and a one percent delay against other parameters. This means that companies who would not accept delays, or companies where reliability was outmost important were excluded from the evaluation of the delay parameter. This may have a serious effect on the magnitude in cost terms of the delay parameter.

Finally, chance of delay is expressed with no qualifications. Are there different types of delay, i.e. critical delays, or less cost sensitive delays. And how are different ways of expressing delays linked with the original parameter reliability, if at all.

Another methodology, which was considered in the same project but without success was a calculation of expected transport time based on stochastic data about traffic loads, travel time, adherence to timetables etc. However, the amount of data necessary to carry out such a calculation could not be made available. It is necessary with complete timetables in case alternative routes should be considered in order to make the trip within the stipulated time window e.g. at late arrival at an intermediate terminal. As an overall planning tool the methodology was unable to provide the necessary simplicity, and therefore the method was given up.

One of the critical points in the model estimations is that the obtained monetary values (Table 5.2) refer jointly to transport producers and transport buyers. An attempt to separate these results in the estimations has failed due to lack of data. Secondly, the aggregation of commodity groups (it has only been assumed that high value commodities are usually transported by road) is hardly applicable for more detailed goods transport planning. Finally, the model fails in determining the values of ‘soft’ variables such as ‘information system’ and ‘customer service’.

‘Soft’ variables are difficult to include in hard core estimations, because they are usually difficult to describe in a way that allows inclusion in mathematical models. It is not sufficient to include “Availability of information system? yes/no”. Usually these ‘soft’ variables have non-significant t-values.

It would be necessary to disintegrate the term information system in a number of sub-variables. These sub-variables need to be discussed with the market players in order to develop the terms as precisely as possible, and subsequent to try to formulate the sub-variables in a meaningful way.

As an example: Information system could consist of a route information system and a fault information system.

Route information could include:

- No route information
- Route information at pre-selected times
- On-line route information (GPS)

Fault reports could include:

- No fault indication
- Response if fault, no information about fault
- Response if fault, information about fault
- Response if fault, information about fault, information about possible remedies

Other aspects could be included as well.

Another problem is to include these discrete variables in SP-experiments based on variations of a continuous nature.

5.2 PROJECT ‘MODE CHOICE MODEL FOR DANISH INTERNATIONAL TRANSPORT’

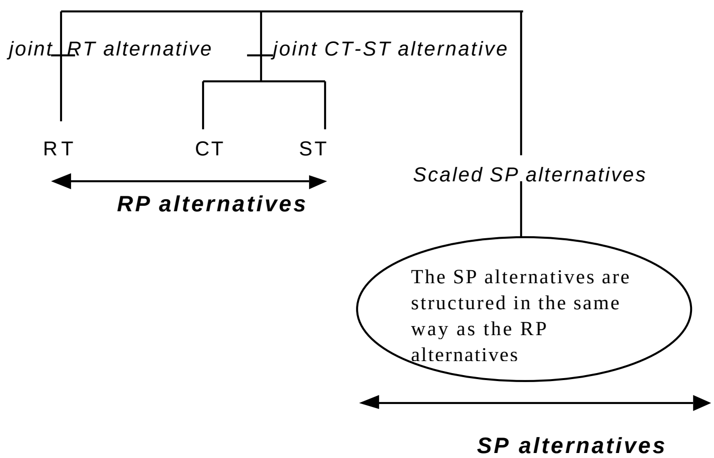
‘Mode choice model for Danish international transport’ is one of the TransForsk ’94 projects. The objective of the model was to forecast the future demands for rail and sea transports relative to lorry transport, when supply variables change.

The model can be described as a NL model based on the combination of RP and SP data. Transport modes included in the model structure are lorry transport (described further as Road Transport (RT)), conventional rail together with the combined road-rail transport (described further as Combined Transport (CT)), and sea transport (ST). Based on the experience from the previous goods transport projects in the country and abroad, CT and ST alternatives are grouped together vs. RT alternative, describing therefore a higher cross-elasticity between the first two modes in relation to RT. It was postulated therefore that the respondents in the sample distinguish first of all between RT and non-RT modes and after that among the non-RT modes they distinguish between CT and ST modes.

Due to the differences in the error standard variance between the RP and SP data, the SP data are finally scaled to the RP data. A simplified model structure is shown in *Figure 5.1*.

The structural parameter (? parameter from equation 11 in appendix 1) is present in a dummy *joint CT-ST alternative* to reflect the higher cross-elasticity of choosing between these two modes relative to the choice between the RT and non-RT alternatives. When this parameter is found to be significantly different from 1, the model structure presented in *Figure 5.1* is justified. It is important to note that the same parameter is also present for the RT alternative (i.e., dummy *joint RT alternative*) to ensure consistency in the scale of the utility functions.

Figure 5.1: The model structure



5.2.1 Model results

Data from three projects were combined in the model structure. These projects are 'Godstransport og kvalitet', 'The Nordic-Link' project and 'Potential of Sea Transport' project. All together, the data set consisted of 11.337 RP and SP observations. 1.012 were RP observations.

The mode choice model in the project consisted of two segments, which are low-value shipments (shipments with values op till 20 DKK/kg) and high-value shipments (shipments with values greater than 20 DKK/kg).

Monetary values of travel time, risk of damage, reliability and frequency variables for two segments are presented in Table 5.1.

Table 5.1 Monetary values for 4 parameters, in DKK (1995 prices)

Parameters	Units	Model for low value commodities	Model for high value commodities
Travel time	DKK/hr	14	46
Risk of damage	DKK/per mille of damage	69	126
Reliability	DKK/per cent of delay	95	139
Travel frequency	DKK/weekly departure	156	266

The value of the structural parameter ? in the model for low value commodities is equal to 1, which means that for this type of commodities respondents do not make a difference when choosing among the three modes. In the model for high value commodities, the ? parameter is 0.459, which indicates that the higher cross-elasticity of choosing between the CT and ST alternatives relative to the RT alternative. This result is to be expected because the high value commodities are usually transported by road transport.

The producers of high value commodities in the sample are willing to pay about 3 times more for saving one hour of transport time between origin and destination than the producers of low value commodities. Value of risk of damage for the high value commodities is about twice as high as the value for the low value commodities. Values of risk of delay and travel frequency are about 50 per cent and 70 per cent, respectively, higher for high value commodities than for low value commodities.

The model was thereafter calibrated in order to match the 1995 Danish export/import per transport mode, divided between low-value and high-value commodities.

Forecasts in the project consisted of model runs where one, or more, variables change values relative to the original values. The forecasting results for one of these scenarios, where the road transport cost was increased for 10% are shown in Figure 5.1.

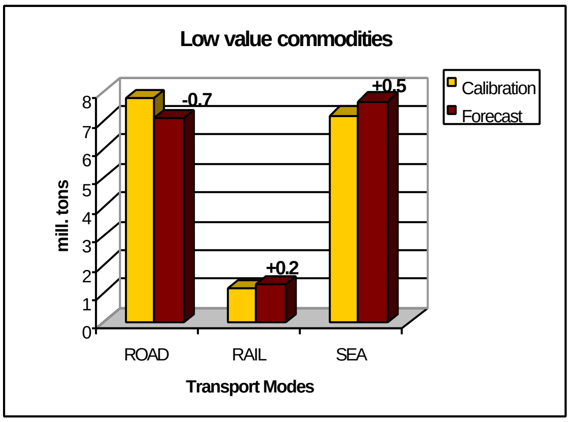
Low-value commodities are more sensitive to transport costs than high-value commodities. Therefore, 700.000 tons of Road Transport freight (about 7%) are transferred to Combined Transport and Sea Transport for low-value commodities in this scenario, while only 300.000 tons (about 3%) were transferred from Road Transport for high-value commodities.

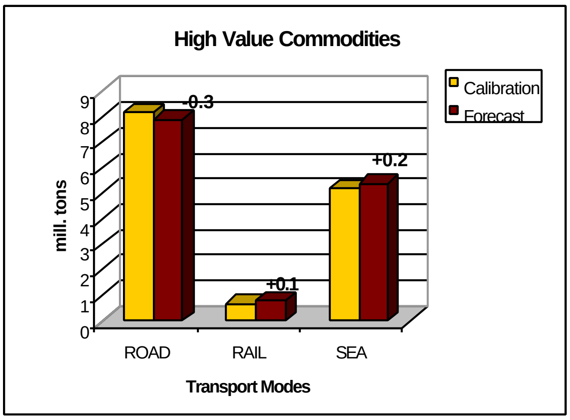
Making a number of runs with the model the following can be concluded about the obtained results:

- *Transport cost* has the far greatest importance for the mode choice in both model segments among 5 chosen variables. *Transport time* variable is less import than *transport cost*, but more important than *risk of damage*, *risk of delay* and *transport frequency*.

- The biggest share of the RT tons, according to the forecasting results, shifts to ST and the smallest share shifts to CT. This points out that ST is more competitive to RT than CT in the Danish export/import. This is in line the conclusions from a domestic goods project described by Henriques ('Transportkæders miljø og omkostninger', Trafikdage på Aalborg Universitetet 1998).

Figure 5.1 Forecasting results for the increased RT for 10%





5.2.2 Discussion of the project results

The model built in this project showed one of the strong sides of the discrete choice models, i.e. data from a number of projects can be successfully joined in a single model structure.

It is not possible to compare directly the obtained values of time and other variables from this model and from the model of 'Godstransport og kvalitet' due to different natures of the two projects. It can be however noticed, for example, that value of travel time by sea transport (14 DKK/hr) is very close to the value of time for low-value commodities (18 DKK/hr). It can be seen from the figure above that there is a relatively higher share of low value goods on sea transport and rail transport than on road transport. Therefore, a low value may be assumed for sea transport. However it is also evident that ship transport is used for big consignments over long distances. Therefore the value of time for ship transport may vary at another level than value of time for other transport modes.

Model forecasts show correctly that low-value commodities are much more sensitive to transport costs changes than high-value commodities. The same goes for other three variables (reliability, risk of damage and frequency).

One problem related to the project of ‘Mode choice model for Danish international transport’ is of special interest for the present discussion. The obtained data in the project is related to the shipments. Even shipments of the same commodity group could vary greatly in volume, weight, value per kg and so on. More than that, shipments of different commodity groups could differ greatly in say volumes. The model is calibrated on the amount of tons exported and imported from/to Denmark in 1995. The number of shipments (applied in model estimation) and the amount of tons transported (applied in model calibration and forecasts) do not necessarily supplement each other. It is difficult to imagine that SP experiments can be presented by anything else but shipments. The calibration data on the other side is seldom available in shipments. As a consequence to this in the forecasts it is shipments that change the transport mode (e.g., an old sea transport changes to rail transport due to improvements in rail transport time). Depending on the volume of the specific shipment more or less tons are shifted from one mode to another. Obviously many more lorry shipments need to change transport mode to achieve the same tonnage shift as a single rail or sea transport.

The model also fails in explaining the importance of variables such as ‘information system’ in the mode choice. Finally, the difference in preferences between transport producers and transport buyers is not included in the model structure.

Another problem inherent in the modal split model is the singular focus on supply variables for the different transport modes. Aspects like organisation of production and transports, and development of transport chains have not been considered. These aspects are also important for the modal split, since they define in which environment to consider the consignments i.e. the objects of the modal split analysis. Because these aspects are excluded the modal split model tend to become rather conservative.

A similar observation was made in the Nordic Link project. The modal split model forecasted a rather slim increase in rail transport if a modern combined transport terminal were established in Aalborg, Northern Jutland. However, a market analysis indicated a further increase of 150% if the combined transport services between Northern Jutland and Copenhagen/Hamburg were designed to fit the need of major companies. An important aspect, highlighted by the interviewed companies was the opening hours of the terminal and the arrival time in the other main terminals. In this sense a travel time by rail between Aalborg and Copenhagen is not just 8 hours, but 8 hours between 21 PM and 05 AM. This particular aspect takes into account the production process and integrates the transport in this process. The established model is not able to handle this aspect.

It should be pointed out, that results obtained in specific projects not necessarily reflects the results obtained in more overall studies. It is possible in the specific projects to direct questions and analysis towards the specific objective of the analysis. Focus on a specific objective may result in more dedicated answers and therefore also other valuations than obtained in general surveys. However, it still requires considerable considerations as to how SP-experiments are formulated and carried out.

5.3 MODE/ROUTE CHOICE MODEL ACROSS FEHMARN BELT

Discrete mode/route choice models were successfully applied in the projects where new infrastructure has been introduced. In the recent years in Denmark the Great Belt, the Øresund and the Fehmarn Belt traffic models were built applying RP and SP type of data. Demand for goods transport across the Fehmarn belt was based on 1.500 RP interviews conducted as road side interviews, 392 SP interviews and the supplementary RP/SP data from the Øresund traffic demand project.

RP interviews on the Fehmarn belt project were completed with lorry drivers in face-to-face paper-based interviews. In these interviews the following information were obtained:

- Vehicle type,
- Type of freight and the load,
- Origin and destination of the trip,
- Total transport time and waiting time, and
- Information of the transport components in the opposite direction.

392 computer-based interviews were completed with freight forwarders and freight carriers in Denmark, Sweden, Germany and The Netherlands. Beside general information about the company, three recently completed trips were described in details. These trips provided the fundament for the following five SP experiments, where in three experiments the respondents were offered an alternative across the Fehmarn fixed-link (i.e., an alternative with a shuttle train or two alternatives with a fixed-link route by road and rail).

Variables included in the SP experiments were:

- Origin-destination transport time,
- Origin-destination transport cost,

- Suitability of the offered capacity (for shippers) or suitability of the terminal (for carriers),
- Time reliability,
- Probability of damage of the shipment,
- Flexibility of handling of specific requests, and
- Information on deviations from the planned route.

Time, cost and suitability variables were always included in the SP experiments while the last four variables were included according to some predefined assumptions. Each experiment included five variables.

5.3.1 Discrete choice models based on RP/SP observations

The Fehmarn Belt mode/route choice model is a NL model. Choice between lorry, conventional rail and combined rail alternatives is on the first level, while the route choice for the lorry alternative via a number of ferry routes, including the fixed-link, is on the second level in the model structure. That structure implies that, for example, a lorry transport via Rødby-Puttgarden has a greater chance in the model to be shifted to another ferry route, or to the fixed-link, than being shifted to rail mode.

The RP part of the data applied in the mode/route choice consisted on 5.190 records, while the SP part of the data consisted on 5.593 records. Models for five segments, based on the aggregation of the commodity groups, were estimated. Table 5.1 shows the obtained values of time per segment in the model.

Table 5.1: Values of travel time per segment, in DKK/hr

Model segments	Value of time
Segment 1: Agricultural products	134
Segment 2: Wood, textiles, chemicals and paper	54
Segment 3: Miscellaneous articles	253
Segment 4: Manufactured products	180
Segment 5: Bulk products	55

A strong side of the Fehmarn belt mode/route choice model is that respondents' perception of the fixed-link (in different forms) was investigated. This means that with the help of hypothetical situations (i.e., SP data) the modellers were able to calculate what the potential of the new crossing is expected to be, isolated from the effects of other variables as for example, travel time and travel costs. This is done by estimating the values of alternative specific constants for the shuttle alternative and the bridge alternative, relative to Rødby-Puttgarden ferry alternative.

The forecasts in the project were made for year 2010 under 4 scenarios:

Ferry service on the Rødby-Puttgarden crossing (2010 reference sc.),
Scenario with a 2-track railway and a 4-lane motorway (sc. 2+4),

Scenario with a 2-track shuttle railway (sc. 2+0), and
Scenario with a 1-track railway and a 2-lane road (sc. 1+2).

For all 4 scenarios was assumed that the total tonnage transported in 2010 would not change due to the existence of the fixed-link. The mode/route choice model has however shown a great shift to the fixed-link from other routes in the 3 fixed-link scenarios. Table 5.2 shows the tonnage transported on the Rødby-Puttgarden crossing in 2010 for 4 scenarios by road and rail alternatives.

Table 5.2: Mode shares for 4 scenarios, in mill. tons

Modes	2010 reference sc.	Sc. 2+4	Sc. 2+0	Sc. 1+2
Road	5.0	5.6 (+10.1%)	5.3 (+5.4%)	5.5 (+9.6%)
Rail	9.9 *	10.8 (+9.0%)	10.8 (+9.1%)	10.7 (+8.5%)
Total	14.9	16.3 (+9.4%)	16.1 (+7.9%)	16.2 (+8.9%)

***** These transports are routed via the Great Belt

The shifts to the fixed-link road and rail alternatives happen in the logit mode/route choice model based partly on the service improvements (e.g. decreased OD travel time, decreased waiting time, improved frequency) and partly on the values of the alternative specific constants for the shuttle and bridge alternatives, which are more positive than the Rødby-Puttgarden ferry alternative specific constant.

5.3.2 Discussion of the project results

It is a standard procedure nowadays to apply discrete choice models in the demand models where a new mode (or route) has been introduced. The Fehmarn belt mode/route choice model was developed in the same framework and structure as the previously developed Great Belt and Øresund demand models. The SP data has proved to be superior to the RP data in predicting the potential of a completely new mode or a completely new route expressed through the values of alternative specific constants for the new alternatives.

The values of travel time for five segments presented in Table 5.1 are somewhat higher than the values of travel time in the first two projects. The most likely reason for that is that travel distances in the Fehmarn belt project are longer in distance than the distances in the other projects. A number of theoretical works in the field of discrete choice models have proved that value of travel time per hour increases with distance.

There are a few critical points related to the model estimations:

- The sampling of the respondents is inadequate in the project. First, less than originally planned respondents were interviewed in total in the project. Second, the portions of respondents belonging to the lorry transport and rail transport were in imbalance, as the planned quotes could not be obtained.

- The LOS files proved to be of a rather poor quality in the project. LOS files are important when calculating values of variables presented in the SP experiments for the alternative options (modes/routes). Consequently, in many cases very poor trade-offs were obtained.
- Even though the commodities were split among five groups in the project it is believed that, for instance, segment 2 includes many sub-groups that can vary substantially in VOT.
- SP experiments were based on shipments while the forecasts were ton-based.

The Fehmarn Belt model suffers from the same weakness as the Danish modal split model, i.e. the interviews do not link the trips with the production process. In this way it is believed that the modal split underestimates the future of combined transport.

Further, a pivot-pointing of the goods transport demand matrices is carried out, which forces the modal split overall into the same pattern whether there is a fixed link or not. The latter is not directly linked with the modal split estimation based on RP-data and SP-data, but does have a considerable effect, since it preserves the modal split as observed in the existing situation for the future.

Appendix 1: Discrete Choice Models – a Theoretical Foundation

1. Theory

Aggregate, or the first generation demand models, dominated the transport planning up to late 1970's. Data collected for a particular mode in an aggregate demand model is established on a regional or national level (i.e., zonal information). A fast increase of car transport including lorry transport since 1970 has caused an increased level of congestion on the national and international roads including the increase in the number of traffic accidents, air pollution and noise pollution. An increased awareness of the growth in the national and international goods transport caused a need for traffic models that ware able to describe effects of important freight transport variables. Disaggregate or discrete choice models were developed as an answer to these demands. Discrete choice models are therefore occasionally called the second generation models and they have been widely applied since the beginning of the 1980's.

Data applied in the discrete choice models are twofold. Firstly, the socio-economic characteristics of the choice maker (e.g., a shipper) refer usually to the commodity produced (or transported), the company's turnover and the size. The second type of data is related to a detailed description of a chosen trip. A trip has its costs, time spent in goods terminals, in-vehicle or driving time, delay time, a certain degree of possibility that a shipment will arrive to the destination in a damaged condition, a certain level of information service offered to the customer together with some other types of services.

Data related to a goods trip can relate to a specific trip that occurred in the past, i.e. a set of Revealed Preference (RP) data, or to a hypothetical trip, i.e. a set of Stated Preference (SP) data. Usually, in discrete choice goods models, a respondent reveals one set of RP data (related to one freight trip) and a number of sets of SP data, which is related to a number of hypothetical variations of the revealed trip. Variations are assigned to the trip components as, for instance, travel costs or in-vehicle travel time. The RP set data is therefore associated with actual behaviour and is, therefore, valuable for the forecasts. This data set is also less susceptible to error and bias as a result of the survey instrument. The problem with the values in the RP data set is that it suffers from a high degree of correlation between the variables. Time and cost parameters in goods models, estimated on the RP data only, are usually of the wrong magnitude due to their high correlation (if everything else is constant, the longer the trip the more costly the trip is).

The major strength of the SP data is that the method of collecting this type of data (i.e., SP experiments) is perfectly controlled by the analyst and the correlation among the variables is therefore essentially eliminated. Parameters that originate from SP data have therefore a high statistical accuracy. Because SP data is not real-life data, no goods forecasts can be based on this type of data. Therefore, a number of methods have been developed in order to combine the strengths of the two types of data in the discrete choice modelling. Essentially all of these methods apply the SP data in the estimation phase (in order to achieve a good model fit) and the RP data in the forecasts.

The basic strength of the discrete choice models is in establishing the functional relationship between the respondent's preferences (or behavioural intentions) and the attributes of model travel alternatives, in the form of a utility function. In travel demand modelling, the utility of an alternative A can be defined as benefit that an individual enjoys by spending his resources by choosing alternative A (alternative A can relate to a mode of transport or to a route). By choosing alternative A, a freight operator assigns different weights to the possible alternatives from a choice set according to the previous experience in shipping his products, type of commodity in question, importance of travel time, travel costs and so on. Beside these factors, the freight operator might have his own reasons or constrains for choosing a specific alternative for the trip in question. Due to the existence of the respondent's choice constraints, the utilities measured in goods transport modelling are indirect utilities of the available alternatives. The most common form of utility of alternative A is linear, where the combination of the variables is additive:

(1) $V_a = a_0 + a_1X_1 + a_2X_2 + \ldots\ldots + a_nX_n$

where:

V_a is utility of alternative A,

a_0 is alternative constant,

X_1, X_2, to X_n are variables of alternative A, and

a_1, a_2, to a_n are model parameters.

The alternative constant represents the mean net effect of all the unmeasured components of the utility function, i.e. the impact of the variables that are not included in the utility function.

For a set of, say, n modes in a goods model, some variables can be common to all alternatives, as for example travel costs and in-vehicle travel time. These variables are called generic variables. Opposite to that, some variables are alternative specific variables, e.g. travel frequency exists in rail transport but not in road transport.

The model parameters represent the relative importance attached to each variable. They are sometimes referred to as 'preference weights' and they are estimated by statistical methods such as maximum likelihood.

Limitation of the number of variables in utility V_a as well as respondents' inconsistencies in their choice behaviour introduce the concept of random utility where an error term is included in the previous formulation of utility. Random utility of alternative A, in the set of n available alternatives, is equal to:

(2) $U_a = V_a + e_a$

where:

U_a is random utility of alternative A,

V_a is deterministic function of the variables of alternative A, and

e_a is an error term (random part of the equation).

This idea of random utility is essential in the way that the random element e_a implies that utility is related to the probability of an individual giving a certain response rather than directly to the response itself. This is essentially a more accurate way of modelling choice behaviour than the approach assuming a complete consistency in the way people perceive or express their preferences. The random utility approach requires, however, more complicated statistical methods in the model estimation.

Discrete choice modelling is based on the random utility theory by which it is assumed that a respondent will choose alternative A, in the choice set of n alternatives, if alternative A corresponds to the highest value of utility among the available utilities. This further means that the probability of choosing alternative A between n alternatives is equal to:

(3) $p_a = p[U_a > U_b, \forall b ? a]$

(4) $p_a = p[e_b < V_a – V_b + e_a] = F[V_a – V_b + e_a, \forall b ? a]$

where:

F[.] is a cumulative density function of the variable in brackets.

Depending on the assumptions for the cumulative density function F[a] (or more correctly, for the probability density function f(a)) of the random variable e_a, it is possible to achieve different probability models. The most commonly applied probability model is logit model.

In the logit model it is assumed that the random variables e_a, e_b, etc., for a set of n available alternatives, are independent and identical across the respondents and alternatives, and they are distributed by the Weibul's distribution (which is an approximation of the normal distribution).

For any random variable x that is distributed by the Weibul's distribution, the probability density function f(x) equals to:

(5) $f(x) = ?e^{-x}$, ?>0

where:

? is non-negative parameter, and

e stands for the base of natural logarithms.

The cumulative function F(x) of the random variable x equals:

$$(6)\, F(x) = e^{-l e^{-x}},\; l > 0$$

The probability of choosing alternative A in the set of two alternatives (alternatives A and B), where the random errors of the alternatives are independent and identically distributed by the Weibul’s distribution is:

$$(7)\, p_a = \int_{e_a} F[V_a - V_b + e_a]\, f(e_a)\, d(e_a)$$

Solving the integral, where the cumulative and the probability density functions are substituted by equations (5) and (6), the probability for choosing alternative A is:

$$(8)\, p_a = \frac{e^{Va}}{e^{Va} + e^{Vb}}$$

For the set of n alternatives (multinomial logit model), the probability of choosing alternative A in a logit model is equal to the ratio of the exponential of the deterministic part of the utility of alternative A and the sum of the exponentials of the deterministic utilities of all available alternatives.

$$(9)\, p_a = \frac{e^{Va}}{\sum_{j=1}^{n} e^{Vj}}$$

The most discussed property of the multinomial logit (MNL) model is the property of independence of irrelevant alternatives (IIA) which can be described as:

“Where any two alternatives have a non-zero probability of being chosen, the ratio of one probability over the other is unaffected by the presence or absence of any additional alternative in the choice set” (Ortuzar and Willumsen, 1995).

IIA property implies that the MNL will calculate the probabilities

$$(10)\, \frac{p_a}{p_b} = \frac{\dfrac{e^{Va}}{\sum_{j=1}^{n} e^{Vj}}}{\dfrac{e^{Vb}}{\sum_{j=1}^{n} e^{Vj}}} = e^{Va-Vb}$$

correctly only in the case that the alternatives in the choice set are independent. If not, a so-called *'red bus – blue bus'* i.e. a problem with two very similar alternatives will occur.

An answer to this problem in the family of logit models is a hierarchical logit (HL) model where the correlated (or more similar) options in the choice set are grouped in hierarchies or nests. Each nest is represented by a composite alternative, which competes with other available alternatives.

Usually alternatives in a freight transport model are not correlated. Therefore, the use of HL models is immediately not needed. However, in more sophisticated freight mode-route choice models, modes of transport are placed on a higher choice level while the available routes, per each transport mode, are placed below the modes. Such a possibility is shown in figure 1.

Figure 1 – HL model with one nest

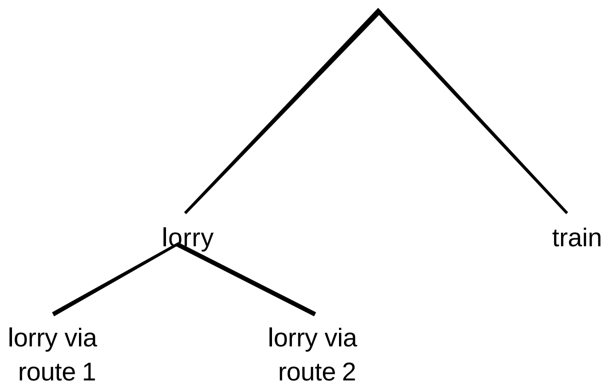


Figure 1 shows that the respondents have 3 choice possibilities: lorry via route 1 (L1), lorry via route 2 (L2) and the train (TR) alternative. Applying the MNL model in this case will produce a wrong result due to the obvious correlation between the two lorry alternatives. The other way of explaining the available alternatives in the figure is that there is a high cross-elasticity between L1 and L2 relative to TR, meaning that a respondent who presently applies L1 will tend to switch to L2 more easily than to TR.

In a HL model, the deterministic utility of a composite lorry (LR) alternative equals to:

(11) $V_{lr} = ? \ln (e^{Vl1} + e^{Vl2})$

where:

V_{lr} is utility function of the lorry composite-alternative,

V_{l1} is utility function of the lorry alternative via route 1,

V_{l2} is utility function of the lorry alternative via route 2, and

T is a parameter to be estimated.

Knowing the value of the lorry composite utility function we get two binomial logit models; one between train and lorry-composite alternatives and the other between two lorry alternatives (via routes 1 and 2).

Probabilities of choosing lorry-composite alternative and train alternative are:

(12) $p_{lr} = e^{Vlr} / (e^{Vlr} + e^{Vtr})$

$p_{tr} = 1- p_{lr}$

Probabilities of choosing lorry alternatives via route 1 and 2, under the constrain that lorry-composite alternative is chosen are:

(13) $p_{l1/lr} = e^{Vl1} / (e^{Vl1} + e^{Vl2})$

$p_{l2/lr} = e^{Vl2} / (e^{Vl1} + e^{Vl2})$

According to the theory, the T parameter must have value between 0 and 1. If T<0 then a possible increase in the utility of alternatives L1 or/and L2, which should increase the value of the utility function of the lorry composite alternative, would actually cause that V_{lr} decreases in value. That would cause a diminishing of probability of selecting the nest. If T=0 then a possible increase in the utility of alternatives L1 or/and L2 would not affect the nest's probability of being selected as V_{lr} would be always 0 (see equation 11). If T>1 then an increase in the utility of alternatives L1 would tend not to affect only its choice probability but also the choice probability of alternative L2. Finally, if T=1 then the HL model will get the shape of a MNL model.

The vast majority of discrete choice models nowadays are MNL and HL models. Beside their straightforward theoretical background, their broad application is a result of the existence of a number of software-packages, which are easy to use. On the other side we find probit models. In the probit model it is assumed that the random variables e_a, e_b, etc., for a set of n available alternatives, are distributed by the normal distribution. If the random error is understood as the sum of a large number of unobserved but independent components, by the central limited theorem the distribution of that stochastic variable would tend to be normal. In the probit model alternatives need not to be independent, i.e. *'red bus – blue bus'* problem cannot occur in the probit model. The probit models are very difficult to compute, and the few available software that allow the estimation of the probit models have a number of serious operational restrictions.

Discrete choice models we know from, for instance, the Great Belt passenger and goods models, the Fehmarn belt passenger and goods models, the Ørestad traffic model are in the shape of HL (hierarchical logit) models.

Few passenger trips are made for the sake of travelling itself, i.e., we travel in order to carry out activities in the trip destination. This is the reason why the mode/route utility functions are called derived or indirect utilities. According to the micro-economic theory of travel behaviour, from the choice set of n alternatives we will choose the alternative with the maximum utility value. The assumption is therefore here, that we behave rationally when choosing a mode/route alternative in the set of n alternatives.

2. Value of travel time in discrete choice models

Let us imagine that a mode choice freight model consists of a train alternative and a lorry alternative, both described by the cost and time variables.

(14) $V_{lorry} = a_0 + c_1 C_{lorry} + t_{lorry} T_{lorry}$

$V_{train} = c_1 C_{train} + t_{train} T_{train}$

where:

V_{lorry} is utility of lorry alternative,

V_{train} is utility of train alternative,

C_{lorry} is lorry travel cost,

C_{train} is train travel cost,

T_{lorry} is lorry travel time,

T_{train} is train travel time,

a_0 is lorry alternative specific constant (to be estimated),

c_1 is cost parameter (to be estimated),

t_{lorry} is lorry time parameter (to be estimated), and

t_{train} is train time parameter (to be estimated).

The model parameters (c_1, t_{lorry} and t_{train}) represent the degree of relative importance attached to each variable by the respondents in the model sample. In the microeconomic theory they are called *'marginal utilities'* of the variables they represent.

Utilities V_{lorry} and V_{train} are measured in utility units, i.e. they are not measured in time or cost units. Assuming that variables C_{lorry}, C_{train}, T_{lorry} and T_{train} have units in DKK (for cost) and hours (for time variables) respectively, then parameters c1 and, t_{lorry} and t_{train} are measured in units 1/DKK and 1/hour, respectively.

By the discrete choice theory, a ratio of any pair of coefficients (marginal utilities) gives a *rate of commodity substitution* or *marginal rate of substitution*. Therefore, the values of travel time (VOT) for lorry alternative equals:

(15) $VOT_{lorry} = t_{lorry} / c_1$

Similarly, the VOT for train alternative is:

(16) $VOT_{train} = t_{train} / c_1$

In our case the VOT for lorry and train alternatives have units DKK/hour. Knowing a VOT for say lorry alternative in the above model, one can find a number of combinations of travel-cost and travel-time variables, for which the corresponding utility values will be of the same magnitude. An example is shown in figure 2.

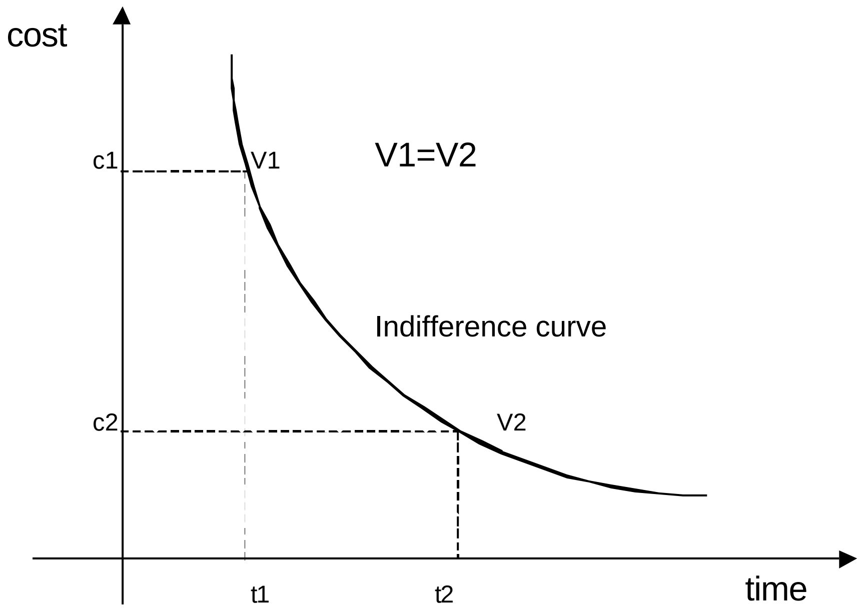


Figure 2 – Indifference curve

The figure shows, that by paying a certain amount of money (i.e., $c_1 - c_2$) for improvement in transport time (i.e., $t_2 - t_1$) the user of transport services will stay on his indifference curve, i.e. he will be equally happy before and after the previously described action (paying for travelling faster). It is therefore the producers of transport services who can profit based on improvements of their services.

REFERENCES

FTC (Fehmarn Belt Traffic Consortium), Fehmarn Belt Traffic Demand Study, Final report, January 1999.

Jovicic, Goran: Study on importance of some Quality Parameteres in Long-distance Goods Transport, Transport og Kvalitet, Transforsk94, 1996

Ortuzar, J. and L.G. Willumsen, 1994: Modelling Transport (John Wiley & Sons, 1994)

Petersen, M. S.; Potentialet for Søtransport, Transportrådet og Trafikministeriet, Maj 1995