FREIGHT TRANSPORT MODELLING

EDITED BY

Moshe Ben-Akiva • Hilde Meersman • Eddy Van de Voorde

Emerald

FREIGHT TRANSPORT MODELLING

FREIGHT TRANSPORT MODELLING

EDITED BY

MOSHE BEN-AKIVA

Massachusetts Institute of Technology, Cambridge, MA, USA

HILDE MEERSMAN

Department of Transport and Regional Economics (TPR), University of Antwerp, Antwerp, Belgium

EDDY VAN DE VOORDE

Department of Transport and Regional Economics (TPR), University of Antwerp, Antwerp, Belgium

Emerald

United Kingdom – North America – Japan
India – Malaysia – China

Emerald Group Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2013

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-78190-285-1

In Memory of Prof. Marvin L. Manheim

Contents

List of Contributors

Megersa A. Abate	Department of Transport, Technical University of Denmark, Lyngby, Denmark
Andreas Balster	Institute of Traffic and Transport, Technische Universität Darmstadt, Germany
Jaume Barceló	Department of Statistics and Operations Research, Technical University of Catalonia, Barcelona, Spain
Moshe Ben-Akiva	Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
Michel Beuthe	Groupe Transport et Mobilité, Université Catholique de Louvain à Mons, Belgium
Denis Bolduc	Département d'économique, Université Laval, Quebec, Canada
Christophe Bouffioux	Groupe Transport et Mobilité, Université Catholique de Louvain à Mons, Belgium
Ennio Cascetta	Dipartimento di Ingegneria Civile, Edile e Ambientale – Universitá degli studi di Napoli "Federico II", Italy
François Combes	Ecole des Ponts (UMR LVMT), France
Antonio Comi	Department of Enterprise Engineering, University of Rome "Tor Vergata", Rome, Italy
Romeo Danielis	University of Trieste, Italy
Gerard de Jong	Significance, The Netherlands; Institute for Transport Studies, University of Leeds, UK and Centre for Transport Studies, KTH Royal Institute of Technology, Stockholm, Sweden
Kate Flanagan	Cintra US, Austin, TX, USA
Hanno Friedrich	Institute of Traffic and Transport, Technische Universität Darmstadt, Germany

Terry L. Friesz — Department of Industrial and Manufacturing Engineering, Pennsylvania State University, PA, USA

Carlos González-Calderón — Center for Infrastructure, Transportation, and the Environment, Rensselaer Polytechnic Institute, New York, NY, USA; Universidad de Antioquia, Medellín, Colombia

Hanna Grzybowska — NICTA (National ICT Australia Ltd) and University of New South Wales, Sydney, Australia

José Holguín-Veras — VREF's Center of Excellence on Sustainable Urban Freight Systems, Center for Infrastructure, Transportation, and the Environment, Rensselaer Polytechnic Institute, New York, NY, USA

Cathérine Krier — KPN Group, Woluwe-St-Lambert, Belgium

Ole Kveiborg — COWI A/S, Lyngby, Denmark

Gernot Liedtke — Institute for Economic Policy Research (IWW), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Rosário Macário — Instituto Superior Técnico, Lisbon University; TIS.PT, Consultores em Transportes, Inovação e Sistemas, S.A., Portugal

Elena Maggi — Departments of Economics, University of Insubria, Italy

Rico Maggi — Institute for Economic Research, University of Lugano, Switzerland

Edoardo Marcucci — DIPES and CREI, University of Roma Tre, Italy

Vittorio Marzano — Dipartimento di Ingegneria Civile, Edile e Ambientale – Università degli studi di Napoli "Federico II", Italy

Lorenzo Masiero — School of Hotel and Tourism Management, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Hilde Meersman — Department of Transport and Regional Economics, University of Antwerp, Antwerp, Belgium

Amir H. Meimand — Department of Industrial and Manufacturing Engineering, Pennsylvania State University, PA, USA

John Mitchell	Department of Mathematics, Rensselaer Polytechnic Institute, New York, NY, USA
Michel Mouchart	Institut de Statistique, Biostatistique et Sciences Actuarielles (ISBA), Université Catholique de Louvain, Belgium
Agostino Nuzzolo	Department of Enterprise Engineering, University of Rome "Tor Vergata", Rome, Italy
Jesús Arturo Orozco	Department of Operations Management, IPADE Business School, Mexico City, Mexico
Andrea Papola	Dipartimento di Ingegneria Civile, Edile e Ambientale – Università degli studi di Napoli "Federico II", Italy
Jay Q. Park	Pantos Logistics, Seoul, Korea
Katherine Rosa	Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
Lucia Rotaris	University of Trieste, Italy
Kees Ruijgrok	TiasNimbas Business School, Tilburg University, The Netherlands
Francesco Russo	Università degli Studi Mediterranea di Reggio Calabria, Italy
Ricardo Sanchez	Cintra US, Austin, TX, USA
Iván Sánchez-Díaz	Center for Infrastructure, Transportation, and the Environment, Rensselaer Polytechnic Institute, New York, NY, USA
Stefan Schröder	Institute for Economic Policy Research (IWW), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
Erika Spissu	Cintra US, Austin, TX, USA
Siri Pettersen Strandenes	Department of Economics, Norwegian School of Economics, Bergen, Norway
Yichen Sun	Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
Lóri Tavasszy	TNO and Infrastructure Systems & Services Department, Delft University of Technology, Delft, The Netherlands

Ellen Thorson	University Transportation Research Center, City College of New York, New York, NY, USA
Tomer Toledo	Technion – Israel Institute of Technology, Haifa, Israel
Eva Valeri	University of Trieste, Italy
Eddy Van de Voorde	Department of Transport and Regional Economics, University of Antwerp, Antwerp, Belgium
Inge Vierth	Swedish Road and Transport Research Institute (VTI) and Centre for Transport Studies (CTS) Stockholm, Sweden
Roberta Vitillo	Dipartimento di Ingegneria Civile, Edile e Ambientale – Università degli studi di Napoli "Federico II", Italy
Qian Wang	Department of Civil, Structural and Environmental Engineering, University at Buffalo, Buffalo, NY, USA
Ning Xu	Systems Development, Delta Air Lines, Atlanta, GA, USA
Bo Zhang	Department of Industrial and Manufacturing Engineering, Pennsylvania State University, PA, USA
Li Zhang	Institute for Economic Policy Research (IWW), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

INTRODUCTION

Chapter 1

Recent Developments in Freight Transport Modelling

Moshe Ben-Akiva, Hilde Meersman and Eddy Van de Voorde

There can be no doubt that the economic globalization of the last decades and the growing need for flexibility in modern enterprises have transformed freight transport and turned it into a major public policy and corporate domain. Freight transportation research has reflected this evolution and is quite justifiably attracting ever closer attention.

Transportation is not just the product of social and economic activity. Good and reliable transport remains a sine qua non for sustained economic growth. Since production and consumption of goods and services are usually physically separated, the distance between the two needs to be bridged by means of at least one mode of transportation.

Similarly, relocating production activities, often from high-cost to low-cost countries, can only be achieved through better, cheaper and more extensive transportation services. The other side of the picture is that an unrestrained expansion of passenger and freight transport will create substantial negative externalities such as air pollution, congestion, accidents and damage to infrastructure. Consequently, if the relevant policies remain absent, the social costs of mobility may exceed the benefits.

Quite a number of international organizations, including the World Bank, IMF, UNCTAD, OECD and many others, have acknowledged the need for effective transport policy. However, implementing and, as the case may be, adjusting such transport policy is not a straightforward proposition. Continuous monitoring and effective insights are required to afford decision-makers the ability to successfully design and pursue transport policies while responding adequately to new challenges. Despite prolific research on passenger transport in the 1970s and 1980s, the pace of economic globalization since the 1990s has caused researchers and policymakers to shift their primary focus to freight transport.

As the late Professor Marvin Manheim emphasized in quite a few of his publications and in his opening address to the 8th World Conference on Transport Research (Antwerp, 12–16 July, 1998), effective and sound resolutions for such issues require a new and broader transport analysis. This book aims to contribute to such

Freight Transport Modelling

ISBN: 978-1-78190-285-1

an analysis by presenting the insights of a wide and international range of experts with a view to pushing forward frontiers in freight transport modelling. Hence, it is intended for transport researchers in general and for those working on freight transport modelling in particular. It guarantees value added for experienced researchers and doctoral students alike. Moreover, the link with transport policy and management will be of interest to transport decision-makers in both government and industry.

In order to capture the complexity of freight transport systems, researchers have proposed a wide array of models. De Jong, Van de Riet, and Kroes (2004) characterized such models as consisting in one or more mathematical–empirical relations designed to describe and explain the behaviour of a transport system. Ultimately, by taking into account that any transport system is subject to exogenous shocks and/or policy measures, these models can provide insight into possible future evolutions in freight transport.

Small and Winston already asserted in 1999 that freight models must also represent special characteristics of transport markets, most importantly the interactions within or involving the transport system. Ideally, a whole range of important factors, such as localization, trade issues, destinations, infrastructure, shipment and parcel size, timing and frequency of shipment, quality of service, transportation mode, routing, costing and inventory holdings, as well as possible interactions with passenger transport, should be entered into the equation.

To complicate matters further, due attention must be paid to the dynamic nature of transport systems. Some decisions need to be taken consecutively and require harmonization in order to optimize the transport and logistics chain. Others are to be taken simultaneously and may involve a high degree of interaction. Finally, freight transport models must allow for a considerable time lag between a decision and its implementation.

Thus far, the majority of freight transport models that have been put forward deal with specific topics and tend to be designed to deal with a limited number of interactions. The main constraint on the development of more elaborate freight transport models is the limited availability of data, especially at the level of individual firms.

Looking back at past decades, the modelling of freight transport demand has evolved from a non-structural, aggregate engineering approach that is conventionally used for traffic management and routing decisions to a structural, disaggregate approach. The aggregate models utilize global data on shippers and shipments to identify general relations resulting from underlying behavioural assumptions. The more sophisticated models rely on flexible functional forms and test such traditional restrictions as homogeneity, economics of scale and separability. With new empirical methods and growing availability of firm-level data, transport modellers have turned their focus on more behavioural disaggregated analyses.

In the literature, one can find various examples of both aggregate and disaggregate freight transport modelling. There are a number of studies providing extensive overviews of the state of the art in freight modelling (see, e.g. Tavasszy, 2006). Three important fields are identified: the modelling of the relationship between transportation and economic activity; logistic decision-making and processes; and the linking of traffic flows and networks.

As Small and Winston (1999) have quite rightly pointed out, '*economists have primarily, though not exclusively, focused on mode choice*'. Over the past two decades, transport researchers have clearly broken with this tradition, as evidenced in this book.

1.1. Content of the Book

The book consists of three parts. The first part deals with freight transport modelling from a global (international) point of view. The second part considers freight modelling from a regional perspective. The final part concentrates on the local/urban level.

The global section encompasses four chapters. The opening contribution, by Hilde Meersman and Eddy Van de Voorde, discusses the relationship between economic activity and freight transport. By applying stability and co-integration tests, the authors show that gross domestic product (GDP) is not the best indicator for modelling this relationship in the long run. There are several reasons for this. Some have to do with the changeable composition of GDP, others have to do with the altered relationship between freight transport and economic activity due to the globalization of the economy, policies aimed at decoupling freight and economic activity and changing business behaviour (time-based competition, labour vs. transport costs, …). All this makes reliable long-term aggregate forecasting of freight transport on the basis of GDP very difficult indeed. A number of alternatives are suggested for estimating a reliable relationship between freight and well-chosen relevant indicators of economic activity. The general conclusion is that more specific disaggregate approaches are needed that are based on detailed microeconomic underpinnings of the behaviour of shippers and freight transport companies.

Ennio Cascetta, Vittorio Marzano, Andrea Papola and Roberta Vitillo discuss Multi-Regional Input-Output (MRIO) models for freight demand simulation at national level. They specify a model with elastic trade coefficients and a multimodal freight supply model on a European geographic scale. From the demand side, an elastic trade coefficient MRIO model is presented and some relevant macroeconomic feedbacks are discussed that are incorporable into the model. From the supply side, a critical review is presented of the complexity of the multimodal freight networks and the corresponding modelling requirements. From a practical standpoint, the implementation of an MRIO model at European level is reported, describing in particular both the database and the supply model for the calculation of transport impedances required by the trade coefficient model. Finally, the authors present some simple applications of the implemented MRIO model with elastic trade coefficients, before drawing conclusions and outlining some research suggestions.

In their contribution to freight choice analysis and market research, Moshe Ben-Akiva and Gerard de Jong start from the observation that most freight transport models applied by national, international and regional authorities ignore important aspects of logistics decision-making, such as the choice of shipment size and the use of consolidation and distribution centres. At the same time, these models often assume transport (mode) optimization at the aggregate zone-to-zone level, for which in reality no agents exist. The Aggregate-Disaggregate-Aggregate (ADA) freight model system tries to overcome these limitations by modelling the generation of trade flows and assignment to networks in an aggregate way, but simulating the logistics decisions at the level of individual firm-to-firm flows. The disaggregate part of the system is the logistics model, where shipment size and transport chain choice (including the use of transhipment centres) are determined by minimizing the total logistics cost. Most countries lack commodity flow surveys that include this information, but the disaggregate logistics model can also be calibrated to more aggregate data on the mode shares by commodity and aggregate zones. Models relying on the latter approach have been developed in Norway, Sweden and Flanders, and are under development in Denmark and for the European Union.

Terry L. Friesz, Amir H. Meimand and Bo Zhang model the dynamic shipper–carrier problem as a differential Stackelberg game. The shippers act as followers operating under the Cournot–Nash behavioural assumption while competing on the sale of a homogenous product in several markets. They choose their strategies simultaneously by maximizing their respective utility functions while assuming their competitors' strategies are fixed. The leader in this problem is a transportation company referred to as the carrier who seeks to maximize its objective function while considering the followers' reactions. The authors show that the differential Nash game describing the shippers' competition may be articulated as a differential variation inequality (DVI). The DVI is then reformulated as a non-linear complimentarily problem (NCP) to characterize the follower's Nash–Cournot equilibrium by the implicit solution of a system of equations. This allows the shippers' response to the carrier's policy to be embedded into the carrier's problem as a set of nonlinear constraints. The penalty method is then used to reduce the number of explicit constraints and finite-dimensional time discretization is employed to approximate the model as a mathematical program that may be solved by the multi-start global optimization scheme found in the off-the-shelf software package GAMS when used in conjunction with the commercial solver MINOS. They also present a small scale numerical example with two followers and four markets.

The second part of the book is comprised of 11 chapters on freight modelling at a regional level. The development of the disaggregate logistics choice model is the subject of the contribution of Moshe Ben-Akiva, Denis Bolduc and Jay Q. Park. They begin with a choice model based on the assumption that a shipper attempts to minimize the total logistics cost that includes transportation as well as order and inventory costs. They estimate their basic choice model using revealed preferences data from a US market research survey and extend this basic model in a number of ways. They apply recent developments in discrete choice analysis to better capture unobserved heterogeneity among shippers and to introduce into the model specification new variables that measure the perceived service qualities to the alternative modes. These extensions are then shown to be more powerful when the revealed preferences data are augmented with stated preferences data collected in the same survey. The results obtained by combining revealed and stated preferences data are shown to be superior to those obtained from revealed preferences only.

Michel Beuthe, Christophe Bouffioux, Cathérine Krier and Michel Mouchart present a systematic comparison of four methodologies for separately analysing individuals' stated preferences relative to the choice of alternative solutions of freight transport. These are defined by the monetary cost and five qualitative attributes: frequency of service, transport time, reliability, carrier's flexibility and damages/losses. As the data consist of alternatives rankings, the models applied are somewhat unusual in the field, at least in transportation analysis: conjoint analysis, UTA-type multi-criteria analysis, rank-ordered conditional logit and neural network analysis. This chapter applies the above four models, some of them with non-linear utility functions, to the individual rankings of nine firms' transport managers in diverse sectors of industry. The alternatives submitted to their judgement are conceived according to an orthogonal fractional factorial design. Each estimation methodology is adjusted and specified to suit the specific data. Over this small set of individual firms, each method applied shows that cost is the most important factor in individual choice making for seven out of the nine transport managers. Reliability is often more important than transport time, even though its relative importance varies from case-to-case. The other factors may play a significant role in some cases. These outcomes are mostly coherent with the earlier results obtained at an aggregate level. The two better-performing methods are the multi-criteria and the neural network analyses, which both involve non-linear partial utility functions.

Gernot Liedtke, Stefan Schröder and Li Zhang assert that, in order to analyse the impacts of policy measures aiming at influencing the behaviour of various logistics actors (such as road pricing instruments, regulation and market interventions), it is necessary to model decision-making in logistics explicitly, including the emergence of spatiotemporal logistics structures. This would be similar to activity-based models for passenger transport that deduce the spatiotemporal movement patterns of individuals from their planning problems. Such types of models consider the complex reactions of individuals coherently and allow for a deduction of individual welfare changes. Multi-agent systems allow for the modelling of the local interactions of these actors. The authors' overview describes some models dealing with the emergence of spatiotemporal structures in freight transport by integrating optimization methods and co-ordination mechanisms. Three modelling efforts are presented: the first model is on the formation of regional logistics agglomerations; the second concerns the formation of tours at a national level; and the third relates to the formation of transport chains and vehicle utilization. Together, they provide an outlook on the prospects of integrating optimization methods and co-ordination mechanisms in a multi-agent system.

Lorenzo Masiero and Rico Maggi deal with accounting for the discrepancy between the willingness to pay (WTP) and the willingness to accept (WTA) in discrete choice models. A key input in cost-benefit analysis is represented by the marginal rate of substitution which expresses the WTP or its counterpart, WTA, for both market and non-market goods. The consistent discrepancy between these two measures observed in the literature is suggestive of the need to estimate reference-dependent models that are able to capture loss aversion by distinguishing the value attached to a gain from the value attached to a loss according to reference-dependent theory. The authors propose a comparison of WTP and WTA measures estimated from models with both symmetric and reference-dependent utility specifications within two different freight transport stated-choice experiments. The results show that the reference-dependent specification outperforms the symmetric specification and they prove the robustness of reference-dependent specification over datasets designed according different attributes levels ranges. They also demonstrate the policy relevance of asymmetric specifications, illustrating the strong implications for cost-benefit analysis in two case studies.

François Combes, Kees Ruijgrok and Lóri Tavasszy start from the observation that, on theoretical grounds, it is already widely accepted that the choice of shipment size influences the choice of mode. However, due primarily to a lack of data, there have been few attempts to empirically test the quality of this relationship. The authors propose a model for the choice of transport mode in which the costs of transport follow the logic of economics of shipment size choice. They specify a discrete choice model where the generalized cost depends on transport distances, mode abstract values of time and continuous shipment sizes. The model is estimated on observations of about 10,000 individual shipments from the French ECHO survey. The results show that mode choice is strongly consistent with the economics of shipment size choice.

Hanno Friedrich and Andreas Balster deal with supply-chain risk analysis with extended freight transportation models. The analysis and management of supply-chain risks has indeed become more important in a world with globally integrated production networks in which extreme events seem to occur more frequently. But transparency, in particular for the overall economic, production and logistics system, is lacking. The authors discuss the possibility of using approaches from freight transportation analysis to analyse consequences of supply-chain risks on overall logistics and freight transportation. A literature review shows that there is a research gap for these models, especially for the analysis of short and medium-term impacts of risks (days to weeks). An overview of possible reactions of economic actors to risks and challenges of modelling these risks recommends that such models should be explanatory, and that they should map time explicitly and probably be specific to sectors and countries.

Tomer Toledo, Yichen Sun, Katherine Rosa, Moshe Ben-Akiva, Kate Flanagan, Ricardo Sanchez and Erika Spissu studies the decision-making process and the factors that affect truck routing. The data collection involved intercept interviews with truck drivers at three rest area and truck stop locations along major highways in Texas, Indiana and Ontario. The computerized survey solicited information on truck routing decisions, the identity of the decision-makers, the factors that affect routing and sources of information consulted in making these decisions. In addition, Stated Preferences (SP) experiments were conducted, in which drivers were asked to choose between two route alternatives. The data was used to study the identity of routing decision-makers for various driver segments and the sources of information used both in pre-trip planning and en-route. A random effects logit model was estimated using the SP data. There are significant differences in the route choice decision-making process among various driver segments, and these decisions are affected by multiple factors beyond travel time and cost. These factors include shipping and driver employment terms, such as the method of calculation of pay and bearing of fuel costs and tolls.

Edoardo Marcucci reviews a set of articles, based on stated preferences techniques, focusing on logistics managers' preferences for freight transport attributes with the intent of assessing the quality of knowledge acquired through applied research, as well as its reliability and transferability. The review indicates that there are some evident shortcomings in the way research has been performed so far, but, at the same time, there appears also to be a high potential if corrective actions are taken. In particular, substantial improvements could be attained by: clearly defining research and reporting protocols; defining and circumscribing who has to be interviewed in the different freight contexts studied; reporting the contractual relationships governing freight movements; reporting freight details (e.g. volume, value, weight); motivating the attribute selection method used. There is a need for systematizing procedures and reporting in applied research by introducing a much higher level of detail and rigor, both in defining the object of measurement and in the experimental design protocols employed.

Megersa A. Abate and Ole Kveiborg discuss a central aspect of freight transportation, that is capacity utilization, in the context of empty running of commercial freight vehicles. They provide an overview of the literature on these topics and distinguish between two types of contribution according to their analytical approach and origin of research. The first approach looks at utilization based on economic theories, such as the firms' objective to maximize profitability, and considers how various firm and haul (market) characteristics influence utilization. The second approach stems from the transport modelling literature and its main aim is to analyse vehicle movement and usage in a transport demand modelling context. A strand of this second group of contributions is the modelling of trip-chain and its implications in terms of capacity utilization. A key lesson to be learned is that it is important to take into account the commercial activity that initiates vehicle movements in evaluating performance. There is room for further enhancement of the modelling exercise by incorporating information regarding the operator in order to provide a stronger behavioural basis for the vehicle movements and utilization analysis.

The purpose of Inge Vierth's contribution is to demonstrate how freight transport time savings and reduced variability in transport time are valuated in CBA and to discuss alternatives. The author suggests a structure where the benefits of transport time savings are related to the reduction of transport resources required, capital tied up in goods while transported, and goods users' opportunity costs. The benefits of less variability in transport time are due to the reduction of transport resources required for a given service and users' opportunity costs. Normative and behavioural approaches are applied to determine the value of transport time savings (VTTS) and the value of reduced variability in transport time (VTTV). The VTTS definition differs between countries. Different VTTS and VTTV constructions make the transfer of results difficult. It is found that the savings in transport resource costs can be calculated with the aid of engineering formulae, provided that the VTTS is limited to the capital tied up in the goods, where market-based approaches can be used, and VTTS(O). The incorporation of VTTV in CBA requires further development of the valuation methods. The use of information on the trade-off between transport costs and inventory costs is one promising approach.

Siri Pettersen Strandenes deals with freight transport pricing models. Freight transport prices typically increase proportionally with distance and with weight or size. Trade models traditionally apply iceberg transport costs, where the value of the good transported is reduced upon arrival by a percentage of the initial value, to reflect transport and other trade costs, whereas freight charges in transport and logistics models reflect physical characteristics such as weight and distance. The author shows that non-linear pricing models can be applied to build priority pricing schemes where shippers pay a surcharge for priority handling of their cargoes. Basing point pricing and unified delivered prices commonly used in freight transport pricing imply spatial price discrimination. Yield management developed by passenger airlines upon deregulation of the airline industry in the United States has spread to air cargo. The EU's recent prohibition on liner conferences makes yield management more probable also in container shipping.

The final part of this book focuses on the local/urban level and includes six chapters. The main objective of José Holguín-Veras, Ellen Thorson, Qian Wang, Ning Xu, Carlos González-Calderón, Iván Sánchez-Díaz and John Mitchell is to provide a comprehensive overview of empirical findings and models that focus on urban freight tours (UFTs). Freight demand modelling researchers and practitioners are cognizant of the need to explicitly model UFTs, and are collecting data and developing models that account for UFT. There are some obvious limitations. The data collected are small and, in most cases, are unable to provide a comprehensive view of UFTs, while the models developed are still in need of significant improvements. Some of the more pragmatic approaches, such as simulations and hybrid models, are likely to require enhancements in their behavioural foundations, as they rely on assumptions that are not always validated. On the other hand, the most theoretically appealing models – such as those based on spatial price equilibrium – still require computational improvements to make them ready for real-life use.

Francesco Russo deals with modelling behavioural aspects of urban freight movements and presents the main behavioural aspects of urban goods movements and the approach to modelling them. The movements in question are generated by the restocking and purchasing decisions of retailers and consumers in an urban context. A general model system to structure and interconnect all the choices, with relative decisions, is presented in the form of a wide-meshed grid in which researchers and practitioners may introduce specific models presented elsewhere, or calibrated ad hoc in a real urban context. The system considers the four main vehicle movements: push and pull, generated by end-consumers and retailers. It also considers a commodity model where the quantity choice model is discussed with its relative implications in the main decisional level for each decision-maker. The advancements relative to the alternative set model and the choice within the set are proposed, with reference to push and pull movements, opening up possibilities for future researches. Due consideration is given to: new discrete choice model; dynamic choice and presence of ITS; dispersion of taste. This chapter attempts to build a bridge between the transport paradigm and the shopping-restocking process and supply planners by means of an integrated and updated outlook.

The contribution of Romeo Danielis, Elena Maggi, Lucia Rotaris and Eva Valeri starts from the assumption that adopting a supply chain approach is crucial to understanding how urban freight distribution works and how urban transport policy alternatives may impact on supply chain performance. In accordance with a recent strand in the literature, this chapter aims at: characterizing the urban supply chains; discussing how an urban supply chain can be modelled, which role actors play and how the coordination issue can be handled; showing how transport choices, in particular between own-account or third-party transport operators, are dealt with in each urban supply chain and by each actor; and analyzing how urban supply chains are affected by the many proposed freight transport policies. Although much progress has been made in this field, with regard to both modelling and empirical analysis, it emerges that further advances are needed in relation to both the *ex-ante* and the *ex-post* evaluation of the private and social efficiency of the different urban supply chains and how they are impacted by local authorities' transport policies.

Rosário Macário recognizes that the diversity of problems associated with urban logistics prevents the organization of urban logistics with a single solution. Diversity of urban profiles and of business models must be successfully matched to solve that complexity. This rational serves as the basis for designing a future master plan for urban logistics, which might be used by local authorities to regulate their territory and to implement measures that enable inducement of better behaviour by the economic agents who take part of the urban supply chain. Therefore a '3-step approach' methodology has been followed. The first step comprises the definition of 'Logistic Profiles'. The second step encompasses the evaluation of different logistics solutions based on different criteria in terms of their suitability for serving different logistics profiles. The third step consists in modelling the changes which will occur if the solutions are implemented. This approach is validated by modelling a first pilot area in Lisbon.

Agostino Nuzzolo and Antonio Comi first classify city logistics measures that city administrators can use to reduce the negative impacts of urban freight transport, in relation to planning (i.e. strategic and tactical/operational). The focus then shifts to models developed to support the assessment of tactical and operational measures. The assessment procedures require the simulation of freight transport demand and hence the estimation of freight vehicle origin-destination (O-D) flows. These O-D flows can be obtained from the simulation of delivery tours. Therefore, this chapter presents a system of models that is able to simulate delivery tours using an aggregate approach. Such models allow one to capture actors' choices that can be influenced by tactical and operational measures. They were calibrated and tested on the basis of surveys conducted in the inner area of Rome, where more than five hundred truck drivers were interviewed.

The chapter by Jaume Barceló, Jesús Arturo Orozco and Hanna Grzybowska deals with making real-time fleet management decisions under time-dependent conditions in urban freight distribution. The design and evaluation of City Logistics applications requires an integrated framework in which all components can work together. Therefore, City Logistics models should account for vehicle routing applications and fleet management models also capable of dealing with the dynamic aspects of traffic flows in the underlying road network, namely when ICT applications are taken into account. The authors develop a methodological proposal based on an integration of vehicle routing models and real-time traffic information. In the computational experiments conducted, a dynamic traffic simulation model is used to emulate the actual traffic conditions, providing, at each time interval, estimates of the traffic state on each link of the road network used, by a real-time fleet management system, in order to determine the optimal dynamic routing and scheduling of the fleet.

Acknowledgements

This book is dedicated to the memory of Prof. Marvin Manheim.

In a book we edited in 2008, Prof. Werner Rothengatter (2008), former chairman of the World Conference on Transport Research Society (WCTRS), characterized Prof. Manheim as '*a most pro-active thinker, always trying to be some steps ahead of the present reality and to anticipate the most important drivers of future development*'. Prof. Rothengatter (2008) rightly stated that Marvin Manheim may be regarded as the spiritus rector of modern activity-based freight modelling and logistics.

Prof. Manheim understood that human beings as individuals or in organizations may accelerate processes, but that they can also act as barriers to development. Bringing people together to form global and worldwide networks is therefore a productive way of reducing barriers, of fostering social cohesion and of enhancing economic progress. Marvin Manheim applied this idea to various fields. Hence it is no coincidence that he was also a founding father of the World Conference on Transport Research (WCTR) and its Society (WCTRS). The development of WCTRS illustrates Marvin Manheim's notion that sustainable networks have to be most flexible and adaptive to new developments. No rigid centrally steered hierarchies, but autonomous working groups of highly motivated and committed individuals, collaborating in a truly international spirit.

In 2008, Marvins's wife, Mary Beth Watson-Manheim (2008), wrote: '*Twenty-five years ago, the first sentence in Marvin's textbook stated: "We live in a world of rapid change." Throughout his career, he never stopped probing and searching for new models of the changes he envisioned across multiple disciplines and problem settings, with a relentless focus not only on theory but also on translating the theory into practical applications. Those who knew Marvin were also influenced by personal qualities, which were interwoven through his research and teaching. Marvin was an inveterate teacher, he was creative, an integrator and a visionary, often seeing connections not immediate obvious to most people. The ability to integrate can also be*

seen in the communities of colleagues he developed, and the sheer enjoyment he found in debating and discussing ideas, while also savouring food and wine with friends and colleagues'.

We sincerely hope that this book will contribute in a profound way to the scientific knowledge of and research into freight transport modelling, in the spirit of Marvin Manheim.

To conclude this introduction, the editors would like to extend their sincere thanks to all those who contributed to the publication of this book. First and foremost, we are grateful to the various authors. They did a great job. We also thank all the referees for their indispensable and much appreciated assistance. Finally, our gratitude goes to all those who contributed logistically to ensuring the success of this publication.

References

De Jong, G., Van de Riet, O., & Kroes, E. (2004). *Rol goederenvervoermodellen in beleid* (p. 96). Den Haag: RAND Europe.

Rothengatter, W. (2008). Marvin Manheim: Founder of the World Conference on Transport Research Society and Spiritus rector of the integration of human resources and technological systems in transport. In M. Ben-Akiva, H. Meersman & E. Van de Voorde (Eds.), *Recent developments in transport modelling: Lessons for the freight sector* (pp. 11–17). Bingley, UK: Emerald.

Small, K., & Winston, C. (1999). The demand for transportation: Models and applications. In J. Gomez-Ibanez, W. B. Tye & C. Winston (Eds.), *Essays in transportation economics and policy. A handbook in honor of John R. Meyer* (pp. 11–55). Washington, DC: The Brookings Institution.

Tavasszy, L. (2006). *Goederenvervoer: verre vriend én goede buur!* (p. 23). Nijmegen: Radboud Universiteit.

Watson-Manheim, M. (2008). The contribution of Professor Marvin L. Manheim to transport and logistics science. In M. Ben-Akiva, H. Meersman & E. VandeVoorde (Eds.), *Recent developments in transport modelling: Lessons for the freight sector* (pp. 19–29). Bingley, UK: Emerald.

GLOBAL (INTERNATIONAL)

Chapter 2

The Relationship between Economic Activity and Freight Transport

Hilde Meersman and Eddy Van de Voorde

Abstract

Traditionally the relation between economic activity and freight transport is used to make forecasts of future aggregate freight flows and volumes. Usually gross domestic product (GDP) is used as an indicator for economic activity in a region or a country. National and international organisations collect data on GDP and also make forecasts of GDP. This explains to some extent the reason for using GDP as indicator for future freight transport.

In this contribution we will show by using stability and co-integration tests that GDP is not the best indicator because its composition changed and is still changing, because some methods to link freight transport to GDP are not suited, and because the link between freight transport and economic activity itself has been changed. There are several reasons for this: the globalisation of the world economy, policies aimed at decoupling freight and economic activity, and changing business behaviour (time based competition, labour vs. transport costs …). This makes reliable long-term aggregate forecasting of freight transport on the basis of GDP very difficult.

We suggest a number of alternatives to estimate a reliable relation between freight and well chosen relevant indicators of economic activity. The general conclusion is that more specific disaggregate approaches are needed based on detailed microeconomic underpinnings of the behaviour of shippers and freight transport companies.

Keywords: Economic activity; co-integration; forecasts

Freight Transport Modelling

ISBN: 978-1-78190-285-1

In his book of 1979 Marvin Manheim put the interrelationship between the transportation system of a region and the socioeconomic system at the heart of transportation systems analysis:

> The transportation system of a region is tightly interrelated with the socioeconomic system .... This interrelationship is fundamental to our view of transportation systems analysis. (Manheim, 1979, p. 12)

Traditionally the analysis is split into two parts: one focussing on the contribution of freight transport to economic growth and economic structures, and one considering the impact of economic activity on the demand for freight transport. The latter is the major topic of this chapter.

2.1. Freight Transport in Europe: General Indicators

It is clear that the transport sector as such takes an important part in generating economic activity. In most industrialised countries, on average 12% of total consumption expenditures are related to transport. The sector generates a considerable part of total value added and contributes directly for approximately 3–4.5% to total employment (Table 2.1). If one adds all the indirect effects, the importance of the transport sector becomes even more pronounced.

Table 2.1: Importance of the transport sector (direct effects) in the EU 27

	Absolute figures		**Shares of total economy**	
	Employment 2009 × 1000	**Value added 2010 mil euro°**	**Employment 2009 (%)**	**Value added 2010 (%)**
Land transport; transport via pipelines	6314	232045	2.8	2.2
Water transport	223	44259	0.1	0.4
Air transport	371	26528	0.2	0.3
Sub-total 1	*6908*	*302832*	*3.1*	*2.9*
Supporting and auxiliary transport activities; activities of travel agencies	3190	155762	1.4	1.5
Sub-total 2	*10098*	*458594*	*4.5*	*4.4*
Total economy	223628	10384518		

°Millions of euro, chain-linked volumes, reference year 2005 (at 2005 exchange rates)
Source: Eurostat National Accounts.

During the last decades inland freight transport in tonne-kilometres (t-kms) followed closely the evolution of the gross domestic product (GDP) in the EU27 as is illustrated by Figure 2.1. It even started to outpace economic growth in 2004, but the global financial and economic crisis hit the freight transport sector much stronger than the overall economic activity. However, in 2010 the growth rate of freight transport was again 5.3% which was considerably more than the 2% GDP growth. A similar evolution can be found in Russia. In the United States goods transport measured in t-kms did not follow the growth of GDP. As a consequence freight in the United States was less affected by the global economic crisis than in Europe or in Russia. In China the exceptional large increase of goods transport in 2008 is striking. This is due to the more than doubling of road freight.

A closer look at the annual growth rates in Table 2.2 reveals that the strong growth of freight transport is mainly attributed to road haulage which expanded over the last 15 years by an annual average of 2.2% for the period 1996–2010 or 3.3% for the period 1996–2007.

As a consequence, road haulage is increasingly the dominant mode of transport in Europe (Figure 2.2). While in 1970 it was less than twice the size of rail transport, in 2010 it is five times as important. This evolution is quite different from the one in the United States and Russia, where rail is the dominant mode for freight transport.

The increasing dominance of road haulage in Europe is illustrated even more sharply by the evolution of the freight intensity relating freight levels to GDP (Table 2.3). Where the relation between total t-kms and GDP has hardly changed and even went down, road freight intensity went continuously up from 0.098 in 1970

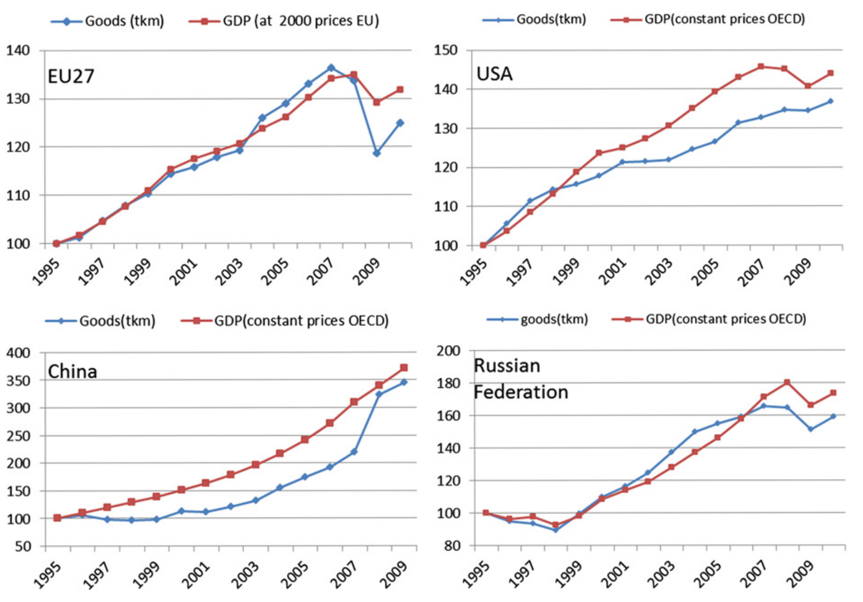

Figure 2.1: Freight transport performance and GDP (1995 = 100).
Source: EU and OECD statistics, 1995–2010.

Table 2.2: Growth of freight transport in t-km (road, rail and inland waterways (IWW)) and of gross domestic product in euro of 2005 (GDP) in the EU27

	Road	Rail	IWW	Total	GDP
1996	1.08	1.55	−1.91	0.98	1.83
1997	3.77	4.43	6.76	4.11	2.73
1998	4.63	−4.16	2.47	2.58	2.99
1999	3.94	−2.26	−1.74	2.30	3.06
2000	3.32	5.23	3.99	3.73	3.90
2001	2.47	−4.39	−1.00	0.90	1.98
2002	3.19	−0.57	−0.02	2.29	1.25
2003	1.22	2.11	−6.78	0.88	1.35
2004	7.18	6.22	10.70	7.21	2.51
2005	2.98	−0.51	1.41	2.25	1.97
2006	2.99	6.34	−0.15	3.39	3.23
2007	3.62	2.89	4.67	3.55	2.99
2008	−1.77	−2.28	−1.22	−1.83	0.52
2009	−10.05	−18.33	−16.30	−11.90	−4.21
2010	3.82	7.84	22.68	5.53	2.00

Source: European Commission, EU Transport in Figures, Statistical Pocketbook 2012.

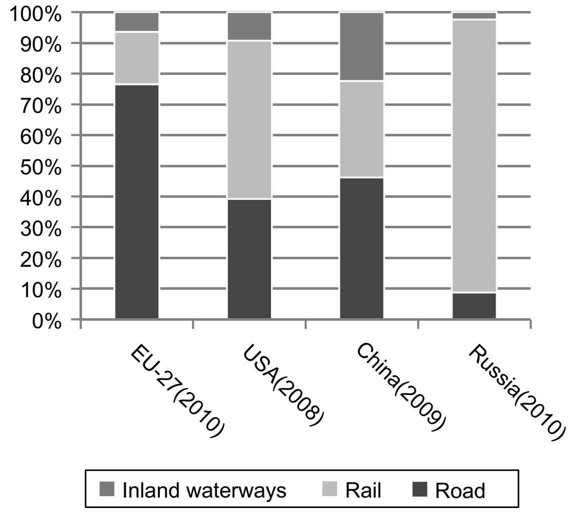

Figure 2.2: Mode shares in freight transport (t-kms).
Source: European Commission (2012).

Table 2.3: GDP freight intensities in the EU15

	Road/GDP	Rail/GDP	Total/GDP
1970	0.098	0.057	0.176
1980	0.106	0.043	0.165
1990	0.113	0.030	0.155
1995	0.122	0.024	0.158
2000	0.125	0.023	0.160
2005	0.126	0.022	0.159
2010	0.109	0.021	0.140

Source: Own calculations based on EU Energy and Transport in Figures 2002, EU Transport in Figures, 2012 and OECD National Accounts.

to 0.126 in 2005 and rail freight intensity went down from 0.057 in 1970 to 0.021 in 2010. This evolution is not without negative social impacts. Although the worldwide economic crisis and the European debt crisis have hit road freight more than GDP, a recovery might bring again the road freight intensities to pre-crisis levels.

The European Commission has expressed its concern with this evolution in a number of White Papers. The 2011 White Paper 'Roadmap to a Single European Transport Area' highlights the importance of an efficient transportation system for the functioning of the internal market and the integration of the European Union (EU) in the global world market (European Commission, 2011). But the negative impacts of the present transport systems should be drastically reduced. Therefore, policy measures are needed to transform the European transport system in such a way that it can safeguard mobility together with a significant reduction of greenhouse gas emissions.

2.2. The Relation between Freight Transport and Gross Domestic Product

In order to be successful the measures taken to reduce the negative impacts of freight transport should be built on a good and clear understanding of the global relationship between freight transport and economic activity. Especially for making forecasts this relationship is crucial. It is at the heart of nearly any type of freight transport model, whether it is based on the traditional four-step approach or the activity based approach, whether it is for global forecasts or sector specific predictions. Whatever approach is chosen, the basic or starting point is most of the time the relation between freight transport demand and economic activity. If this relation is not represented in an appropriate way, it will weaken the rest of the model and the forecasts. Therefore, it should be modelled carefully using the most suitable variables, data and techniques.

In the past it was often assumed that there was a simple one-to-one relationship between GDP or value added and the demand for freight transport. To a large extent the use of GDP was inspired by its availability and the fact that most national and international institutions made forecasts for GDP. The elasticity was generally put equal to one. Although this simple relationship seemed to work rather well, it became clear that it cannot be applied uniformly to all transport modes, all regions and all commodity types.

A first indication of this changing relationship is the fact that models developed in the past have underestimated the actual growth of freight transport. Table 2.4 gives an example of the forecasts based on a model which was built in 1990 for forecasting road haulage in a number of European countries, and the actual realised growth of road haulage (Meersman & Van de Voorde, 1999). The forecasts for the total growth over the period 1990–1999, based on an error-correction model, were made in 1990 for a scenario with an annual GDP growth of 1% and for a scenario with 3% GDP growth. With an actually realised average annual GDP growth of hardly 1% over the period 1990–1999, it is clear that the model did severely underestimate the growth of road haulage. This means that the relationship between economic activity and freight transport has changed over time, especially in the 1990s.

So the research questions are clear: Is the elasticity of freight transport demand with respect to economic activity equal to one for all modes and is it stable over time? Are there differences among countries? Are there differences among commodity sectors? Which indicator for economic activity should be used?

Table 2.4: Road haulage forecasts for the period 1990–1999 and actual growth of road haulage

Country	**Scenarios for GDP growth (%)**	**Forecast growth (%)**	**Actual growth (%)**
Belgium	1	25.76	50
	3	43.44	
Germany	1	38.49	87
	3	63.95	
France	1	12.6	34
	3	34.4	
The Netherlands	1	18.85	53
	3	29.63	
United Kingdom	1	6.54	15
	3	20.73	

Source: Meersman and Van de Voorde (1999, pp. 23–48).

To treat the first two questions, panel data estimation techniques can be used. They have the advantage that they combine data for a number of countries over a number of time periods and endow regression analysis with both a spatial and temporal dimension. The common feature of this type of data is that in general the number of time periods over which the countries are observed is relatively limited. The major panel data estimation methods are constant coefficients or pooled regression, fixed effects methods and random effects methods. The choice between those methods is not an easy one, but as is mentioned by Verbeek (2012) the fixed effects methods are more appropriate when the temporal dimension is rather small and when the cross-section units cannot be viewed as a random draw which is the case for countries.

To study the stability of the freight elasticity over time, data on GDP, total freight transport and road haulage for 11 European countries are collected for the years 1970, 1980, 1990, 2000 and 2010.[1] To see whether the elasticity of freight transport demand with respect to GDP, has changed over time, the following specifications were considered:

$$\text{lntot}_{it} = \alpha_t + \beta_t \text{lngdp}_{it} + \varepsilon_{it} \tag{2.1}$$

and

$$\text{lnroad}_{it} = \gamma_t + \delta_t \text{lngdp}_{it} + \upsilon_{it} \tag{2.2}$$

with

lntot_{it} the logarithm of total freight transport in country i in year t;
lnroad_{it} the logarithm of total road freight transport in country i in year t;
lngdp_{it} the logarithm of the GDP in country i in year t;
α_t β_t γ_t δ_t coefficients which might vary over t;
ε_{it} ν_{it} stochastic error terms.

Both equations were estimated using the panel data fixed effects estimators using dummy variables DT such that Eqs. (2.1) and (2.2) can be written as:

$$\text{lntot}_{it} = \sum\nolimits_t \alpha_t \text{DT}_{it} + \left(\sum\nolimits_t \beta_t \text{DT}_{it}\right) \text{lngdp}_{it} + \varepsilon_{it} \tag{2.3}$$

and

$$\text{lnroad}_{it} = \sum\nolimits_t \gamma_t \text{DT}_{it} + \left(\sum\nolimits_t \delta_t \text{DT}_{it}\right) \text{lngdp}_{it} + \upsilon_{it} \tag{2.4}$$

1. The countries in the sample are Belgium, Germany, Greece, Spain, France, Ireland, Italy, the Netherlands, Austria, Portugal and the United Kingdom. GDP is measured in constant prices, freight transport in t-kms. The data are from the OECD (http://stats.oecd.org/).

Table 2.5: Pooled and fixed effects estimation results for the relation between total freight traffic and road haulage, and gross domestic product (11 EU countries, 1970–2010)

Dependent variable	**lntot**				**lnroad**		
	Coefficient	**Standard error**	***t*-Value**		**Coefficient**	**Standard error**	***t*-Value**
Pooled estimation							
α	−2.85	0.57	−4.98	γ	−5.24	0.67	−7.81
β	1.07	0.045	23.72	δ	1.21	0.053	22.95
Log likelihood	−34.54				−43.29		
Adjusted R^2	0.91				0.91		
S.E. of regression	0.462				0.542		
Sum squared residuals	11.308				15.55		
Fixed effects estimation							
α_{70}	−2.082	0.736	−2.830	γ_{70}	−5.178	1.949	−2.656
α_{80}	−2.362	0.806	−2.930	γ_{80}	−4.596	1.525	−3.014
α_{90}	−2.808	0.966	−2.906	γ_{90}	−4.863	1.822	−2.669
α_{00}	−3.797	1.602	−2.371	γ_{00}	−5.839	2.415	−2.418
α_{10}	−5.271	2.276	−2.316	γ_{10}	−6.119	2.330	−2.625
β_{70}	1.023	0.055	18.728	δ_{70}	1.209	0.151	7.981
β_{80}	1.036	0.057	18.080	δ_{80}	1.164	0.116	10.057
β_{90}	1.066	0.070	15.312	δ_{90}	1.184	0.135	8.744
β_{00}	1.142	0.117	9.760	δ_{00}	1.265	0.178	7.115
β_{10}	1.239	0.167	7.418	δ_{10}	1.276	0.171	7.469
Log likelihood	−30.23				−42.78		
Adjusted R^2	0.91				0.89		
S.E. of regression	0.463				0.582		
Sum squared residuals	9.667				15.26		

$$\text{where} \quad \text{DT} = 1 \ \text{if} \ T = t \quad \text{with } T \text{ and } t = 1970, 1980, 1990, 2000, 2010$$
$$= 0 \ \text{if} \ T \neq t$$

The results, corrected for heteroscedasticity, are reported in Table 2.5.

It is clear that, gradually the relation with economic activity has changed since 1970 and this change is the largest in the nineties although the changes are not statistically significant. The elasticities of freight transport based on the estimates in Table 2.5 are represented in Figure 2.3. In Meersman and Van de Voorde (2008)

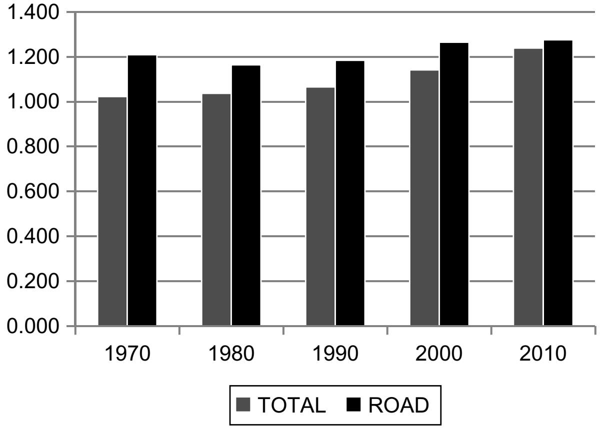

Figure 2.3: Evolution of the elasticity of freight transport with respect to GDP based on estimation results in Table 2.5.

the results pointed to a more significant change in the freight transport–GDP relation. The difference is clearly due to the change in database for the road haulage figures.[2]

The previous approach does not take into account the full dynamics which are present in the relation between freight and economic activity and which can be very different on a country level. This is illustrated in Figure 2.4 with annual data for France, Germany and the United Kingdom for the period 1970–2010. Not only is there a difference in the slope of the trend line, but also in the cyclical fluctuations around the trend line. This suggests that the relation should be modelled on an individual country base.

The main question is whether there is a stationary long-term relation between freight transport and GDP on a country level. One way to check the stability of the relationship is by means of the CUSUM test and the CUSUM of squares test (Brown, Durbin, & Evans, 1975). Table 2.6 gives an overview of the CUSUM and the CUSUM of squares tests for 19 countries. The equation on which the tests are based for each country, *i*, is

2. In Meersman and Van de Voorde (2008) the dataset consisted of EUROSTAT data which for road freight contained the number of t-km performed on the territory of the reporting country. From 1995 the road freight transport statistics are reported by Member States for vehicles registered in their country. This implies that the figures will no longer reflect all the freight transport on the territory of the reporting country. For countries with large amounts of transit flows, there can be a considerable difference between the t-kms performed by vehicles registered in the country and the t-kms actually performed on the territory as is for instance the case for Belgium. As it was not possible to find a consistent database on the t-kms performed on the countries' territory, we decided to use the OECD data which are available for a longer time period.

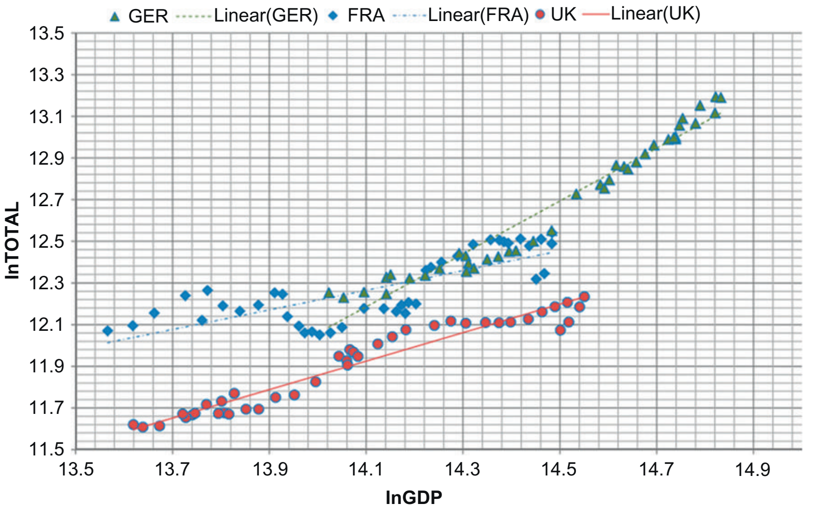

Figure 2.4: Total freight transport and GDP in Germany, France and the United Kingdom.
Source: Based on data from the OECD, 1970–2010.

$$\text{lntot}_{it} = \alpha_i + \beta_i \text{lngdp}_{it} + \varepsilon_{it} \tag{2.5}$$

extended for some countries with dummy variables for the years for which the OECD reports a break in the data.

Another way to check whether there is a stable long-term relation between freight transport and GDP is to test whether they have a common trend which keeps their long-term evolution together. Formally this is done by testing whether they are co-integrated. As only two variables are involved, it is possible to use the Engle–Granger approach (Engle & Granger, 1987) using the appropriate critical values of Davidson and MacKinnon (1993). The results are presented in Table 2.7. The conclusion is quite clear: it is not possible to find a common trend in freight transport and GDP over the period 1970–2010 for the majority of the countries under consideration.

A first consequence is that it is not possible to make reliable long-term forecasts for total freight transport in t-km on a country level based on economic activity represented by GDP. A second consequence is that only a short-term relationship can be modelled in a statistical reliable way. As for all the countries under consideration freight transport and GDP are both integrated of order one, the relation has to be estimated using first differences of the variables.[3]

3. Again dummy variables are used for the years in which the OECD reports a break in the freight data series.

Table 2.6: CUSUM and CUSUM of squares test for stability of the relation between total freight transport in t-kms in logarithms (LNTOT) and GDP in logarithms (LNGDP)

Sample 1970–2010 (USA 1980–2010)		**STABLE AT 5%?**	
Country	**Equation**	**CUSUM**	**CUSUM OF SQ**
Australia	LNTOT = − 6.77 + 1.45*LNGDP	NO	
Japan	LNTOT = 2.36 + 0.68*LNGDP	NO	
Turkey	LNTOT = − 6.24 + 1.37*LNGDP		NO
United States	LNTOT = − 13.40 + 1.79*LNGDP	NO	
Austria	LNTOT = 0.72*LNGDP + 1.45 + 0.52*D95		NO
Belgium	LNTOT = 1.08*LNGDP − 2.67		NO
Denmark	LNTOT = 0.80*LNGDP + 0.02		
Finland	LNTOT = .85*LNGDP + 0.44 − 0.06D95 − 0.06D03	NO	
France[a]	LNTOT = 0.06*LNGDP + 11.37 + 0.27*D95		
United Kingdom	LNTOT = 0.66*LNGDP + 2.51		NO
Germany	LNTOT = 1.24*LNGDP − 5.34	NO	NO
Greece	LNTOT = 0.76*LNGDP + 0.12		NO
Ireland	LNTOT = 0.95*LNGDP − 1.87	NO	NO
Italy	LNTOT = 1.64*LNGDP − 10.96–0.39*D97	NO	
Luxemburg	LNTOT = 0.08*LNGDP + 6.38		NO
Netherlands	LNTOT = 0.59*LNGDP + 3.46		NO
Portugal	LNTOT = 0.89*LNGDP − 1.25	NO	
Spain	LNTOT = 0.76*LNGDP + 1.32 + 0.29*D02		
Sweden	LNTOT = 0.69*LNGDP + 2.20		

[a]Coefficient of LNGDP is not significantly different from zero.

$$\Delta \text{lntot}_{it} = (\alpha_i) + \beta_i \, \Delta \text{lngdp}_{it} + \varepsilon_{it} \tag{2.6}$$

The results of the fixed effect estimations are reported in Table 2.8.

The short run elasticities of total freight transport with respect to GDP vary among countries but are for most of the countries larger than one. Although this relation gives an idea of how freight transport will react to changes in economic activity, the danger is in using it to make long-term forecasts.

2.3. Why is this Relationship between Freight and GDP Changing?

Although nobody will question the fact that freight transport is a consequence of economic activity, the relation is a dynamic and complex one. There are a set of

Table 2.7: Results of Engle–Granger co-integration test for freight transport (t-km) and economic activity measured alternatively by gross domestic product for the period 1970–2010

	test-statistic	Cointegrated?		test-statistic	Cointegrated?
Australia		YES**	Germany	– 2.62	NO
Japan	– 2.78	NO	Greece	– 2.96	NO
Turkey	– 3.23	YES*	Ireland	– 2.70	NO
United States[a]	– 2.81	NO	Italy	– 2.56	NO
Austria	– 2.53	NO	Luxemburg	– 1.94	NO
Belgium	– 1.15	NO	Netherlands	– 3.15	YES**
Denmark	– 3.19	YES**	Portugal	– 1.69	NO
Finland	– 3.82	YES*	Spain	– 2.51	NO
France	– 3.09	NO	Sweden	– 2.81	NO
United Kingdom	– 2.05	NO			

*at 5%; **at 10%.
[a]Sample 1980–2010.

Table 2.8: Fixed effects estimates of the short run elasticity of freight with respect to GDP

	$\hat{\beta}$		$\hat{\beta}$
Australia	1.275	Germany	1.986
Japan	0.573[ns]	Greece	– 0.039[ns]
Turkey	1.270	Ireland	2.096
United States	0.654[ns]	Italy	1.154
Austria	2.216	Luxemburg	1.448
Belgium	2.024	Netherlands	1.627
Denmark	1.530	Portugal	2.272
Finland	0.885	Spain	1.375
France	2.341	Sweden	1.564
United Kingdom	1.377		

[ns]Not significant.

factors which have an impact on the changing intensity of this relationship. In what follows a non-exhaustive overview is given of a number of influential factors such as certain logistical developments, the altered role of government, and the aspect of capacity and capacity utilisation.

2.3.1. International Logistical Developments

With the growing significance of international trade, more and more companies have in recent years become involved in what has come to be referred to as international logistics.[4] Globalisation of production and trade generates substantial goods flows between countries which must be dealt with as efficiently as possible. Furthermore, there is a growing tendency in international logistics towards supply-chain management (SCM) and time-based competition. Finally, developments in informatics and especially EDI have made the organisation of international networks for goods and documents a lot more simple and efficient.

International logistics inevitably displays a certain complexity, not least because it involves various actors, each with their own objectives and characteristic behaviour. With a view to retaining and/or improving their own competitive position, a number of these players are keen to maximise the benefits of possible cooperation. In itself, this can lead to better and more streamlined operations.

These operations unfold at various levels, and give rise to various channels: (1) the international transaction and payment channel, (2) the international distribution channel through which the goods physically move and (3) the documentation-communications channel.[5] The resulting problems which may arise in terms of international goods flows can be illustrated quite clearly in a European context, where – even over relatively short distances – many frontiers may have to be crossed. One of the most significant problems is that national transport systems are often inadequately geared to each other.[6] An adequate co-ordination of goods flow and document flow can prevent unnecessarily long and considerably costly delays. It is in this sense that information technology is able to make an increasingly important contribution to international logistics (Manheim, 1999, p. xix).

---

4. International logistics distinguishes itself from domestic logistics in a number of ways. The dissimilarities are mostly due to the fact that goods are moved between countries that may differ from each other considerably. Young (1995) referring to Wood, Barone, Murphy, and Wardlow (1995), rightly asserts in this context that 'international logistics is not the same as domestic logistics with perhaps the addition of one or more international border crossings. International logistics certainly contains border crossing issues, but also must accommodate a significantly higher level of complexity comprising various combinations of cultural, political, technological, and economic variables'.

5. Besides a smooth physical movement of goods through an international network, sufficient attention needs to be paid to terms of sale and terms of payment, as these become much more complicated when buyer and seller are from different countries and as more intermediaries are involved in the transaction. Indeed, differences in commercial and insurance law may be considerable. It may demand a lot of time and effort before all parties involved are able to reach consensus on time and place for the transfer of ownership and responsibility for insurance of the goods. Moreover, goods may be in transit for some time, certainly in case of maritime traffic, so that terms of payment can be crucially important.

6. The SORT-IT study (Strategic Organisation and Regulation in Transport – Interurban Travel), which was conducted at the request of the European Commission, shows that interconnection and interoperability are not merely impeded by technical differences. In fact, organisational and regulatory problems often constitute a much more serious barrier to efficient goods traffic in Europe (Beaumont, Preston, & Shires, 1996).

The key to success in international trade is logistical support (Meersman & Van de Voorde, 2001). A properly functioning logistics system can, after all, cut costs considerably and may thus contribute towards improving the competitive position of a firm. This explains why so much significance is attached to studying the characteristics of international logistics systems and the environment in which they operate. International logistical activities clearly take place in an environment that has changed a great deal and indeed continues to evolve all the time. Consignors must take this into account, as such developments have a significant impact on the relation between costs and services. A divergent evolution and response pattern in different countries can give rise to the previously observed empirical results.

A number of important developments affecting the international flow of goods and international logistics are: globalisation of the production process, growing competition in international trade, SCM strategies, time-based competition and growth of computerisation, EDI and global e-commerce. Let us consider these aspects in some greater detail.

2.3.1.1. Globalisation of the production process Globalisation implies international mobility of goods and services, as well as persons and capital (Meersman & Van de Voorde, 2006, p. 1). Essentially, any decision about the relocation of production or parts of the production process to a country other than the country where the product is sold depends on cost considerations. The following question is of central importance in this respect: do the economies of scale offered by factories specialising in the production of certain components for a global market outweigh the economies of scope offered by factories that produce more extensive packages for a local market?

The answer to this question depends on the following factors:

– the extent of modulation and standardisation of the production process,
– the evolution of the local consumption level,
– the possibility of spreading existing technologies geographically,
– the share of the transport costs in the overall cost structure.

On the one hand, a rising local consumption, stimulated by economic growth, will lead to a sufficiently high demand to allow local production. As a consequence, companies will go international and invest a considerable amount of their financial assets abroad. Figure 2.5 offers an overview of the trans-nationality index of the top 100 transnational companies (TNCs) for 1993, 1997, 2004 and 2011. The trans-nationality index is calculated as the average of foreign assets to total assets, of foreign sales to total sales, and of foreign employment to total employment. Over the period 1993–2011, TNCs became increasingly internationalised, selling more abroad and generating more employment abroad.

However, declining real transport costs, possibly enhanced by an increasing value of the goods transported and a declining ratio of weight against volume, is conducive to a concentration of production (or part of it) in specialised factories. Declining

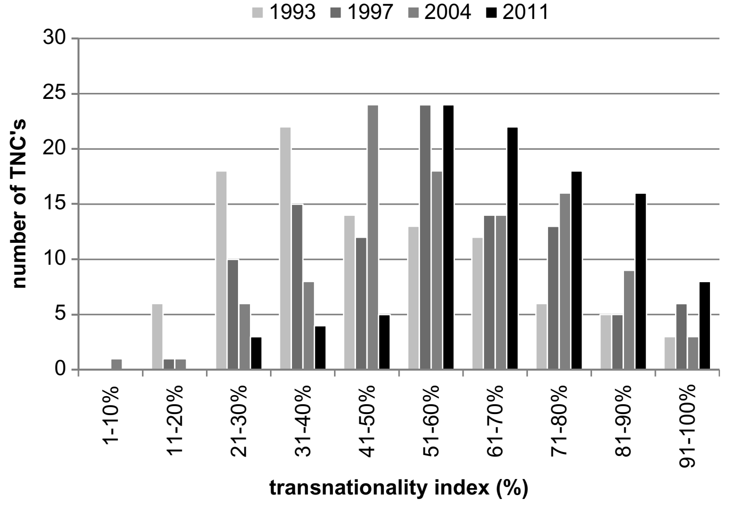

Figure 2.5: Distribution of the trans-nationality index of the top 100 transnational companies (TNCs).
Source: Based on UNCTAD (1995, 1998, 2005, 2012).

telecommunication and computer costs may contribute further to a smoother internationalisation of the production process.

2.3.1.2. Growing international trade Besides globalisation of the production process, there is also a trend towards globalisation of product markets. World trade has been growing at an average rate which is considerably greater than the growth rate of GDP (Figure 2.6). Especially in the nineties world trade was booming. Factors that stimulate international trade play an important role in this respect. Policies aimed at the reduction of tariffs and quotas have led to an increase of the bilateral trade flows between countries.

But it is the ICT revolution which shaped the new globalisation era. As described by Baldwin (2011) the telecom and internet revolutions made it possible to unbundle different production stages spatially. An important consequence is the increase of international trade in intermediate goods. From the mid-1980s the value and growth rate of intermediate goods trade is quite similar to trade in final goods (Sturgeon & Memedovic, 2010, p. 10).

2.3.1.3. Supply-chain management strategies Another important development is the growing importance of SCM strategies. This concept is based on establishing relationships with partners further up or down the logistics chain. The purpose is twofold. On the one hand, one aims for quality improvement of the logistics product. This implies, among other things, greater reliability, a smoother goods flow through the chain, and more efficient connections between the various links in the chain. On

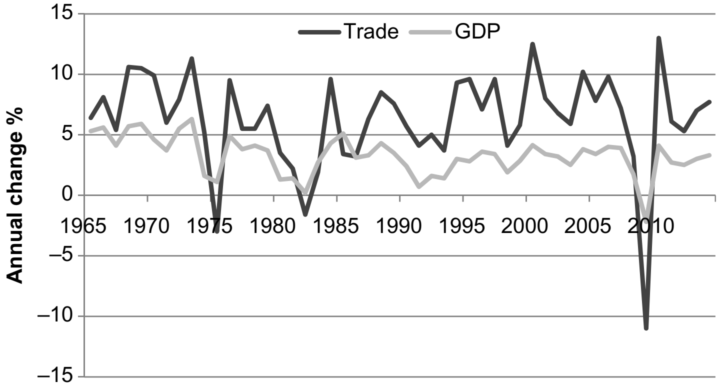

Figure 2.6: Rate of change in world GDP and world trade (estimates start after 2011).
Source: IMF World Economic Outlook, 2012.

the other hand, one strives to realise this at the lowest possible cost for the chain as a whole. This can only be achieved if one has overall control over the logistics chain.

By way of illustration, it may be interesting to consider the structure of a seaport. Every port is, in itself, a chain consisting of consecutive links, while the port as a whole is a link in a global logistics chain. Over the course of time, the relative significance within the port of those separate links has clearly changed. This is due to, among other things, important efficiency-enhancing technological developments, such as the increasing degree of containerisation, the growing dimensions of vessels, quicker cargo-handling etc. Consequently, it no longer suffices to concentrate on one or even a few links in the chain.

This evolution in the function and purpose of seaports obviously has a number of consequences for their organisation and management (Meersman, Van de Voorde, & Vanelslander, 2003). Port operations involve a substantial number of parties, at the policy-making, the managerial and the operational levels. These different players may be subsumed under a single company, as in the case of certain private ports in the United Kingdom, or they may represent a multitude of enterprises and institutions within the port.

The port as a physical entity is managed by a port authority, which is in turn monitored and/or regulated by a higher, often public and political or administrative authority. The authorities may thus be represented to various degrees in the port.

In addition to a port authority, and depending on the size of the port, there are usually a considerable number of companies who have established themselves in the port area. Empirical research has focused on gaining insight into these various port actors and how they interrelate (cf. Coppens et al., 2007). A first empirical application relates to the port of Antwerp.

The port authority occupies a central position. The other actors may be roughly divided into two groups: the port users and the service providers. Among the port

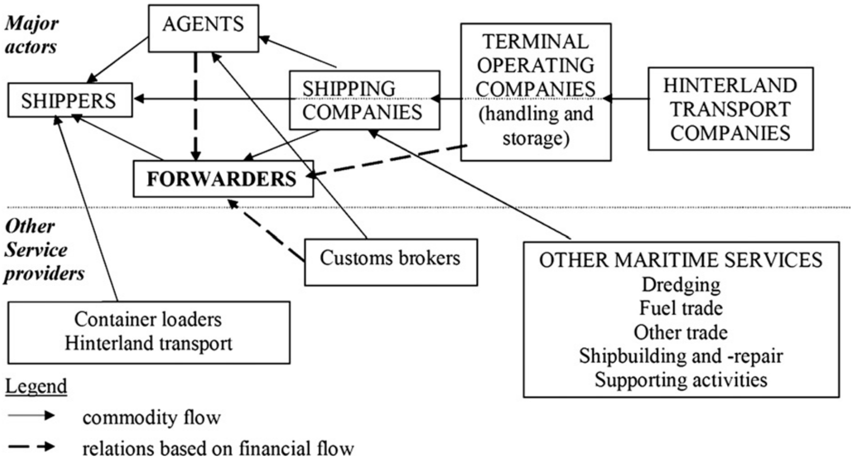

Figure 2.7: Adjusted relations between port actors: Commodity and financial flow point of view.
Source: Coppens et al. (2007).

users are, first and foremost, the shipping companies. Also belonging to this group are the shippers and industrial enterprises who are established within the port perimeter and have lands in concession. The service providers are a heterogeneous group: pilots, towage services, agents, forwarders, ship repairers, suppliers of foodstuffs and spare parts, waste reception facilities and bunkerers. Stevedores, who are increasingly evolving into Terminal Operating Companies, constitute a special case. They provide services (transhipment, storage, stripping and stuffing …) to shipping companies and shippers, for which they are effectively remunerated. At the same time, they pay the port authorities for terminal concessions.

Figure 2.7 illustrates quite clearly how the large number of actors involved in port activities, each with their own objectives, gives rise to a strong degree of heterogeneity within and between ports.[7] As a result, the consequences of increasing globalisation for the various market players will not be equally far-reaching and tangible. A much more far reaching consequence is that the competitive position of a port, and hence the freight flows it will attract and generate, is no longer solely in the hands of the port authority. It will also be determined by the strength of the other actors. Their choices and strategies will therefore have an impact on the magnitude and the routing of the freight flows to and from the hinterland.

7. As a consequence of this strong heterogeneity, the objectives of a port authority, for example, are in part determined by the extent to which that authority is subjected directly or indirectly to outside influences, external control, or direct competition from other ports. It is therefore not surprising that these objectives often differ greatly and have been known to shift significantly over longer periods of time (Suykens, 1986, p. 108).

2.3.1.4. Time-based competition Differentiation between products and services is often founded on so-called time-based competition (TBC). In essence, this concept boils down to developing strategies for production and delivery with the purpose of supplying customers with products in as little time as possible. Cost, value and speed are no longer regarded as possible trade-offs for each other, but as objectives in their own right, to be realised through effective TBC programmes (Hise, 1995).

The significance of time-based competition is especially apparent in the growing success of companies such as the so-called 'integrators' (i.e. DHL, UPS, Fedex, TNT), who provide a range number of services (including strategic supply management, return repair inventory) and specialise in speedy delivery.

2.3.1.5. Growth of EDI and global e-commerce The introduction of internet, e-commerce and e-business has important consequences for international logistics. From 2001 to 2011 the mobile phone subscriptions increased from 15.5 to 85.7 per 100 inhabitants on world scale, but there are still considerable differences between the developed and developing countries (Table 2.9). By the end of 2010, there are an estimated 5.3 billion mobile cellular subscriptions worldwide, including 940 million subscriptions to 3G services. Access to mobile networks is now available to 90% of the world population and 80% of the population living in rural areas. Over the last five years, developing countries have increased their share of the world's total number of internet users from 44% in 2006 to 62% in 2011. Today, internet users in China represent almost 25% of the world's total internet users and 37% of the developing countries' internet users. (ITU, ICT facts and figures, The World in 2010, 2011)

As a consequence of this development, production for stock will decline, while production on order will increase. This will result in transportation of smaller batches and growing problems with regard to reverse logistics.

Table 2.9: Global ICT developments, 2001–2011, per 100 inhabitants

	2001	2005	2010	2011
Mobile-cellular telephone subscriptions				
World	15.5	33.9	77.1	85.7
Developed	47.1	82.1	114.5	122.3
Developing	7.9	22.9	68.9	77.8
Individuals using the internet				
World	8.0	15.7	29.2	32.5
Developed	29.4	50.8	66.8	70.2
Developing	2.8	7.7	21.0	24.4
Fixed-telephone subscriptions	16.6	19.1	17.8	17.3
Active mobile-broadband subscriptions		11.2	15.7	
Fixed (wired)-broadband subscriptions	0.6	3.4	7.7	8.5

Source: ITU World Telecommunication (2012)/ICT Indicators database.

2.3.1.6. Where will these international logistics developments end? Each of the aforementioned developments clearly affects the international logistics process. While globalisation of production and product markets enhance the significance of international logistics, the growing interest in SCM and time-based competition tend to have an impact on the organisation and structuring of the international logistics chain. The emergence of worldwide commerce on the internet not only leads to a greater need for international logistics systems, but also to important procedural changes. This puts increasing pressure on a provision of services that is aimed at speed, reliability, frequency and low costs. Obviously, the repercussions will vary depending on the goods category, transport mode and geographical area.

2.3.2. The Role of Government

The role of the government in the freight transport sector is built around three targets: the reduction of the negative external impacts of freight transport, the implementation of fair and efficient pricing, and to support infrastructure investments. Over the decades, a political custom has developed in many countries to regard interventions in the transport sector as a favoured tool for achieving all kinds of objectives in areas well beyond the realm of transportation.

Transport policy measures, especially infrastructure investments, have been used to enhance the economic development of certain regions, to boost the competitiveness of other sectors of industry, to provide social assistance to the needy, and to stimulate employment. More often than not, the question of whether transport measures were actually the most efficient means of achieving those diverse objectives was simply never raised (Blauwens, De Baere, & Van de Voorde, 2012, p. 382). In terms of means deployed, this is certainly a very narrow viewpoint. After all, one tends to rely one-sidedly on transport-related measures without ever weighing up the option of intervening in other areas of the economy.

There are however explanations for why transport has become the favoured area for all manner of general public interventions (Blauwens et al., 2012, p. 383). A first explanation concerns the nature of transport itself: transport requires infrastructure, an area where government must inevitably play a role. Examples that come to mind are competencies with regard to expropriation, the designation of the path of new transport links and policing. A second historical explanation concerns the monopoly power which the railway industry used to enjoy in the 19th century.

This way, a trend was set. Infrastructure became a tool for boosting employment, for serving local interests, for influencing the price of land, for attracting industry, for enhancing tourism (Blauwens et al., 2012, p. 384). It is a trend that prevails until this day, including within the EU. Typical examples are the enormous investments that are either planned or have already been made in relation to the Trans European Networks, or so-called TENs. Here, too, quite substantial subsidies have been or are still being made available for investment projects inspired not only by existing transport needs, but also by considerations of economic development, integration and social cohesion. In the EU, the total investment on transport infrastructure during the period 2000–2006 was €859 billion (EC, 2009). The cost of EU infrastructure development to match the demand for transport has been estimated at over €1.5 trillion for 2010–2030. The completion of the TEN-T network requires about €550 billion until 2020 out of which some €215 billion can be referred to the removal of the main bottlenecks (http://ec.europa.eu/transport/themes/infrastructure).

In the past the negative effects of an unrestrained growth of transport inspired politicians to design measures to decouple freight and economic growth, current policies are much more designed around safeguarding mobility in a sustainable way. In the White Paper of 2011 (European Commission, 2011) the focus is on innovation, pricing and funding of infrastructure in order to reach the following goals by 2050:

- No more conventionally fuelled cars in cities;
- 40% use of sustainable low carbon fuels in aviation; at least 40% cut in shipping emissions;
- A 50% shift of medium distance intercity passenger and freight journeys from road to rail and waterborne transport;
- All of which will contribute to a 60% cut in transport emissions by the middle of the century.

2.3.3. The Capacity Issue

Scale increases in operations and goods flows, coupled with further internationalisation, have shown quite clearly that sufficient and free capacity is becoming increasingly important. Capacity encompasses not only capital stock, but also capacity utilisation and available capacity.

The significance of available capacity is quite apparent in the maritime sector, more specifically in the manner that international shipping companies decide which ports and port terminals to call at and to include in shipping schedules.

Terminals are important links in the logistics chain. However, shipping companies are often larger and more powerful, and are thus able to impose their rules of play: the economic benefits that ship owners seek to achieve through large-scale expansion and corresponding cost reductions must not be wasted through quayside bottlenecks or, in a subsequent phase, bottlenecks in hinterland transportation. Ship owners will not accept any waiting times and/or other time losses.

The position of the ship owners is perfectly understandable. First and foremost, the required investments in fleet renewal are enormous. Moreover, the goods flows involved are very substantial. Consequently, any impediment due to lack of available capacity implies loss of time and money.

This pressure from ship owners has, in the first instance, affected the traditional structure of what used to be known as stevedoring businesses, but which have since developed into much more complex terminal operating companies. In most cases, this evolution was occasioned by a need for substantial investment capital that could no longer be made available by the original owners. A wave of mergers, takeovers and externally financed expansion projects ensued, causing a concentration movement, coupled with market entries by new, primarily Asian, market players such as PSA and Hutchison Whampoa. This evolution at once created a buffer against possible vertical integration on the initiative of ship owners.

From the perspective of the ship owners, the trend towards ever-greater concentration among terminal operators obviously poses an economic risk: less mutual competition, lower productivity growth, longer handling times for cargoes, and – above all – higher rates. Ship owners will thus no longer be confronted with different, competing terminal operating companies, but with larger players who are active in various ports and who are therefore able to negotiate package deals. Consequently, port competition will, most probably, gradually shift from the level of individual port authorities to that of private terminal operators, that is large groups that offer regional networks.[8]

It is within this framework that one has to place the investments in port and terminal capacity, which quickly translated into additional overcapacity. The consequence is strong capacity growth at a moment in time when, globally, there is already a problem of overcapacity and when, in certain European ports, some terminals are seriously underused.

2.4. Modelling the Relation between Freight and Economic Activity: There is more than GDP

The relation between freight transport and GDP as measure of economic activity is clearly far more complicated than what traditional models would suggest. This has consequences for the way in which reliable long-term freight forecasts can be obtained. Although GDP is the main indicator of economic activity on a country level, it is too general to be used in most of the aggregate freight models because it consists for a large part of value added generated in the services sector. The problem can be solved partly by using more relevant indicators for the freight generating economic activity which can differ by country, by goods category and by mode. Additionally it would be ideal to allow also for indicators which capture the reasons for the changing transport intensity. Sometimes this will imply a careful modelling of the components of the logistics chains and of the choices which are made on different levels in this chain, but sometimes a careful selection of proxies for economic activity can already lead to a vast improvement in the aggregate models.

8. The question arises how shipping companies will respond to such an evolution. As they are at risk of losing relative market power, they may be expected to focus increasingly on acquiring so-called 'dedicated terminals', possibly under joint ventures with terminal operators who are active in the local market.

As an example freight transport in France is modelled using the production in the manufacturing sector as relevant indicator for economic activity. Additionally the impact of the globalisation is introduced by using the export performance and the import penetration of France.[9]

It is possible to find a co-integrating, long run equilibrium relation between total freight, the index of manufacturing production, the export performance and the import penetration rate. This relation was estimated using fully modified least squares to get reliable standard errors and was used to estimate the error correction model with the Engle–Granger method.

The final results are given in Table 2.10.

In the long run freight transport performance in France is not only related to the activity in the manufacturing sector, but also to the indicators of international trade. In the short run the growth rate of freight transport is driven mainly by the growth in the manufacturing sector corrected for deviations of the long run relation.

Another example which illustrates the importance of using suitable indicators for economic activity in the freight sector can be found in Kupfer, Meersman, Onghena, and Van de Voorde (2011) which models the relation between world air freight and economic activity. As an alternative for GDP, the authors use world merchandise exports together with the share of manufactures in the value of world merchandise exports.

If one descends to a more disaggregate level of analysis, the relation between freight transport and economic activity can be more elaborated. Janssens et al. (2003) give a detailed analysis of the relation between on the one hand the number of tonnes loaded and unloaded in the port of Antwerp, and, on the other hand, economic activity, for a number of commodity groups for the period 1950–1990. As relevant indicators they use imports and exports of the Belgo-Luxembourg Economic Union (BLEU). They also introduce quay length as indicator for port capacity and real wages in the port of Antwerp.

The relations were estimated with annual data covering the period 1951–1990 for 10 commodity types using error–correction models. The resulting trade-elasticities are reported in Table 2.11.

At this disaggregate level it is clear that the relation between throughput and trade differs over the good categories. This can be explained to a large extent by the fact that for some good categories, the growing international trade of the BLEU is mainly with neighbouring countries without having to call at the port of Antwerp.

If one wants to continue using GDP as an indicator, it will be necessary to use estimation methods which allow for structural changes and which can incorporate changes in the freight intensity of the economies under consideration. One potential

9. Both indicators are defined by the OECD. For the calculation of the export performance of goods, the growth rate of the total exports of goods of a particular OECD country has been divided by the growth rate of the imports of goods of the rest of the OECD. The import penetration rate is measured as the ratio between imports and domestic demand. It shows to what degree domestic demand is satisfied by imports.

Table 2.10: ECM-estimation for freight transport in France

$$\Delta \ln TOT_t = \underset{(0.030)}{-0.082}\,\Delta D82_t + \underset{(0.026)}{0.145}\,\Delta D95_t + \underset{(0.093)}{1.024}\,\Delta \ln MANU_t$$
$$\underset{(0.183)}{-0.53}\,(\ln TOT_{t-1} - \underset{(0.48)}{9.52} + \underset{(0.013)}{0.128}\,D82_{t-1} - \underset{(0.020)}{0.192}\,D95_{t-1} - \underset{(0.106)}{0.283}\,EXPP_{t-1}$$
$$- \underset{(0.003)}{0.0092}\,IMPP_{t-1} - \underset{(0.117)}{0.523}\,\ln MANU_{t-1})$$

$R^2_{adj} = 0.81$ $DW = 1.72$ $\text{logLikelihood} = 89.21$ Sample : 1971 – 2009

With

lnTOT	log total freight t-km
lnMANU	log index of manufacturing production
EXPP	export performance (goods)
IMPP	import penetration rate (goods)
D82, D95	dummy variables for years in which OECD reports breaks in the freight data series

figures between brackets are standard errors

Table 2.11: Trade elasticity of port throughput in the port of Antwerp for different goods categories based on data for 1950–1990

Goods category		**Import elasticity of tonnes unloaded**		**Export elasticity of tonnes loaded**	
		short run	**long run**	**short run**	**long run**
cat.0	Agricultural products and living animals	0.33	0.14	0.59	0.93
cat.1	Food and cattle feeding products	0.38	0.29	0.31	0.86
cat.2	Solid mineral fuels	1.75	0.74	0.99	0.89
cat.3	Petroleum and derivatives	1.04	0.88	0.95	0.59
cat.4	Ore, metal waste, roasted pyrite	1.22	0.97	0.45	0.90
cat.5	Iron, steel and non-ferrous metal	1.18	0.97	0.50	0.53
cat.6	Raw minerals and products; building material	1.00	0.51	0.58	0.27
cat.7	Fertilisers	0.80	0.94	0.77	0.22
cat.8	Chemicals	0.55	0.68	0.17	0.39
cat.9	Vehicles, machinery and remaining	0.86	1.25	0.15	0.34
Total		1.26	1.15	0.56	0.72

Source: Adapted from Janssens, Meersman, and Van de Voorde (2003).

direction is using methods which allow for discrete or continuous time-varying coefficients. The coefficients can be modelled as a function of time, but it will be more interesting to model them as a function of variables which explain the changing relation between freight and GDP.

The problem with most of the aforementioned approaches is that they have to rely upon long-term forecasts of the economic indicators. As they are not always available, one has to use or build models for generating these forecasts or one has to work with long-term scenarios for the evolution of the indicators as is already applied in a number of countries for different modes.

2.5. Conclusion

Freight transport is a derived demand. In a world of specialisation and trade inputs have to be transported to production units the output of which has to be delivered to other production units, distribution centres, wholesalers, warehouses, retailers and consumers. Commodities and goods have to be transported throughout this entire chain.

Traditionally this relation between economic activity and freight transport is used to make forecasts of future aggregate freight flows and volumes. Usually GDP is used as an indicator for economic activity in a region or a country. It is the basis for national income, and hence also for consumption. National and international organisations collect data on GDP and also make forecasts of GDP. This explains to some extent the reason for using GDP as indicator for future freight transport.

However, it is difficult to find a common trend over the past 40 years on an aggregate level between the evolution of freight and of GDP. As a consequence the relation cannot be used for long-term forecasting. There are several reasons for this instability of the freight–GDP relation. The composition of GDP changed and is still changing, some methods to link freight transport to GDP are not suited and the link between freight transport and economic activity itself is changing due to policies (decoupling) and business behaviour (time based competition, labour vs. transport costs …).

This imposes more complexity on demand models which can cope with the evolutions in freight transport. At an aggregate level panel data techniques are very promising because they do not only allow for more diversification of the relation between freight transport and economic activity on a country level, but also shifts over time can be introduced. Also varying and stochastic parameter models can open up a number of new perspectives in quantifying the relation between freight transport and economic activity. The main purpose is to estimate not only that relation, but also the relation between the elasticities and a number of factors which have an impact on the magnitude of the elasticities.

To get a grip on the changing relation between the economy and freight transport, disaggregate models are necessary. They require not only information for different commodity groups, but have to be built starting from a detailed microeconomic underpinning. Behavioural shipper and carrier models will help to understand the decision and choice processes that determine which, how much and in what way commodities will be moved. In this perspective, transport can be considered as part of the logistics process and a kind of production factor. The basic idea is that companies do not regard as output the mere products of their production units, but the commodities arriving at their final destination within a certain time period. To realise this output not only traditional inputs such as labour and capital are needed, but also transportation becomes a crucial production factor. This allows for an analysis of the relation between the economic activity of a company or group of companies and the generalised cost of the logistics chain.

To be successful, this approach requires a constant updating of insights into new and innovative developments in the field of logistics, changes in the respective roles of the various actors involved, such as the authorities at the national and, even more so, supranational levels, and capacity policy at all levels. Add to this the single principle that is increasingly put forward within the context of the EU: fair and efficient pricing, in combination with a battery of supporting measures. Ultimately, then, not much has changed since Marvin Manheim published his Fundamentals of Transportation Systems Analysis back in 1979.

References

Baldwin, R. (2011). *21st century regionalism: Filling the gap between 21st century trade and 20th century trade rules*. CEPR Policy Insight No. 56. Retrieved from http://www.cepr.org/

Beaumont, H., Preston, J., & Shires, J. (1996). *SORT-IT Strategic Organisation and Regulation in Transport, Summary of country reports, Version 0.2*. Research Project for the Commission of the European Communities, DGVII. Institute for Transport Studies, University of Leeds.

Blauwens, G., De Baere, P., & Van de Voorde, E. (2012). *Transport Economics* (5th rev. ed.). Antwerp: De Boeck Ltd.

Brown, R. L., Durbin, J., & Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society B*, *37*, 149–163.

Coppens, F., Lagneaux, F., Meersman, H., Sellekaerts, N., Van de Voorde, E., Van Gastel, G., … Verhetsel, A. (2007). *Economic impact of port activity: A disaggregate analysis*. The Case of Antwerp. Working Paper No. 110. National Bank of Belgium, Brussels. Retrieved from http://www.nbb.be/doc/ts/publications/wp/wp110En.pdf

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. New York, NY: Oxford University Press.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276.

European Commission. (2009). *First intermediate report 'evaluation of cohesion policy programme 2000–2006, work package transport'*. Retrieved from http://ec.europa.eu/transport/themes/infrastructure/doc/2009_intermediate_report_dg_regio.pdf

European Commission. (2011). *Roadmap to a single European transport area – Towards a competitive and resource-efficient transport system*. (COM (2011) 144 final of 28 March 2011). Retrieved from http://ec.europa.eu/transport/themes/strategies/2011_white_paper_en.htm

European Commission. (2012). *EU Transport in figures*. Statistical pocketbook. Retrieved from http://ec.europa.eu/transport/facts-fundings/statistics/doc/2012/pocketbook2012.pdf

Hise, R. T. (1995). The implications of time-based competition on international logistics strategies. *Business Horizons*, *38*, 39–46.

IMF. (2012). *World economic outlook database*. Retrieved from http://www.imf.org/external/pubs/ft/weo/2012/02/weodata/index.aspx

ITU World Telecommunication. (2012). *ICT Indicators database*. Retrieved from http://www.itu.int/ITU-D/ict/statistics/

Janssens, S., Meersman, H., & Van de Voorde, E. (2003). Port throughput and international trade: Have port authorities any degrees of freedom left? In R. Loyen, E. Buyst & G. Devos (Eds.), *Struggling for leadership: Antwerp–Rotterdam port competition between 1870–2000* (pp. 91–113). New York, NY: Physica-Verlag.

Kupfer, F., Meersman, H., Onghena, E., & Van de Voorde, E. (2011). World air cargo and merchandise trade. In R. Macário (Ed.), *Critical issues in air transport economics and business* (pp. 98–111). London: Routledge.

Manheim, M. L. (1979). *Fundamentals of transportation systems analysis*. Cambridge, MA: The MIT Press.

Manheim, M. L. (1999). The next challenge in transportation research: Enhancing the application of the mind. In H. Meersman, E. Van de Voorde & W. Winkelmans (Eds.), *World transport research: Transport modes and systems* (Vol. 1, pp. ix–xi). Oxford: Pergamon.

Meersman, H., & Van de Voorde, E. (1999). Is freight transport growth inevitable? In ECMT (Ed.), *Which changes for transport in the next century?* (pp. 23–48). Paris: OECD Publications Service.

Meersman, H., & Van de Voorde, E. (2001). International logistics: A continuous search for competitiveness. In A. Brewer, D. Hensher & K. Button (Eds.), *Handbook of logistics and supply-chain management* (pp. 61–77). New York, NY: Pergamon.

Meersman, H., & Van de Voorde, E. (2006) *Dynamic ports within a globalised world*. 17th International ECMT/OECD Symposium on Transport Economics and Policy, Benefiting from Globalisation. Transport Sector Contribution and Policy Challenges, 26 pp. (CD-ROM).

Meersman, H., & Van de Voorde, E. (2008). The relationship between economic activity and freight transport. In M. Ben-Akiva, H. meersman & E. Van de Voorde (Eds.), *Recent developments in transport modelling* (pp. 69–92). Bingley, UK: Emerald Group Publishing Limited.

Meersman, H., Van de Voorde, E., & Vanelslander, T. (2003). The industrial economic structure of the port and maritime sector: An attempt at quantification. In Cetena/Atena (Eds.), *NAV2003: International Conference on Ship and Shipping Research, Conference Proceedings (Volume III)* (pp. 2.3.1–2.3.19). Palermo: Cetena.

OECD Statistics. Retrieved from http://stats.oecd.org/

Sturgeon, T. J., & Memedovic, O. (2010). *Mapping global value chains: Intermediate goods trade and structural change in the world economy*. UNIDO Development policy and strategic research branch. Working Paper 05/2010. Retrieved from http://www.unido.org/fileadmin/media/documents/pdf/Publications/110923/WP052010_Ebook.pdf

Suykens, F. (1986). Ports should be efficient (even when this means that some of them are subsidized). *Maritime Policy and Management*, *13*(2), 105–126.

UNCTAD. (1995, 1998, 2005, 2012). *World Investment Report*. Geneva: United Nations Publications.

Verbeek, M. (2012). *A guide to modern econometrics* (4th ed.). Wiley.

Wood, D. F., Barone, A., Murphy, P., & Wardlow, D. L. (1995). *International logistics*. New York, NY: Chapman and Hall.

Young, R. R. (1995). Book review international logistics. *Transportation Journal*, *34*, 72–73.

Chapter 3

A Multimodal Elastic Trade Coefficients MRIO Model for Freight Demand in Europe

Ennio Cascetta, Vittorio Marzano, Andrea Papola and Roberta Vitillo

Abstract

The chapter proposes the specification of a multi-regional input-output (MRIO) model with elastic trade coefficients and a multimodal freight supply model at an European geographic scale. After a brief review of the MRIO modelling approach, some important theoretical issues related to both demand and supply modelling were discussed.

Specifically, from the demand side, an elastic trade coefficient MRIO model is presented and some relevant macroeconomic feedbacks incorporable into the model are discussed. From the supply side, a critical review of the complexity of the multimodal freight networks and of the corresponding modelling requirements is presented. From a practical standpoint, the implementation of a MRIO model at European level is reported, describing in particular both the database used to feed the model and the supply model applied for the calculation of transport impedances required by the trade coefficient model. Finally, simple applications of the implemented MRIO model with elastic trade coefficients are presented, and some conclusions and research developments are drawn.

Keywords: MRIO models; multimodal freight supply modeling; Euro-Mediterranean freight DSS

3.1. Introduction

The appraisal of plans and projects for the development of transport infrastructures and services is a key topic for researchers and practitioners, both at a national and

Freight Transport Modelling

ISBN: 978-1-78190-285-1

international scale. This is particularly crucial in the view of economic and industrial systems, whose dynamics are rapidly and strongly influenced by freight transport systems policy and governance, leading to substantial effects on the market accessibilities and on the competitiveness of national economies. In turn, freight demand is substantially determined by the overall economic structure of the countries (or regions) under analysis. In that respect, a significant amount of models and methodologies has been proposed over the years, both for quantifying the economic impacts of freight transport plans on the economy and for a proper estimation of freight flows given the underlying economic/industrial structure.

Within this context, the multi-regional input-output (MRIO) approach is well-developed in its theoretical foundations and widely adopted in the practice: up-to-date and exhaustive reviews are reported amongst others in Juri and Kockelman (2005) and Cascetta, Marzano, and Papola (2008). Interestingly, even if the original extension of the Leontief input-output model to the case of multiple regions with interacting economies – provided by Izard (1951) – was mainly aimed at modelling the economic system, several significant improvements were made over the years to better model transport demand and related feedbacks:

- The Chenery-Moses (see for instance Miller & Blair, 1985) formulation simplified the Izard (1951) model, by introducing the less data-demanding trade coefficient specification adopted throughout the paper;
- The integration with random utility models for the estimation of trade coefficients (RUBMRIO), addressed amongst others in De La Barra (1989), Cascetta (2001) and Jin, Kockelman, and Zhao (2003);
- The taxonomy of several MRIO formulations accordingly with the assumptions on import/export flows to/from the external of the study area, as analysed in Marzano and Papola (2004) and Cascetta et al. (2008);
- The incorporation of selling price estimation/adjustment procedures through fixed-point approaches, as proposed for instance by Cascetta (2001) and Zhao and Kockelman (2004).

As a result, MRIO models are a reliable and stable modelling tool, useful as a freight demand model (i.e. providing o-d flows as a function of economic and transport variables) and as a basic macroeconomic impact model (i.e. providing production and GDP as a function of economic and transport variables), as described in detail by Cascetta et al. (2008). In addition, they are easily incorporable within transport-oriented decision support systems (DSS), as it will be discussed in detail in Section 3.3.

Notwithstanding, amongst the DSS implemented in Europe for transport modelling and appraisal, there are no significant and advanced applications of MRIO models, in spite of its acceptable compromise between computational effort and transport modelling requirements. On the other side, there are valuable complex macro-economic models mostly oriented to transport-based cost-benefit analyses, such as IASON (2004) and SASI (Wegener, 2008), which are characterised by a commodity disaggregation and a transport component not entirely appropriate for dealing with the inherent complexity of freight transport systems.

Starting from these premises, this chapter proposes the specification of a random utility based MRIO (RUBMRIO) model with elastic trade coefficient and a multimodal freight supply model at an European geographic scale. In more detail, Section 3.2 describes the mathematical framework underlying the MRIO model. Section 3.3 reports the implementation of the MRIO model at the European level, describing in particular the database used to feed the model and the supply model applied for the calculation of transport impedances required by the trade coefficient model. Finally, Section 3.4 presents simple applications of the implemented MRIO model with elastic trade coefficients and Section 3.5 draws some conclusions and research developments.

3.2. The MRIO Model: Mathematical Framework

3.2.1. Fixed Trade Coefficients Model Specification

Following a taxonomy proposed by Marzano and Papola (2004) and Cascetta et al. (2008), several basic formulations of the MRIO model can be specified, according to different hypotheses on the import/exports between each zone within the study area and the rest of the world. For instance, assuming endogenous imports – that is an increase in final demand satisfied by an increase in internal (i.e. within-zone) production, in imports from the other zones in the study area and also in imports from the rest of the world – and endogenous exports – that is international exports satisfied both by internal production and imports – the following model formulation can be derived:

$$\mathbf{X} = (\mathbf{I} - \mathbf{TA})^{-1}[\mathbf{T}(\mathbf{Y} + \mathbf{Y_{RW}})] \tag{3.1}$$

where $\mathbf{X}$ is the vector of production values X_i^m for each zone i and for each commodity sector m, $\mathbf{I}$ the identity matrix, $\mathbf{T}$ the matrix of trade coefficients t_{ij}^m representing the acquisition percentages of goods of sector m in zone j from zone i, $\mathbf{A}$ the matrix of technical coefficients a_i^{mn} representing the value of goods of sector m needed for producing one value unit of goods of sector n in zone i, $\mathbf{Y}$ and $\mathbf{Y_{RW}}$ the vectors respectively of final demand Y_i^m and exports outside the study area $Y_{\mathbf{RW}Ti}^{\ m}$ for sector m in zone i. For different specifications the reader may refer to Cascetta et al. (2008): the specification (3.1) will be adopted throughout the paper without loss of generality.

All variables included in the model (3.1) can be directly obtained from the i-o tables of each zone but trade coefficients, whose estimation requires knowledge of trade flows between each pair of countries within the study area and to each country from the rest of the world. In that respect, Marzano and Papola (2004) showed that the internal trade coefficient for each zone i and for each sector m (i.e. the acquisition percentages t_{ii}^m) can be derived directly from the i-o table of zone i, through the formula:

$$t_{ii}^m = 1 - \frac{J_{RWi}^m + J_{SAi}^m}{\sum_n K_i^{mn} + Y_i^m + Y_{RWi}^m} \tag{3.2}$$

where J_{RWi}^m and Y_{RWi}^m are respectively the imports and the exports of zone i for sector m with reference to the rest of the world, J_{SAi}^m the imports of zone i for sector m from the other zones within the study area, Y_i^m the final demand of sector m in i and K_i^{nm} the intermediate re-usage of goods of sector m for production of sector n in i. As it will be shown in Section 3.3, Eq. (3.2) can be coupled effectively with observed trade flows between zones within the study area, in order to obtain a full observation of trade coefficient for the base year.

As mentioned in the introduction, the MRIO model (3.1) has a twofold practical relevance. Firstly, once calculated the production vector $\mathbf{X}$ through (3.1), the o-d matrix in value $\mathbf{N}$ — whose generic element N_{ij}^m represents the trade flow in value of commodity m between zones i and j — can be obtained through the relationship:

$$\mathbf{N} = \mathbf{TA}\mathrm{dg}(\mathbf{X}) + \mathbf{T}\mathrm{dg}(\mathbf{Y}) \tag{3.3}$$

wherein $\mathrm{dg}(\cdot)$ is the diagonalisation operator. The matrix $\mathbf{N}$ can be in turn transformed into a freight o-d matrix in quantities (e.g. tonnes/year) based on the availability of observed value-to-quantity transformation coefficients for the base year. Secondly, since the unit value added va_i^m for sector m in zone i is given by:

$$va_i^m = 1 - \sum_n a_i^{nm} \tag{3.4}$$

the gross domestic product GDP_i of zone i can be calculated by definition as:

$$GDP_i = \sum_m X_i^m va_i^m = \sum_m X_i^m \left(1 - \sum_n a_i^{nm}\right) \tag{3.5}$$

3.2.2. *Modelling Elastic Trade Coefficients*

The key conceptual vehicle incorporating the freight transport system into the basic MRIO (3.1) is the explicit modelling of trade coefficients as a function of the performances of the freight transport system. The assumption of elastic trade coefficients allows simulating explicitly the effects of changes of transport system performances both on trade patterns (e.g. updating o-d matrix via Eq. (3.3)) and on the macroeconomic fundamentals (e.g. calculating the variation in GDP through Eq. (3.5)). The development of MRIO model with elastic trade coefficients started in the context of the transport-land use interaction models: a state of the art in this field can be found, for instance, in the contributions of Min, Kockelman, Zhao, and Jin (2001) and Timmermans (2003). As reported amongst others by Cascetta (2001), the rationale is to model variations in trade coefficients through a discrete choice model, for example a Multinomial Logit specification:

$$t_{ij}^m = \frac{\exp(V_{ij}^m/\theta^m)}{\sum_k \exp(V_{kj}^m/\theta^m)} \tag{3.6}$$

wherein V_{ij}^m represents the systematic utility of acquiring inputs of sector m from zone i to use it in j and θ^m the Logit variance parameter.

From a behavioural standpoint, the choice dimension for each economic agent in j is represented actually by the choice of an 'elementary origin' within zone i. Therefore, the theory of distribution models with size functions (Ben-Akiva & Lerman, 1985; Cascetta, 2009; Richards & Ben-Akiva, 1975) can be suitably adapted to this choice context. According to that, zone i can be seen as a compound alternative composed of the aggregation of a number EO_i^m of elementary alternatives (i.e. elementary origins) for sector m and, therefore, a Nested Logit model may be usually used to model t_{ij}^m, yielding:

$$t_{ij}^m = \frac{\exp(V_{ij}^m/\theta^m + \delta^m Y_i^m)}{\sum_k \exp(V_{kj}^m/\theta^m + \delta^m Y_k^m)} = \frac{\exp[(V_{ij}^m + s_i^m)/\theta^m]}{\sum_k \exp[(V_{kj}^m + s_k^m)/\theta^m]} \tag{3.7}$$

wherein the systematic utility V_{ij}^m is common to all the elementary alternatives in zone i and s_i^m is the expected maximum perceived utility related to the choice of the elementary origin within i, given by:

$$s_i^m = \theta_{eo}^m Y_i^m = \theta_{eo}^m \ln \sum_{e=1}^{EO_i^m} \exp\left(V_{e|i}^m/\theta_{eo}^m\right) \tag{3.8}$$

wherein $V_{e|i}^m$ is the systematic utility of the elementary origin e within i and θ_{eo}^m the related Gumbel variance parameter. Consistently, in Eq. (3.7) $\delta_m = \theta_{eo}^m/\theta_m$. Eq. (3.8) can, in turn, be rewritten as:

$$s_i^m = \overline{V}_i^m + \theta_{eo}^m \ln EO_i^m + \theta_{eo}^m \ln\left[\frac{1}{EO_i^m}\sum_{e=1}^{EO_i^m} \exp\left[\left(V_{e|i}^m - \overline{V}_i^m\right)/\theta_{eo}^m\right]\right] \tag{3.9}$$

where $\overline{V}_i^m$ is the average utility of all elementary origins for sector m within i. Notably, all terms but the last[1] in the right-hand side of Eq. (3.9) can be easily determined. For instance, assuming that $V_{e|i}^m$ is given by a mass attribute M_e (e.g. the number of employees in the elementary origin e), it occurs:

$$\overline{V}_i^m = \frac{1}{EO_i^m}\sum_{e=1}^{EO_i^m} V_{e|i}^m = \frac{1}{EO_i^m}\sum_{e=1}^{EO_i^m} M_e = \frac{M_i}{EO_i^m} \tag{3.10}$$

being M_i the corresponding mass value for the overall zone i. In addition, if the number of elementary origins EO_i^m is unknown, a proper size function can be specified as a proxy.

1. Note that this term is a measure of the dispersion of systematic utilities of the elementary origins within i and disappears if those utilities are all equal.

Interestingly, whatever specification of the trade coefficient model (e.g. either (3.6) or (3.7)), the systematic utility of each zone/elementary origin should be expressed ideally as a function of: selling prices b_i^m of sector m in region i, transport costs (intended in a broad sense) d_{ij}^m from i to j, aforementioned size variables, possibly other dummies typically adopted in the literature[2] aimed at capturing specific effects. Such variables may help addressing one of the main issues in the estimation of the trade coefficient model, that is the need of reproducing the relative magnitude of the internal trade coefficient (i.e. t_{ii}^m) if compared to the other trade coefficients (i.e. t_{ij}^m with $i \neq j$). Furthermore, as broadly evidenced by the literature on distribution models (see for instance Cascetta, 2009; Ortuzar & Willumsen, 2001), the presence of size variables and dummies – if used parsimoniously and designed effectively – may lead to a more stable model with realistic elasticities. According to this general specification, two main operational issues in the implementation of the MRIO model with elastic trade coefficients arise, respectively the calculation of the selling prices b_i^m and of the transport costs d_{ij}^m. Both aspects will be discussed in more detail in Section 3.2.3 and Section 3.3, respectively.

In this framework some trade coefficient model specifications presented in the literature can be reviewed. De La Barra (1989) and Hunt (1993) set V_{ij}^m equal to minus the acquisition cost of m from i, c_{ij}^m, that is the sum of the selling price of m in i, b_i^m, and the generalised cost d_{ij}^m for transporting m from i to j, without any attractiveness or size variable:

$$V_{ij}^m = -c_{ij}^m = -(b_i^m + d_{ij}^m) \tag{3.11}$$

Cascetta, Di Gangi, and Conigliaro (1996) and Cascetta and Iannò (1998) estimated a trade coefficient model wherein d_{ij}^m is calculated as a mode choice logsum and production values are introduced as size attributes, but selling prices are not taken into account in their model. Jin et al. (2003) applied the specification (3.11) with a mode choice logsum for d_{ij}^m and provide an aggregate estimation of model parameters, which show a large variability amongst sectors. Marzano and Papola (2004) proposed the following expression for V_{ij}^m:

$$V_{ij}^m = \begin{cases} \beta_1^m d_{ij}^m + \beta_2^m b_i^m + \beta_3^m Y_{SAi}^m & \text{if } i \neq j \\ \beta_1^m d_{ii}^m + \beta_2^m b_i^m + \beta_4^m A_i^m & \text{if } i = j \end{cases} \tag{3.12}$$

where d_{ij}^m is calculated as mode choice logsum[3] and the mass attributes are differentiated as a function of the origin zone. In more detail, if the acquisition zone

2. Relevant examples may include a same zone dummy (i.e. equal to 1 if $i=j$), a common boundary dummy (i.e. equal to 1 if i and j share a border), trade agreement specific dummies (i.e. equal to 1 if i and j have a specific commercial and/or trade agreement) and so on.

3. Road transport is assumed to be the only transport mode available for intra-zonal shipments, therefore d_{ii}^m is the generalized transport cost by road calculated based on a distance given by the average radius of the zone.

differs from the destination zone, the total export Y_{SAi}^{m} of the acquisition zone towards all the other zones in the study area is used as proxy of its availability of sector m, whilst if the acquisition zone coincides with the destination zone, the following variable is used as proxy of the total internal availability of sector m:

$$A_i^m = X_i^m + J_{RWi}^m - Y_{SAi}^m - Y_{RWi}^m \tag{3.13}$$

where the symbols were previously defined. The original specification adopted in this chapter for the applications presented in Section 3.4 is given by the following equation:

$$V_{ij}^m = \beta_1^m Y_{ij}^m + \beta_2^m b_i^m + \beta_3^m \ln X_i^m + \beta_4^m \delta_{SZ} \tag{3.14}$$

wherein Y_{ij}^m is the logsum of the mode choice context for the o-d pair i-j and the sector m, b_i^m is expressed in €/tonne, X_i^m in k€ and δ_{SZ} is a 'same zone' dummy (i.e. equal to 1 if $i = j$). The model (14) was estimated through an aggregated GLS procedure on the basis of 2011 COMEXT data (see Section 3.3.1), leading to the values reported in Table 3.1.

It is worth noting that all the coefficient have the expected sign, i.e. positive proportionality of logsum and production and negative proportionality of selling prices; clearly, there are no prior expectations for the sign of the same-zone dummy variable. Furthermore, a comparison of logsum coefficients across commodities reveals robust results, for example higher elasticities to transport costs for perishable and high-value goods.

3.2.3. Estimation of Selling Prices

As proposed by Cascetta (2001), and as mathematically addressed by Zhao & Kockelman (2004), selling prices b_i^m can be identified endogenously trough a fixed point specification by coupling Eq. (3.6) with further equations defining selling prices as a function of acquisition costs and technical coefficients, and in turn acquisition costs as a function of trade coefficients, selling prices and transportation costs. More specifically, based on such equations, a double-step fixed point approach for the estimation/application of the RUBMRIO model can be adopted (see for instance Cascetta et al., 2008).

Two main drawbacks affect the endogenous calculation of selling prices. The first is related to theoretical issues in the structure of some of the equations underlying the fixed point approach. This aspect is still under current investigation, and opens interesting perspectives towards new enhanced and effective RUBMRIO specifications. The second relates to the practical interpretation of the selling prices, which actually cannot be defined as 'prices' in a proper sense, being rather country- and commodity-specific dummies. For this reason, the applications proposed in Section 3.5 are based on a MRIO model applied with elastic trade coefficients but with fixed selling prices, observed from statistical data.

Table 3.1: Results of trade coefficient model estimation (missing figures: parameters not significant at 95% level)

Commodity sector (CPA nomenclature)	**β_1**	**β_2**	**β_3**	**β_4**
Products of agriculture, hunting and related services	5.839	–	6.000	– 0.217
Products of forestry, logging and related services	2.366	– 1.403	2.114	–
Fish and other fishing products; services incidental of fishing	0.430	–	0.665	– 0.236
Coal and lignite; peat	0.055	– 6.000	0.182	1.621
Crude petroleum and natural gas	0.351	– 0.108		0.283
Uranium and thorium ores	0.990	– 0.990	0.990	–
Metal ores	0.152	– 0.749	0.125	0.921
Other mining and quarrying products	1.079	– 5.619	0.714	–
Food products and beverages	3.110	– 6.000	4.216	– 0.284
Tobacco products	0.033	–	0.064	2.153
Textiles	0.062	–	0.448	1.117
Wearing apparel; furs	0.052	–	0.220	1.347
Leather and leather products	0.073	– 0.004	0.372	0.913
Wood and products of wood and cork (except furniture)	0.842	– 0.295	0.570	0.004
Pulp, paper and paper products	0.484	–	0.892	– 0.285
Printed matter and recorded media	1.924	– 0.311	1.107	–
Coke, refined petroleum products and nuclear fuels	1.728	– 6.000	2.747	– 0.409
Chemicals, chemical products and man-made fibres	0.035	–	0.279	1.736
Rubber and plastic products	0.946	– 0.184	1.497	– 0.460
Other non-metallic mineral products	5.176	– 1.214	6.000	– 0.287
Basic metals	5.690	–	6.000	– 0.839
Fabricated metal products, except machinery and equipment	4.157	– 3.167	6.000	– 0.404
Machinery and equipment n.e.c.	0.067	– 0.013	0.385	1.181
Office machinery and computers	2.670	– 3.026	3.408	3.538
Electrical machinery and apparatus n.e.c.	0.056	– 0.024	0.490	1.255
Radio, television and communication equipment and apparatus	0.098	– 6.000	5.674	1.271
Medical, precision and optical instruments, watches and clocks	2.809	– 3.683	2.542	5.069
Motor vehicles, trailers and semi-trailers	0.065	– 0.017	0.429	1.032
Other transport equipment	0.154	– 2.554	–	1.098
Furniture; other manufactured goods n.e.c.	0.043	– 0.012	0.155	1.714

In that respect, it is important to underline that observing selling prices is not always possible, being this possibility strictly dependent on the study area and on the zonization of the model. In general, when dealing with European-based MRIO models with national zones, the already mentioned EUROSTAT COMEXT database provides trade data between each pair of zones i and j in value and quantity, and therefore an estimation of the unit selling price can be determined. Importantly, it should be noted that, in accordance with the so-called mirror trade theory, trade flows in value and in quantity between i and j are recorded twice in the COMEXT database, once as exports from i to j and once as imports of j from i: however, exports are conventionally calculated FOB and imports CIF,[4] therefore selling prices should be estimated using export values.

3.2.4. Some Equilibrium Feedbacks in the RUBMRIO Model

The double-step estimation/application procedure for the calculation of the selling prices mentioned in Section 3.2.3 and described in Cascetta et al. (2008) is only an example of the possible equilibriums which may arise depending on the chosen MRIO specification.

For instance, if mass attributes are expressed as a function of production values or other input-output variables, such as in the specification (3.12)–(3.13), a fixed point problem arises, in which a feedback between the MRIO model (3.1) and the trade coefficients model (3.6) is introduced. Furthermore, in most of the practical applications of MRIO models to future scenarios, exports Y^m_{RWi} and imports J^m_{RWi} to/from the rest of the world may be determined through a gravity model, in which the GDP of each zone is normally included as explanatory variable (see for instance Gallo, Simonelli, & Marzano, 2012; Kepaptsoglou, Tsamboulas, Karlaftis, & Marzano, 2009). Again, this induces a further fixed point since total GDP for each zone depends on the outputs of the MRIO model itself via Eq. (3.5). Those equilibriums are currently under analysis in terms of both theoretical properties and implementation patterns.

Finally, it is worth underlining that a wide range of further feedbacks between relevant MRIO variables and external socio-economic and land-use variables/indicators is available in the literature, in the form of either equilibrium or dynamic process models: the reader may refer to the literature of transport-land use interaction models (e.g. De La Barra, 1989; Hunt, 1993) and also to transport-oriented contributions with the related references (e.g. Jin et al., 2003; Juri & Kockelman, 2005). Interestingly, such wide range of equilibrium feedbacks within the structure of the MRIO framework allows modelling various policy horizons. For instance, short run applications (i.e. within a time horizon of 1–3 years) may be carried out just through a MRIO model with trade coefficients elastic to transport

4. FOB (free on board) and CIF (cost, insurance, freight) are two common INCOTERMS regulating the way which international trade transactions can be defined.

level-of-service attributes, whilst long run applications (i.e. involving a time horizon farther than 5 years) may be modelled through MRIO specification accommodating also some of the macroeconomic feedbacks described above.

Finally, it is also worth mentioning that, once embedded the equilibrium feedbacks described above into the MRIO model, a thorough comparison/contrast of the obtained modelling results with those provided by different macro-economic approaches (see Section 3.1) will be a crucial step for defining guidelines to researchers and practitioners in performing freight demand modelling and related cost-benefit analyses.

3.3. MRIO Implementation at European Level

Following the theoretical background described in Section 3.2, this section deals with the implementation of a MRIO model at European level within the context of a wider transport DSS developed by the authors at an Euro-Mediterranean scale,[5] schematically illustrated at a glance in Figure 3.1.

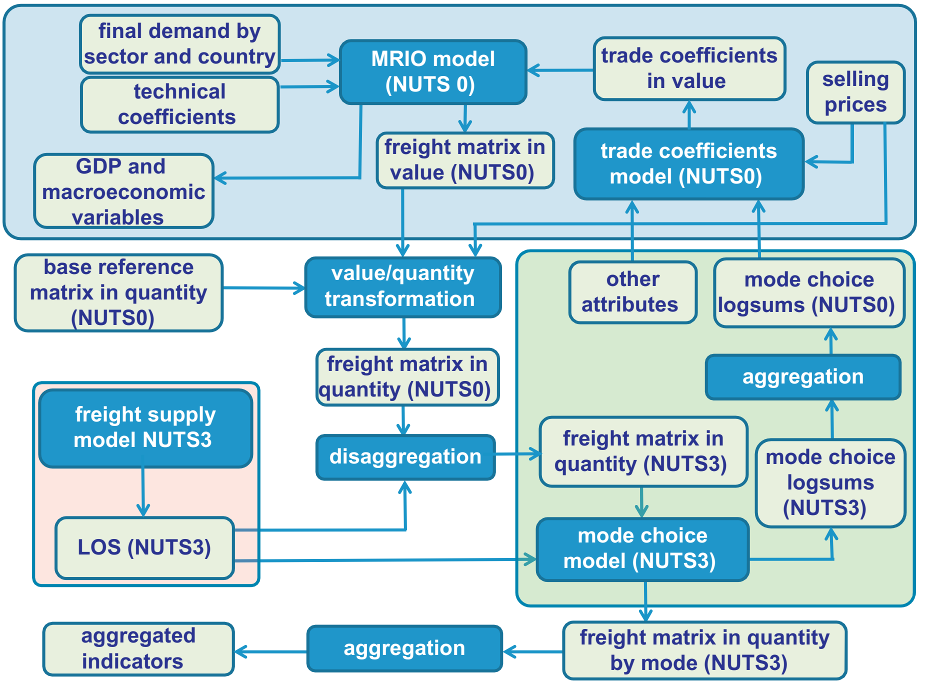

Figure 3.1: Basic structure of the proposed European MRIO-based DSS.

5. In fact, this DSS was developed by the authors in research contexts requiring a much wider study area with respect to the current MRIO application.

The DSS is made up by two main sub-components: the former is a MRIO model with elastic trade coefficients and exogenous fixed selling prices; the latter is a sequence of transport models basically able to model freight transport supply and freight mode choice. Importantly, a two-way interaction exists between these two components: the MRIO model provides o-d flows in value, eventually transformed into quantities through exogenous fixed value-to-quantity ratios, to be used as inputs for the transport models. Vice versa, the transport models calculate the transport level of service attributes which enter as inputs to the trade coefficient model, in turn feeding the MRIO model. It is worth underlining that there is a remarkable discrepancy in the geographical disaggregation of the various sub-models, also conditional on the characteristics of the corresponding available input data. For instance, the MRIO model operates at NUTS0 level[6] (i.e. countries as zones) whilst the supply and the mode choice models require a NUTS3 zonization. This implies performing aggregations and disaggregation, mainly based on gravitational models and/or on origin/destination zone masses.

The most relevant components of the DSS depicted in Figure 3.1 for the purposes of this chapter are the MRIO model with the related database and the multimodal freight supply model. The remaining components, specifically the mode choice model and the value/quantity transformation procedure, have been implemented through 'traditional' models already available in the literature, leaving as a future research step the incorporation of more complex 'logistics' models. In particular, the mode choice model is the one described by Cascetta et al. (2008) for Italy, that is a Multinomial Logit consignment model estimated on Italian disaggregate consignment data. Furthermore, for each mode (or combination of modes) the shortest path has been taken into account for calculating the corresponding modal impedances. In turn, the value/quantity transformation is performed through observed coefficients from COMEXT source for the base year (see Section 3.3.1).

Consistently, in the following, Section 3.3.1 illustrates the main characteristics of the implemented MRIO model and of the related database, whilst Section 3.3.2 deals with the multimodal European freight supply model. More details on the DSS are reported in Marzano (2010).

3.3.1. *Model Database and Specification*

This section provides some details on the implementation of the European MRIO model. In accordance with the above, a NUTS0 country-based zonization of the study area was performed: therefore, a zone will be denoted in the following as a country for the sake of simplicity. For EU countries, both data sources required for

6. The NUTS classification of the geographical units adopted by the European Union is fully reported in the EUROSTAT RAMON dataset. Basically, the NUTS0 level corresponds to the country level, whilst the most detailed NUTS3 level corresponds for instance to the Italian provinces.

the MRIO implementation (i.e. i-o tables and trade flows) are provided by EUROSTAT through the so-called ESA95 and COMEXT datasets respectively.

The ESA95 database contains input-output tables for most of the 27 EU countries, encompassing 59 commodity sectors classified accordingly with the CPA nomenclature.[7] For each country, tables are normally available for different years, therefore year 2005 has been chosen as reference horizon since it guaranteed maximum geographical coverage. As a result, 22 European countries are included in the study area[8]: Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Netherlands, Portugal, Poland, Slovak Republic, Slovenia, Spain, Sweden and United Kingdom.

The COMEXT database provides input/output flows from each EU country to/from each other country of the world. Data are available on year and monthly bases up to 2012 and are expressed both in monetary value and in quantities, with disaggregation in more than 4000 sectors accordingly with the NC8 commodity nomenclature. Therefore, in order to harmonise this dataset with the other data provided by the ESA95 i-o tables, a correspondence between the NC8 and the CPA classifications has been established, and trade flows have been aggregated in order to calculate trade coefficients for the same 59 CPA sectors. Then, trade coefficients t_{ij}^{*m} with $j \neq i$ have been determined through the expression:

$$t_{ji}^{*m} = \frac{F_{ji}^{m}}{\sum_{h} F_{hi}^{m} + F_{RWi}^{m}} \tag{3.15}$$

where F_{ji}^{m} and F_{RWi}^{m} are the flow of goods of sector m to i from j and from the rest of the world respectively, derived from the harmonised COMEXT database. Since trade coefficients (3.15) sum up to one without the internal trade coefficient (3.2), a normalisation was performed in order to achieve the condition $t_{ii}^{m} + \sum_{j \neq i} t_{ij}^{m} + t_{RWi}^{m} = 1$. Such observations on actual trade coefficients are needed for the estimation of the trade coefficient model (see Table 3.1).

Based on the described database for the base year (current scenario), and in accordance with the theoretical review reported in Section 3.2, a MRIO model with elastic trade coefficients modelled through specification (3.12)–(3.13) has been implemented. Obviously, in order to apply such trade coefficient model to the European context described above, a correspondence table between the commodity sectors in Table 3.1 and the 59 commodity sectors of the COMEXT data source must be established. Furthermore, selling prices are determined exogenously, again based on the COMEXT database, using the procedure described in Section 3.2.3.

---

7. The classification of productive activities is described in detail in the EUROSTAT RAMON dataset, together with corresponding tables which allow transformation to/from other relevant nomenclatures (e.g. NST/R).

8. An extension of such study area is quite straightforward, since other statistical bodies and public research bureau (e.g. World Trade Organization (WTO) and World Bank) provide i-o tables for the majority of the world countries. Obviously, a harmonization with the ESA95 dataset should be performed.

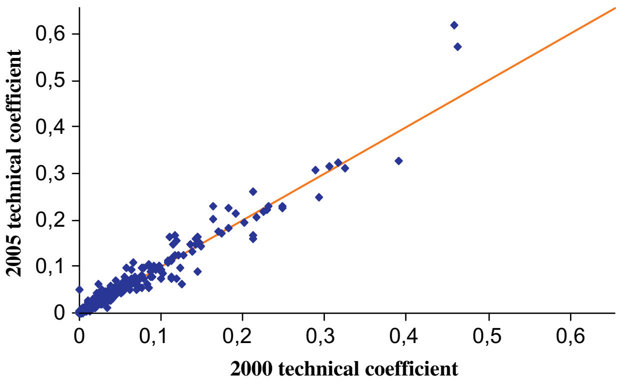

Figure 3.2: Variation of technical coefficients for Italy between 2000 and 2005.

Before continuing, it is important to stress that an investigation of the hypothesis of constant technical coefficients across years has been performed. Notably, this analysis shows the temporal variability of technical coefficients to be very scarce, at least for a time horizon of ten years: this seems to be consistent with their interpretation of coefficients representing the technical pattern of economic production of a country for each sector. By way of an example, Figure 3.2 reports the comparison of technical coefficient between years 2005 and 2000 for Italy (the situation for the other countries within the study area is very similar). Notably, the stability of their values is significant, whilst the only exhibiting a greater variation refer to the sector of production of energy.

3.3.2. Multimodal European Freight Supply Model

An important issue in the practical implementation of an European MRIO Model is related to the calculation of transport costs to be included in the trade coefficients model. Unfortunately, modelling freight supply at European level is a challenging task, not properly addressed by the existing literature, for some reasons. Firstly, there is a significant presence of non-additive transport costs for each single mode, for example the EU regulation EC 561/2006 on rest/stop times for road drivers[9] or non-additive fares/tariffs for rail and sea modes. In addition, there is need to account for freight characteristics (e.g. for road/sea Ro-Ro transport: 1 driver/2 drivers, own account/hiring, accompanied/unaccompanied and so on) impacting on the performances of each mode: this implies that, for each mode, any feasible combination of relevant characteristics (named in the following a 'supply segment') should

9. Substantially, the daily driving period shall not exceed 9 hours, with an exemption of twice a week when it can be extended to 10 hours. Furthermore, 2 drivers can sum basically their driving times.

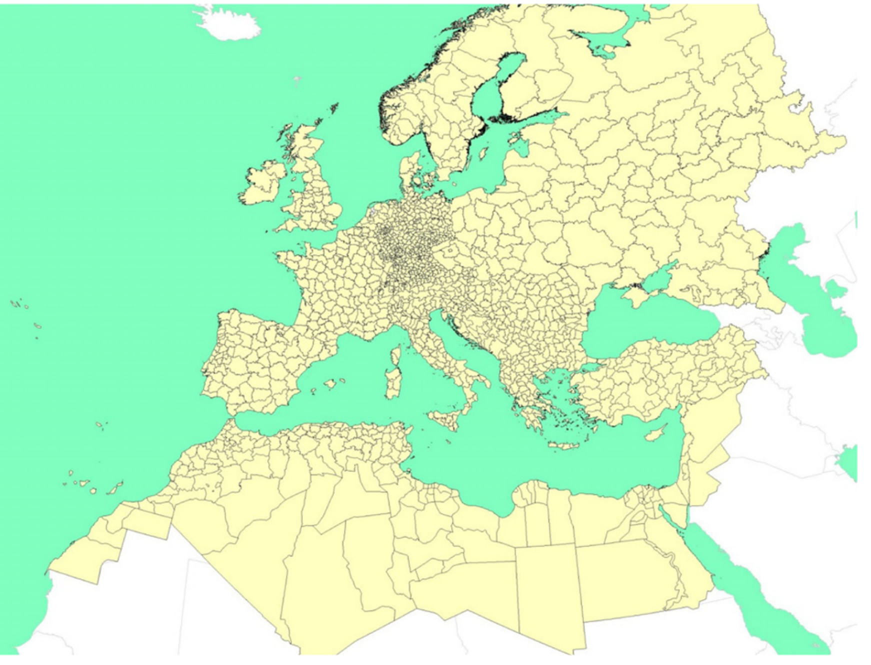

Figure 3.3: Study area and zonization of the DSS for the Euro-Mediterranean basin.

be explicitly identified, and a different supply model implemented for each supply segment. Finally, the multimodal integration of the single available modes is not straightforward, because of the already mentioned presence of single mode non additive impedances.

For these reasons, this section provides brief details about the multimodal European freight model used for the calculation of the transport impedances to be adopted into the elastic trade coefficient model. The supply model refers to a study area encompassing 57 Euro-Mediterranean countries with a zonization corresponding to the NUTS3 geographical level for EU Countries and to the administrative regional[10] level for the remaining countries. As a result, 1508 traffic zones are defined (see Figure 3.3). In accordance with this zonization, four different supply models are implemented for road, rail, maritime and inland waterways freight modes, respectively. In general, the implementation of a supply model requires the definition of the topological and of the analytical characteristics of the network: a brief review of the methodology, of the hypotheses and of the structure of the supply model for each mode is reported in each of the following sub-sections (from 3.3.2.1 to 3.3.2.4 for road, rail, sea and inland waterways, respectively). Finally, Section 3.3.2.5 deals with the integration of the single-mode supply models into a multimodal model.

10. The actual definition of the regional administrative level may obviously differ amongst countries.

3.3.2.1. Road supply model The topological model for the road mode was implemented by taking into account all relevant infrastructures representing significant connections between zones within the study area, for a total amount of 704.989 km of infrastructures, modelled through a graph made by 63.109 links and 50.870 nodes. Each road link is associated with physical characteristics used as explanatory variables in the impedance functions. For this aim, a functional classification of links into five clusters was defined: motorways, double-carriageway highways, single-carriageway national roads, regional roads, local roads. Each link is associated also with a deviousness index and a slope class, the latter derived from advanced geo-spatial analyses based on the availability of detailed altitude geo-grids.

Different analytical models are implemented for different road vehicle types: car, light commercial vehicles, medium commercial vehicles, and heavy commercial vehicles. Consistently, free flow link speeds are defined for each vehicle type, road type, and also country class. Average (i.e. recurrent) congested speeds, derived from previous consultancy studies carried out by the authors, are taken into account as well for the main European metropolitan areas.

In turn, link travel times are calculated as the sum of running times (t_r) and stopping times (t_w), the former calculated on the basis of the length of the link and of the above mentioned link speeds, the latter depending on specific link-related issues, for example customs procedures at borders, port and intermodal terminal operations, and so on. These stopping times have been validated through a detailed comparison with observed values for the Italian context and to/from some relevant EU o-d pairs. The availability of link travel times allows calculating the shortest additive[11] time path $T_{\text{od}}^{\text{add}}$ for each o-d pair, that is between each pair of o-d zones. Importantly, the already mentioned EC 561/2006 regulation on rest/stop times for road drivers leads to a further non-additive travel time component, which can be calculated easily by means of a specific algorithm providing the stopping travel time $T_{\text{od}}^{\text{stop}}$ on the basis of the additive time $T_{\text{od}}^{\text{add}}$, so as to obtain finally the total travel time $T_{\text{od}}^{\text{stop}} + T_{\text{od}}^{\text{add}}$. Importantly, $T_{\text{od}}^{\text{stop}}$ is a non decreasing function of $T_{\text{od}}^{\text{add}}$, i.e. it is a sub-additive cost function: practically, this means that the shortest time-additive path is the same as the overall shortest path, being just the overall path impedance different.

Similarly, the calculation of travel cost is carried out by considering the following components: time-dependent costs (e.g. drivers, vehicle amortisation, value of travel time savings); length-dependent costs (e.g. parametric tolls, fuel, maintenance); other link-specific costs (e.g. border duties). The cost of the return trip has been explicitly accounted for by adding an extra cost on 'unbalanced' o-d relationships (i.e. relationships characterised by unbalanced flows). Information about unbalanced o-d relationships is available from the international import-export statistics previously mentioned. Other link-specific costs related to border duties and vignettes, or to

11. A link time component is said to be additive when the total path time component can be calculated as the sum of the time components of all links belonging to that path. This is for instance the case of the running and waiting times.

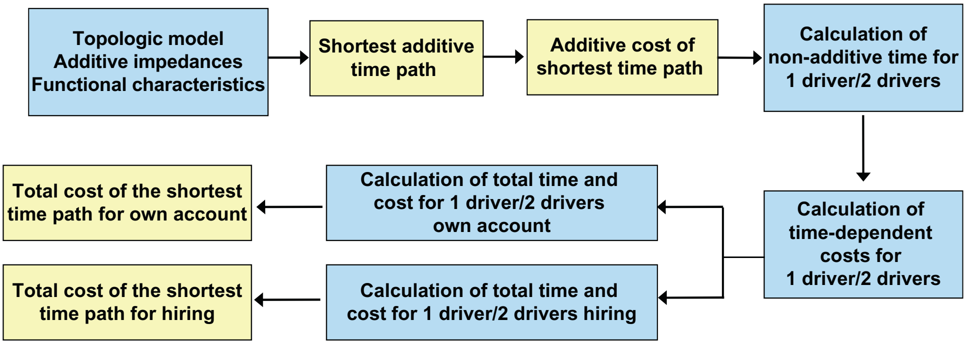

Figure 3.4: Example of calculation of impedances for road freight using the DSS.

link-specific fares, are taken into account, together with the presence of state aids for road transport in some European countries which may lead to reduced costs. In general, all cost parameters were determined on the basis of various studies and surveys available in Italy and in Europe (e.g. CONFETRA, DG-TREN, Italian Ministry of Transport), relevant literature and proprietary data. In general, the calculation of costs depends on the hypothesis of either own transport or hiring: for instance, VTTS is not considered for hiring, whilst some other costs (e.g. amortisation and insurances) are not perceived by the own transport supply segment.

Summarising, thanks to the aforementioned sub-additivity property, for any given o-d pair the shortest time-additive path can be calculated, then the sub-additive impedances can be calculated through post-processing of the additive shortest paths, and finally relevant costs depending on the specific supply segment can be calculated as well: a graphical example of such calculation is reported in Figure 3.4.

3.3.2.2. Rail supply model The topological model for the rail mode takes into account all relevant railways lines within the study area, leading to a total of 406.975 km, modelled through a graph encompassing 90.259 links and 83.462 nodes. Each rail link is associated with physical characteristics in accordance with the following classification: electrification (yes/no), number of tracks, gauge (six different clusters), allowance for freight transport (yes/no). A similar classification is introduced for the 5.449 freight stations/terminals in the study area, of which 321 have direct connection to ports and 257 to the inland waterway network. The backbone of the main international rail terminals is represented by 200 main terminals.

In terms of analytical model, a commercial running speed[12] was associated to each link accordingly with its type and with the country classification. Since rail mode is discontinuous in time and space, a proper modelling of each available 'service' would

12. In more detail, the actual commercial running speed for freight trains was collected for some countries, whilst for the remaining average values have been used, depending also on the number of tracks and the electrification type.

be desirable for the realistic implementation of an analytic model. In that respect, since there is a substantial lack of data for rail freight transport, a simplified approach is adopted: all the 200 main international terminals are assumed to be connected with each other through direct services, whilst all the remaining terminals (local terminals) are connected with the closest international terminal, in order to mimic a hub-and-spoke network structure. From this assumption, travel times have been calculated as the sum of running times (t_r) and waiting times (t_w): the former is determined on the basis of the mentioned link speed, the latter has been assumed equal to 6 hours for connections between main international terminals and to 48 hours for connections implying the presence of one or more local terminals. The described procedure is applied separately for two different rail freight 'services': combined and traditional (i.e. not unitised). Finally, travel times are turned into prices, through regressions estimated on disaggregated data and/or available in the literature.

Notably, the adopted price functions are not proportional to the distance; therefore, but they can still be classified as sub-additive: as a result, the same approach described in detail for road transport (Section 3.3.2.1) can be adapted straightforwardly for the calculation of terminal-to-terminal rail shortest paths.

3.3.2.3. Maritime supply model The implementation of the topological model for the maritime mode is performed by considering firstly all the 491 ports within the study area with active freight services, and then building a graph of all possible routes connecting those ports: this objective was achieved through advanced automatic GIS-based procedures.

The development of the analytical model for maritime transport faces the same issues of rail transport, that is a proper simulation of the discontinuity in space and in time for access and egress. Differently from the railway transport, however, detailed and reliable databases reporting actual maritime services and their timetables are available from various sources.[13] In more detail, three main different sources can be taken into account: the European ShortSea Network (ESN), the AXS-Alphaliner and the Containerisation international databases. As a result, a database of maritime services for the study area was built, covering 9.933 services with the following information available: sequence of called ports, shipping line(s) and liner agent(s), vessel capacity (for container services), transit time, monthly frequency. In terms of impedances, net travel times for each type of maritime service are calculated by considering average speeds; time stops at ports for loading/unloading operations and for other duties are also taken into account.

Finally, travel times are turned into prices, through regressions estimated on disaggregated data and/or available in the literature. The adopted price functions

13. This is particularly true for liner services, whilst for tramp services a simplified approach similar to that described for rail transport should be necessarily applied.

are sub-additive as a result, the same approach described in detail for road transport (Section 3.3.2.1) can be adapted straightforwardly to calculate travel times and costs for a given type of maritime service between each pair of ports, providing information about the sequence of services used and the consistent sequence of ports called.

3.3.2.4. Inland waterways supply model The topological model for the inland waterway network is made up by 1041 links and 966 nodes, of which 364 are inland ports and/or quays with connection with road and/or rail modes. The physical characteristics taken into account for their impact on the functionality of the inland waterways are the nature of the link (natural/artificial) and its classification in accordance with the EU regulations. Unfortunately, information about water direction is not available in the current version of the DSS. Therefore, a simplified analytical model is implemented, wherein an average travel time is calculated for each link assuming a commercial speed of 5 knots independently of the direction, and an average cost is assumed by considering an average fare of 1€/km for an intermodal transport unit.

3.3.2.5. Integration of modal networks into a multimodal model The single-mode supply models described in the previous sections can be effectively integrated in order to calculate performances and impacts related to intermodal/multimodal transport services. The integration procedure is based on the presence, within each graph, of mode transfer nodes: for instance, if a given port is connected with rail and road networks, both rail and road topological models will have a specific intermodal node tagged with the *id* of that port. Therefore, starting from this premise, specific shortest path procedures, implemented with ad hoc programming codes, can be defined in order to calculate the times and costs for any combination (i.e. sequence) of modes.

In that respect, it should be noted firstly that the sub-additivity property (Section 3.3.2.1) does not hold anymore for paths on the multimodal network. In order to overcome this issue, a 'virtual' multimodal network should be implemented, wherein each mode is represented by a single 'virtual' monomodal network with macro-links representing overall o-d trips. This allows converting the sub-additive impedances on the real monomodal network into additive impedances on the corresponding macro-links of the virtual monomodal network.

Furthermore, a very frequent application of the multimodal supply model is to calculate impedances to be used as input attributes for a mode choice model: this requires estimation of the overall impedances for each given predetermined sequence of modes (e.g. a combination road-sea-road or road-sea-rail-road) within the mode choice set. Such problem can be solved by introducing a proper 'duplication' of the terminal nodes allowing modal change, so that the forward star of a given terminal related to a mode s in the predetermined sequence is made up just by macro-links related to the mode $s + 1$ in the sequence.

By way of an example, the shortest path for an intermodal road-sea-road service connecting a given o-d pair can be calculated through the following steps:

- Calculation of the shortest paths from the origin o to each port p_o and from each port p_d to the destination d by road, using the road supply model under the relevant assumptions (e.g. heavy vehicles, accompanied transport, one driver, hiring);
- Calculation of the shortest path between each pair of ports p_o-p_d by sea, using the maritime supply model under the relevant assumption (e.g. only Ro-Ro services, fares for accompanied transport);
- Implementation of a 'virtual' intermodal graph made up by all possible combinations of type o-[*road*]-p_o-[*sea*]-p_d-[*road*]-d, and calculation of the shortest path on such network.

Similar procedures can be effectively adopted for any intermodal combination involving also rail and inland waterway networks.

3.4. Applications

The DSS encompassing the MRIO model with elastic trade coefficient described in Section 3.2 was applied, by way of an example, on two different hypothetical future scenarios, involving likely changes in the generalised transport costs depending on specific market conditions and/or new governance policies:

- Scenario 1: 10% increase in road fuel costs, for example coming either from an increase of oil price or from a reduction of public subsidies for lorries;
- Scenario 2: 20% reduction of maritime fares, so as to mimic a policy of public subsidies towards motorways of the sea within the European Union.

For each scenario, the changes in generalised transport costs were implemented into the supply model described in Section 3.3, whose outputs entered in turn the whole modelling chain reported in Figure 3.1. A first result is reported in Table 3.2, which illustrates the absolute values of attracted and generated tonnes/year by country in the base scenario and the corresponding percentage variations for both the project scenarios.

Intuitively, the Scenario 1 (which penalises road transport) implies a general reduction of the total generated/attracted tonnes by country (−1.16% on the overall total trade), whilst the Scenario 2 (which supports sea transport) leads to a general increase of traded tonnes (+2.30% overall). However, looking at the detailed results of Table 3.2, a substantial heterogeneity emerges, with some apparently counter-intuitive results, for example slight increases of generated/attracted tonnes in some countries for the Scenario 1, and vice versa for the Scenario 2. This is consistent with the common sense evidence that transport and macroeconomic impacts generated by

Table 3.2: Results of DSS application for the project scenarios: tonnes/year for the base scenario and percentage variations in the project scenarios (disaggregation by attracted/generated tonnes)

Country	Million tonnes/year		Percentage variation			
	Base		Scenario 1		Scenario 2	
	Attracted	Generated	Attracted	Generated	Attracted	Generated
Austria	56.5	38.2	–2.8%	–2.0%	–0.3%	–0.5%
Belgium	174.3	163.5	–2.3%	–1.9%	0.1%	–0.9%
Czech Republic	31.6	42.4	–1.2%	–4.1%	–0.5%	–0.7%
Denmark	26.2	32.7	–2.0%	–0.7%	–0.1%	–2.2%
Estonia	3.3	9.0	–1.8%	–0.5%	2.0%	16.1%
Finland	15.2	17.9	–0.3%	0.2%	7.9%	12.4%
France	184.8	151.5	–0.5%	–1.1%	1.6%	1.4%
Germany	305.6	261.4	0.2%	–3.3%	1.3%	1.2%
Greece	11.8	7.4	–0.3%	–0.1%	5.4%	10.6%
Hungary	16.7	17.8	–2.5%	–0.5%	–0.3%	–1.0%
Irish Republic	29.3	12.1	–0.7%	–0.1%	1.8%	3.6%
Italy	94.5	77.8	–0.6%	–0.4%	2.0%	5.1%
Lithuania	3.2	10.8	–1.0%	–1.1%	5.0%	14.6%
Netherlands	177.7	258.7	–3.8%	0.4%	1.3%	–0.1%
Poland	38.5	55.5	–0.8%	–1.0%	0.8%	–0.3%
Portugal	26.2	17.6	–0.3%	–0.3%	16.3%	4.3%
Slovakia	15.1	21.8	–5.4%	–1.1%	–0.7%	–0.8%
Slovenia	12.1	7.6	–1.3%	–0.4%	0.9%	1.7%
Spain	86.4	62.9	–0.5%	–0.4%	9.7%	7.4%
Sweden	42.6	56.4	0.3%	0.6%	5.5%	7.9%
United Kingdom	98.1	126.7	–0.2%	–0.3%	4.6%	8.8%
Total	**1449.8**		**–1.16%**		**2.30%**	

change in transport costs are very difficult to predict – with remarkable difficulties in drawing prior expectations – since a number of factors may play different and contrasting roles. In particular, countries with some economic competitive advantages tend to benefit from a decrease of transport costs (and vice versa), since in such a way they are able to increase their exports: in other words, given the total final demand in the study area, an export increase by some countries will necessarily imply a reduction in the exports of some other countries. Moreover, such competitive advantage may arise as a consequence of substantially different factors (availability of raw materials, of developed technologies, of low labour cost etc.) which are implicitly accounted for in the technical coefficients. Finally, a variation of transport costs – for example a tout court decrease – impacts differently on the various countries depending on their geographical position within the study area, that is of their initial accessibility with respect to the other countries. For all these reasons, as it is known, an accessibility increase does not imply necessarily an economic improvement for every zone in the study area. In that respect, traditional gravity-based approaches are not able to recognise such complex interactions, being based on a straightforward inverse causal relationship between transport costs and trade flows, whilst a MRIO based approach is capable to recognise unexpected trade patterns and economic interactions. Nonetheless, it is important to underline that, to the authors' opinion, research is still needed in this field and many efforts should still be done both for consolidating the instruments described so far – especially the prediction capability of the trade coefficient model – and for improving them, by considering all the other feedbacks mentioned in Section 3.2. Clearly, once completed an improvement of MRIO models towards such economic feedbacks, a thorough and exhaustive campaign of experiments aimed at comparing/contrasting the results with other modelling approaches will be performed as well.

In order to show the output capabilities of the DSS reported in Figure 3.1, modelling results for the base and the project scenarios are reported also in terms of distribution of the overall tonnes/year traded by distance band both in absolute terms and in percentage variations (Table 3.3) and depicted in Figure 3.4. By the sake of completeness, Table 3.3 reports also the total trade by road, in order to complement the analyses reported in Table 3.4. Furthermore, a clearer representation of the impacts of the two project scenarios on the distribution of the modal split of the traded tonnes by distance band is reported in Table 3.4. Interestingly, transport-related results of the project scenarios are robust and easily interpretable. For instance, since Scenario 1 penalises road transport (i.e. the mode prevailingly

Table 3.3: Million tonnes/year traded by distance band and by mode in the base scenario (overall/road) and overall tonnes traded in the project scenarios (absolute values and percentage variations)

Distance band	**Million tonnes/year by distance band**					
	Base scenario		**Scenario 1**		**Scenario 2**	
	Total	**Road**	**Total**	**% Var tonnes**	**Total**	**% Var tonnes**
up to 499 km	389.9	386.8	381.3	*−2.20%*	387.6	*−0.60%*
500–999 km	537.9	492.4	530.9	*−1.29%*	537.5	*−0.07%*
1000–1499 km	279.2	209.9	278.7	*−0.19%*	288.7	*3.38%*
1500–1999 km	140.6	71.4	140.2	*−0.30%*	148.3	*5.43%*
2000–2499 km	65.7	34.9	65.3	*−0.62%*	76.8	*16.97%*
2500 km and more	36.3	13.1	36.4	*0.27%*	44.1	*21.52%*
Total	**1449.6**	**1208.5**	**1432.8**	***−1.16%***	**1482.9**	***2.30%***

Table 3.4: Modal split in the base scenario and in the project scenarios (absolute values and variations)

Distance band	Base scenario	Scenario 1		Scenario 2	
	Modal split (road)	Modal split (road)	Modal split var.	Modal split (road)	Modal split var.
up to 499 km	99.20%	99.15%	– 0.05%	99.04%	– 0.16%
500–999 km	91.55%	90.80%	– 0.75%	89.41%	– 2.14%
1000–1499 km	75.16%	72.70%	– 2.47%	68.36%	– 6.80%
1500–1999 km	50.77%	46.05%	– 4.71%	36.84%	– 13.92%
2000–2499 km	53.13%	47.99%	– 5.14%	26.34%	– 26.79%
2500 km and more	36.11%	30.21%	– 5.89%	13.94%	– 22.17%
total	**83.37%**	**81.63%**	**– 1.73%**	**77.07%**	**– 6.30%**

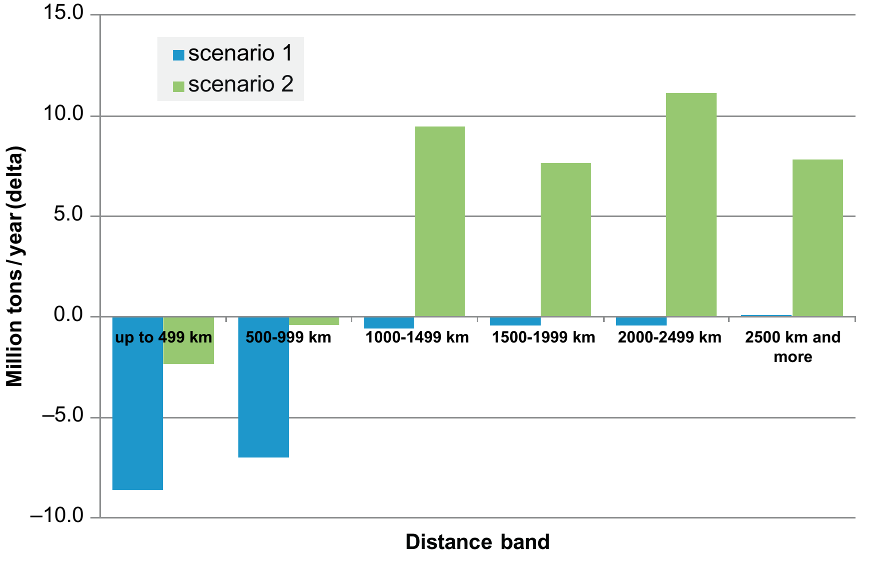

Figure 3.5: Absolute variation of traded tonnes between the project scenarios and the base scenario by distance band (values in million tonnes/year).

used for short trips) whilst Scenario 2 supports short-sea shipping (i.e. a mode mostly used for medium to long range trips), the model provides consistently a reduction of tonnes/year for shorter distances for Scenario 1 and an increase of tonnes/year for longer distances for Scenario 2 (see Table 3.4 and Figure 3.5).

3.5. Conclusions

This chapter dealt with the proposition of a multimodal elastic trade coefficients MRIO model for freight demand in Europe. After a brief review of the MRIO modelling approach, some important theoretical issues related to both demand and supply modelling were discussed. Specifically, from the demand side, an elastic trade coefficient MRIO model was presented and some relevant macroeconomic feedbacks incorporable into the model discussed. From the supply side, a discussion on the complexity of multimodal freight networks and on the corresponding modelling requirements was presented, and a multimodal supply model for the whole Euro-Mediterranean area presented. Finally, an application to hypothetical scenario analysis revealed that complex interactions between countries and sectors may be revealed by a MRIO approach.

Importantly, some research needs worth to be assessed emerged throughout the research. Firstly, the proposed multimodal supply modelling approach represents an improvement over the existing models implemented into the available DSS: therefore, it is worth quantifying both direct (i.e. calculation of skim matrices and aggregated accessibility indicators) and indirect (i.e. application of demand models based on the skims provided by the supply model) errors coming from using traditional approximated approaches in place of the more realistic proposed approach. Furthermore, the long-term macroeconomic feedbacks discussed in Section 3.2.4 should be fully addressed from both theoretical (i.e. analysis of existence, uniqueness and proposition of solution algorithms) and operational (i.e. database implementation and data requirement issues) standpoints, and then embedded into the general DSS structure described at the beginning of Section 3.3, in order to reproduce correctly such long-term impacts also from the transport side.

References

Ben-Akiva, M., & Lerman, S. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA: MIT Press.

Cascetta, E. (2001). *Transportation systems engineering: Theory and methods*. Boston, MA: Kluwer.

Cascetta, E. (2009). *Transportation systems analysis: Models and applications*. New York, NY: Springer.

Cascetta, E., Di Gangi, M., & Conigliaro, G. (1996). *A multi-regional input-output model with elastic trade coefficients for the simulation of freight transport demand in Italy*. Proceedings of the 24th PTRC, London.

Cascetta, E., & Iannò, D. (1998). *Calibrazione aggregata di un sistema di modelli di domanda merci a scala nazionale*. Proceedings of the Annual Meeting on ''Metodi e tecnologie dell'ingegneria dei trasporti'', Reggio Calabria, Italy.

Cascetta, E., Marzano, V., & Papola, A. (2008). Multi regional input-output models for freight demand simulation at a national level. In M. Ben-Akiva, H. Meersman & E. Van de Voorde

(Eds.), *Recent developments in transport modelling: Lessons for the freight sector*. Bingley, UK: Emerald.

De la Barra, T. (1989). *Integrated Land Use and Transport Modelling*. Cambridge, MA: Cambridge University Press.

Gallo, M., Simonelli, F., & Marzano, V. (2012). An empirical comparison of parametric and non-parametric trade gravity models. *Transportation Research Record – Journal of the Transportation Research Board*, *2269*, 29–41. doi: 10.3141/2269-04

Hunt, J. D. (1993). *A description of the MEPLAN framework for land use and transport interaction modeling*. Proceedings of the 73rd TRB Meeting, Washington, DC.

IASON. (2004). *Final project report*. Retrieved from www.transport-research.info/Upload/Documents/200608/20060821_162810_15461_IASON%20Final%20Report.pdf

Izard, W. (1951). Interregional and regional input-output analysis: A model of space economy. *The Review of Economics and Statistics*, *33*, 318–328.

Jin, L., Kockelman, K., & Zhao, Y. (2003). *Tracking land use, transport, and industrial production using random-utility based multizonal input-output models: Applications for Texas trade*. Proceedings of the 82nd TRB Meeting, Washington, DC.

Juri, N. R., & Kockelman, K. (2005). *Evaluation of the trans-texas corridor proposal: Application and enhancements of the RUBMRIO model*. Proceedings of the 84th TRB Meeting, Washington, DC.

Kepaptsoglou, K., Tsamboulas, D., Karlaftis, M. G., & Marzano, V. (2009). Free trade agreement effects in the Mediterranean Region: An analytic approach based on SURE gravity model. *Transportation Research Record – Journal of the Transportation Research Board*, *2097*, 88–96.

Marzano, V. (2010). Maritime traffic and induced pressure on hinterland transport networks. An integrated modelling exercise: the case of Northern Adriatic ports. (V. Marzano & L. Lonza Eds.) JRC Scientific Technical Reports. European Commission, Joint Research Centre, Institute for Environment and Sustainability, Ispra.

Marzano, V., & Papola, A. (2004). Modelling freight demand at a national level: theoretical development and application to Italian demand. Proceedings of 2004 ETC Conference, Strasbourg.

Miller, R. E., & Blair, P. D. (1985). *Input-output analysis, foundations and extensions.* Englewood Cliffs, NJ: Prentice-Hall.

Min, Y., Kockelman, K. M., Zhao, Y., & Jin, L. (2001). Models of location, trade & travel: an application of random utility based interzonal input-output models. Proceedings of the 48th Annual North American Meetings of the Regional Science Association International.

Ortuzar, J. de D., & Willumsen, L. G. (2001). *Modelling transport*. New York, NY: Wiley.

Richards, M., & Ben-Akiva, M. (1975). *A disaggregate travel demand model*. Lexington: D.C. Heath.

Timmermans, H. (2003). *The saga of integrated land use-transport modelling: how many more dreams before we wake up?* Proceedings of the 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.

Wegener, M. (2008). *SASI model description*. Working Paper 08/01. Spiekermann & Wegener Urban and Regional Research, Dortmund, Germany.

Zhao, Y., & Kockelman, K. M. (2004). The random utility based multiregional inptut-output model: Solution existence and uniqueness. *Transportation Research*, *38B*, 789–807.

Chapter 4

The Aggregate-Disaggregate-Aggregate (ADA) Freight Model System

Moshe Ben-Akiva and Gerard de Jong

Abstract

Most freight transport models that are used by national, international, and regional authorities are lacking important aspects of logistics decision-making such as the choice of shipment size and the use of consolidation and distribution centers (DCs). At the same time, these models often assume transport (mode) optimization at the aggregate zone-to-zone level, for which in reality no agents exist. The aggregate-disaggregate-aggregate (ADA) freight model system tries to overcome these limitations by modeling the generation of trade flows and assignment to networks in an aggregate way, but simulating the logistics decisions at the level of individual firm-to-firm flows. The disaggregate part of the system is the logistics model, where shipment size and transport chain choice (including the use of transshipment centers) are determined by minimizing the total logistics cost.

To estimate a disaggregate logistics model, data are required on weight, transport chains (a sequence of modes, preferably with transshipment locations), origin and destination of the shipment, as was done for shipment size and transport chain choice models on the Swedish Commodity Flow Survey. Most countries do not have commodity flow surveys that include this information, but the disaggregate logistics model can also be calibrated to more aggregate data on the modal shares by commodity and aggregate zones. Models that use the latter approach have been developed in Norway, Sweden, and Flanders and are under development in Denmark and for the European

Freight Transport Modelling

ISBN: 978-1-78190-285-1

Union (Transtools 3), and could potentially be developed for other regions where only aggregate freight transport data are available.

Keywords: Freight transport model; logistics; disaggregation; micro-simulation

4.1. Introduction

Disaggregate models – defined here as models using observations at the level of the traveler, the traveling group, the business establishment or the shipment – have several advantages over aggregate models (which use groupings of those units as observations, e.g., groupings by geographic zone). Disaggregate models can be based on a foundation in behavioral theory, can include more detailed policy-relevant variables and do not suffer from the aggregation biases of aggregate models. Nevertheless, there are perfectly valid reasons why some of the components of a model system are modeled in an aggregate fashion. In this chapter, we propose an "aggregate-disaggregate-aggregate" (ADA) model system for freight transport.

The context is that of a model system at the international, national, or regional scale, designed by or for public authorities. Such international, national, and regional freight transport models are used for different purposes, including:

- Forecasting transport demand (and through this, emissions from traffic, traffic safety, etc.) in the medium to long run under various scenarios;
- Testing transport policy measures, such as road user charging;
- Predicting the impacts on traffic (and traffic-related measures as mentioned above) of the provision of new infrastructure (roads, railway lines, canals, bridges, tunnels, ports, public freight terminals).

In the ADA model system, the production to consumption (PC) flows and the network model are specified at an aggregate level for reasons of data availability. Between these two aggregate components is a logistics model that explains the choice of shipment size and transport chain, including mode choice for each leg of the transport chain. This logistics model is a disaggregate model at the level of the firm, the decision-making unit in freight transport.

Most (inter)national or regional freight transport model systems are lacking logistics elements, such as the determination of shipment size or the use of distribution centers (DCs). Exceptions are the SMILE and SMILE + model in The Netherlands (Bovenkerk, 2005; Tavasszy, van de Vlist, Ruijgrok, & van de Rest, 1998), the SLAM model for Europe (SCENES Consortium, 2000; TNO, 2008), the EUNET model for the Pennine Region in the UK (Jin, Williams, & Shahkarami, 2005) and other regions in the UK (Bates, 2011), the model for Oregon (Hunt, 2003; Hunt et al., 2001; PbConsult, 2002), the work of Liedtke (2005), which includes an application to German long-distance markets and the FAME model for freight transport in the United States (Samimi, Mohammadian, & Kawamura, 2010). Reviews of these developments are given in Tavasszy, Ruijgrok, and Davydenko (2012) and de Jong, Vierth, Tavasszy, and Ben-Akiva (2012).

Section 4.2 of this chapter explains the structure of the ADA model system. The various components of this model system are treated in Section 4.3, focussing on the disaggregate ("middle") part of the ADA system (the logistics model) and on the disaggregation that comes immediately before the disaggregate part and the aggregation that comes directly after it. The logistics model's data requirements are discussed in Section 4.4. Estimation/calibration and validation issues for the logistics model are discussed as well (see Section 4.5). These first five sections all present the model system in general terms, not in terms of an application to a specific study area. In Section 4.6, various applications of the ADA model in the context of the freight transport model systems for Norway, Sweden Flanders, and Denmark are presented. Finally, a summary and conclusions are provided in Section 4.7.

4.2. The ADA Model Structure

4.2.1. The General Concept

Figure 4.1 is a schematic representation of the structure of the freight model system. The boxes indicate model components. The top level of Figure 4.1 displays the aggregate models. Disaggregate models are at the bottom level. So in the ADA model system, we first have an aggregate model for the determination of PC flows, then a disaggregate "logistics" model, and finally another aggregate model for network assignment. In Section 4.2.1, we explain the general concept of the ADA model system. The relation between the first aggregate part (zone-to-zone PC flows) and the disaggregate part is further treated in Section 4.2.2, and in Section 4.2.3 the boundary lines between the disaggregate logistics model and the last aggregate part (assignment) are discussed.

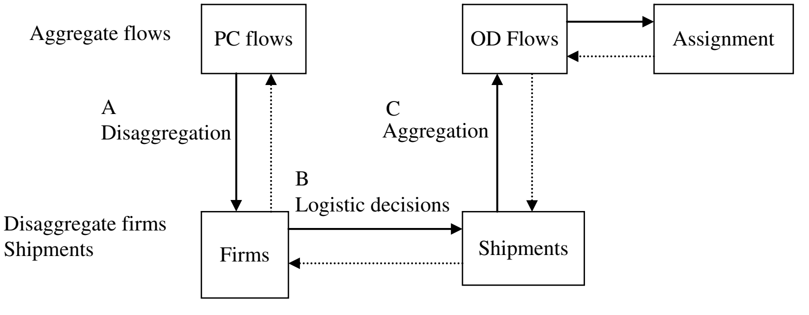

Figure 4.1: ADA Structure of the (Inter)national/regional Freight Transport Model System.

The model system starts with the determination of flows of goods between production (P) zones and consumption (C) zones (being retail for final consumption; and further processing of goods for intermediate consumption). These models are commonly based on economic statistics (production and consumption statistics, input–output tables, trade statistics) that are only available at the aggregate level (with zones and zones pairs as the observational units). Indeed, to our knowledge, no models have been developed to date that explain the generation and distribution of PC flows at a truly disaggregate level.

Most existing freight transport model systems include submodels for generating PC or origin-destination (OD) matrices (possibly by mode) and routines for assigning either one of these matrices to the networks (unimodal or multimodal). As explained in Section 4.2.2, assignment of PC flow to the networks would not be correct. A relatively new phenomenon of the ADA model is the inclusion of a logistics model that on the basis of PC flows produces OD flows for network assignment. The logistics model consists of three steps:

A. Disaggregation to allocate the flows to individual firms at the P and C end;
B. Models for the logistics decisions by the firms (e.g., shipment size, use of consolidation and DCs, modes, loading units, such as containers);
C. Aggregation of the information per shipment to OD flows for network assignment.

This model structure allows for logistics choices to be modeled at the level of the actual decision-maker, along with the inclusion of decision-maker attributes.

The allocation of flows in tons between zones (step A) to individual firms are, to some degree, based on observed proportions of firms in local production and consumption data, and from a registry of business establishments. The logistics decisions in step B (to be modeled as a random cost choice model) are derived from minimization of the full logistics costs, including the transport costs.

The aggregation of OD flows between firms to OD flows between zones provides the input to a network assignment model, where the zone-to-zone OD flows are allocated to the networks for the various modes. Assignment can, in principle, be done at the level of individual vehicles (microscopic or mesoscopic models for simulating route choice, see Ben-Akiva, Bottom, Gao, Koutsopoulos, & Wen (2007); the Oregon model also has assignment at the level of individual vehicles). In such cases, the ADA model would be an ADD (aggregate-disaggregate-disaggregate) model (see Figure 4.2).

Most model systems perform an assignment of aggregate zone-to-zone flows (possibly with several user classes) to the networks in order to use available software, to keep the model tractable and to keep the run time manageable. This approach is also used because the network level is the level at which validation/calibration data are usually available (e.g., traffic counts at various locations/screenlines). On the other hand, vehicle-level data that can be used for the estimation of micro-level network models are becoming more common.

There can also be backward linkages, as seen in Figure 4.1 (the dashed lines). The results of network assignment can be used to determine the transport costs that will

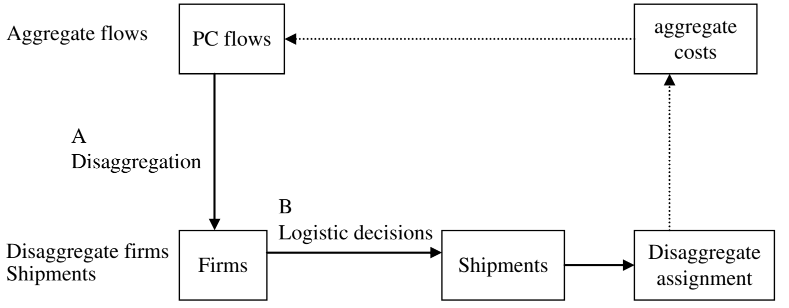

Figure 4.2: Proposed ADD Structure of the (Inter)national/regional Freight Transport Model System.

be part of the logistics costs which are minimized in the disaggregate logistics model. The logistics costs for the various OD legs can be summed over the legs in the PC flow (and aggregated to the zone-to-zone level by an averaging over the flows). These aggregate costs can then be used in the model that predicts the PC flows (for instance, as part of the elastic trade coefficients in an input–output model). In case of assignment of individual vehicles in an ADD model, one might calculate aggregate transport costs at the PC level by adding the costs of the different vehicles that are involved in the same PC chain, and then averaging over PC flows (Figure 4.2, dashed lines).

4.2.2. *Relation between the PC Flows and the Logistics Model*

The PC flows between the production locations *P* and the consumption locations *C* are given in tons by commodity type. The consumption locations here refer to both producers processing raw materials and semifinished goods and to retailers. The logistics model then serves to determine which flows are covered by direct transports and which transports will use ports, airports, consolidation centers (CCs), DCs, and/or railway terminals. It also gives the modes and vehicle types used in the transport chains. The logistics model, therefore, takes PC flows and produces OD flows. An advantage of separating out the PC and the OD flows is that the PC flows represent what matters in terms of economic relations – the transactions within and between different sectors of the economy. Changes in final demand, international and interregional trade patterns, and in the structure of the economy, have a direct impact on the PC patterns. Also, the data on economic linkages and transactions are in terms of PC flows, not in terms of flows between producers and transshipment points, or between transshipment points and consumers.

Changes in logistics processes (e.g., the number and location of depots) and in logistics costs have a direct impact on how PC flows are allocated to logistics chains, but only indirectly (through the feedback effect of logistics choices and network assignment) impact the economic (trade) patterns. Assigning PC patterns to the

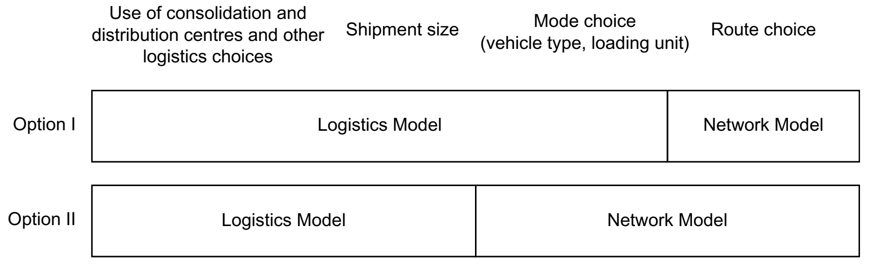

Figure 4.3: Two Options for Combining the Logistics and the Network Model.

networks would not be correct. For instance, a transport chain road-sea-road would lead to road OD legs ending and starting at ports instead of a long-haul road transport that would not involve any ports. A similar argument holds for a purely road-based chain that uses a van first to a CC, then is consolidated with other flows into a large truck, and finally uses a van again from a DC to the C destination. In this scenario, the three OD legs might be assigned to links differently than would be the case for a single PC flow. Therefore, adding a logistics module that converts the PC flows into OD flows allows for the trade-off between inventory, order and transport costs (endogenous shipment decisions) and for a more accurate assignment. The data available for transport flows (from traffic counts, roadside interviews and interviews with carriers) also are at the OD level or screenline level, not at the PC level.

4.2.3. Relation between the Logistics Model and the Network Assignment

In several existing freight transport model systems, a network model carries out both the modal split and the assignment to the networks (in a multimodal assignment). If the assignment in the ADA model system would be multimodal, the logistics model would not have to cover mode choice. This is labeled Option II in Figure 4.3. A better approach would be to include mode choice in the logistics model, and restrict the assignment to be unimodal. In the latter approach, the mode choice would be determined together with the other logistics choices (e.g., shipment size, number of legs in a transport chain, terminals used). This is Option I in Figure 4.3. The outputs of the logistics model will then be in terms of vehicle or vessel flows (not just tons) between OD pairs.

4.3. Model Specification

In this section we describe each component of the ADA model: generation of PC flows; logistics model in three steps; and assignment. Additional detail on the logistics model can be found in de Jong and Ben-Akiva (2007).

4.3.1. Generation of PC Flows at the Aggregate Level

The type of model used here can be a multiregional input/output (MRIO) model or a regionalized national input/output model (Cascetta, 2001; Hunt & Abraham, 2005; Marzano & Papola, 2004), or a spatial computable general equilibrium (SCGE) model (Bröcker, 1998; Hansen, 2010; Ivanova et al., 2007; Tavasszy, Thissen, Muskens, & Oosterhaven, 2002; Vold & Jean-Hansen, 2007). Input data required for these models are input–output statistics (preferably multiregional), production and consumption statistics by economic sector and international trade statistics. If the resulting PC flows would be in monetary values, conversion to tonnes would have to be done before going to the logistics model. Conversion factors can be based on data sets that include value and weight information for the same shipments (such as Commodity Flow Surveys) or trade/transport flows (e.g., customs data).

4.3.2. The Logistics Model

In Section 4.3.2 we discuss the three steps (A, B and C; see Figure 4.1) of the logistics model.

4.3.2.1. Disaggregation to Firm-to-Firm Flows (Step A) Step A in the logistics model (conversion of zone-to-zone flows to firm-to-firm flows) is not a choice model, but rather a prerequisite so that logistics choices can be captured at the actor level. Instead of modeling trade between zones, this step makes modeling trade between firms possible. (These firms are manufacturers, wholesalers or retailers.)

The aggregate representation of the PC flows that are produced by the first aggregate model of the ADA system, and that are input to the disaggregation step A, is as follows:

Flows of goods in tons per year, by

- r, zone of the sender (production zone)
- s, zone of the receiver (consumption zone)
- k, commodity type.

Step A disaggregates this to a disaggregate representation, characterised by:

Flows of goods in tons per year, by

- m, sending firm (located in zone r),
- n, receiving firm (located in zone s),
- k, commodity type.

Three general approaches to generate a disaggregate population or sample of firm-to-firm flows can be distinguished:

1. Re-weighting-use an existing sample or population and re-weight using marginal distributions (i.e., the row and column totals);
2. Synthetic-draw from a sequence of conditional distributions;

3. Hybrid-begin with re-weighting and enrich the set of characteristics using synthetic draws.

The reweighting approach is the simplest, but a sample of actual firm-to-firm flows is only rarely available. The United States and Sweden have a Commodity Flow Survey (CFS) sample, but these are samples of shipments for a limited time period (e.g., one to three weeks in Sweden). In this case, we are looking for all supplier/receiver pairs on an annual basis. Therefore, in practically all applications, a synthetic or a hybrid approach for step A is developed. In Annex 1 we give a practical example (from the implementation in Norway) of how the disaggregation step can be worked out.

4.3.2.2. The Logistics Decisions at the Disaggregate Level (Step B) Step A produces disaggregate supplier–receiver relations (a business relation between two firms in which one is the sender of a good and the other the receiver). Each relation has an annual flow of goods in tons by commodity type. Even for a small area, there are millions of such relations. To reduce runtime, firm-to-firm relations are sampled, and expansion factors are used to obtain population estimates. For each relation, step B simulates the logistics decisions (micro-simulation), and adds the outcomes of these to the level of a firm-to-firm relation.

The different logistics decisions included in step B are:

- Frequency/shipment size (so inventory decisions are endogenous);
- Choice of loading unit (e.g., containerized or not);
- Use of DCs, freight terminals, ports and airports and the related consolidation and distribution of shipments. The locations of these transshipment points are taken as given, what is determined here is their use. This also gives the number of legs in the transport chain;
- Mode/vehicle type used for each leg of the transport chain. The choice set may contain: air transport, road transport, rail transport and maritime transport (possibly each with different vehicle/vessel types). Economies of scale in transport (larger vehicles have lower unit costs) are taken into account in the cost functions for the vehicle types.

What step B of the logistics model does is to add dimensions to the disaggregate representation that was produced by step A. The full disaggregate representation after step B, consists of:

Shipments of goods in number of shipments, tons, ton-kilometers, vehicle-kilometers and vehicle/vessels per year, by

- m, sending firm (located in zone r),
- n, receiving firm (located in zone s),
- k, commodity type,
- q, shipment size,
- l, transport chain type (number of legs, mode and vehicle/vessel type used for each leg, terminals used, loading unit used).

The basic mechanism in the model for decision-making on all these choices is the minimization of total logistics costs. The total annual logistics costs *G* of commodity *k* transported between firm *m* in production zone *r* and firm *n* in consumption zone *s* of shipment size *q* with transport chain *l* (including number of legs, modes, vehicle types, loading units, transshipment locations) are:

$$G_{rskmnql} = O_{kq} + T_{rskql} + D_k + Y_{rskl} + I_{kq} + K_{kq} + Z_{rskq} \tag{4.1}$$

where:
G: total annual logistics costs
O: order costs
T: transport, consolidation and distribution costs
D: cost of deterioration and damage during transit
Y: capital costs of goods during transit
I: inventory costs (storage costs)
K: capital costs of inventory
Z: stockout costs

In this minimization, it is assumed that the subscripts for the specific firms *m* and *n* (and also, for instance, firm size) do not matter. This assumption may be relaxed to accommodate economies of scale in for instance warehousing and ordering. Also, variation in the discount rate for the inventory capital costs and of other preferences between firms may be included.

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

The decision-making thus takes place at the level of the individual sender to receiver relation. Whether these decisions are taken by the sender or by the receiver will vary from sector to sector and even within sectors. One way to look at it is to regard the sender–receiver combination (and also the carrier if the transport is contracted out) as a single decision-making unit, carrying out the minimization of the total logistics costs of this firm-to-firm flow. The idea of joint overall optimization by shippers and carriers was supported by experimental economics (Holguín-Veras, Xu, de Jong, & Maurer, 2011).

Eq. (4.1) is expanded as follows (see RAND Europe, Solving and INRO, 2004; RAND Europe and SITMA, 2005):

$$\begin{aligned} G_{rskmnql} = {} & o_k \cdot \frac{Q_k}{q_k} + T_{rskql} + i \cdot j \cdot g \cdot v_k \cdot Q_k + \frac{i \cdot t_{rsl} \cdot v_k \cdot Q_k}{365} \\ & + (w_k + (i \cdot v_k)) \cdot \frac{q_k}{2} + a \cdot ((LT \cdot \sigma^2_{Qk}) + (Q^2_k \cdot \sigma^2_{LT}))^{1/2} \end{aligned} \tag{4.2}$$

where:
o: the constant unit cost per order
Q: the annual demand (tons per year)

q: the average shipment size
i: the discount rate (per year)
j: the fraction of the shipment that is lost or damaged (might vary between modes)
g: the average period to collect a claim (in years)
v: the value of the goods that are transported (per ton)
t: the average transport time (in days)
w: the storage costs per unit per year
a: a constant to set the safety stock in such a way that there is a fixed probability of not running out of stock. For medium/high frequency products, a common assumption is that the demand (and lead-times) follows a Normal distribution. *a* will then be: $a = F^{-1}(CSL)$, where F^{-1} is the inverse Standard Normal Distribution and *CSL* is the cycle service level, which is the probability that the stock will not be empty during a replenishment cycle.
LT: expected lead-time for a replenishment (time between placing the order and replenishment)
σ_{LT}: standard deviation for the lead-time
σ_{Qk}: the standard deviation for the yearly demand

The first term on the right-hand size (RHS) of Eq. (4.2) gives the order cost, the second the transport cost, the third the cost of damage to the goods during transport, the fourth the capital costs on the goods in transit, the fifth the capital and storage costs of the (average) inventory and the last term represents the safety stock cost.

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

4.3.2.3. Aggregation to Zone-to Zone Flows (Step C) In step C of the logistics model, transport chain legs of individual firm-to-firm shipments for the same commodity type are aggregated by origin zone (on the basis of all sending firms and transshipment locations in the zone for the commodity) and destination zone (on the basis of all receiving firms and transshipment locations in the zone for the commodity). In this way we obtain OD flows in vehicles and tons.

The aggregate representation that is produced in step C is as follows:
Flows of vehicle/vessel units per year, by

- *r*, origin zone (production zone or zone used for transshipment),
- *s*, destination zone (zone used for transshipment or consumption zone),
- *k*, commodity type (and distinguishing empty vehicles),
- *v*, mode (e.g., road transport) or vehicle/vessel type (e.g., heavy truck).

If the disaggregation was done properly, and if step B has not introduced errors, this is simply a matter of straightforward summation over shipments.

In reality not every shipment in a firm-to-firm flow is optimized and transported by itself, but multiple shipments from different firm-to-firm flows are often combined into a single vehicle (which we call ''consolidation''). This reduces the cost burden on the individual firm-to-firm flow, because now costs can be shared with other sender-receiver pairs. In all the ADA models developed so far, we allow for consolidation between transshipment locations: these terminals not only serve for modal transfer, but also for consolidating shipments (and deconsolidating these before final delivery). The degree of consolidation depends to a large extent on the presence of other cargo that could be moved between the same terminals, and this can be determined by an initial model run. The next model run then takes the predicted terminal to terminal flows from the first run and uses these to determine the share that an individual shipment has to pay of the transport cost of consolidated shipments, and so on, applying the logistics model in an iterative fashion.

Empty vehicle flows are calculated as follows: the loaded trips are first calculated as described above, and then vehicle balances between zones are used to let vehicles return from where they came, with specific shares for empty and loaded return trips. In this formulation, the probability that some of the empty capacity will also be used for transporting goods in the opposite direction is taken into account.

4.3.3. Assignment to Networks

Standard aggregate network assignment software can be used to carry out an assignment of vehicles or tons (ADA model), or a (less standard) disaggregate assignment (ADD model). The latter is usually a simulation-based procedure, where individual vehicles are loaded one-by-one onto the network. This simulation can be microscopic, and includes all the movements of each vehicle on the network in detail (e.g., including lane changes) or mesoscopic, where more aggregate relationships (such as speed-density curves) are used to model individual vehicle movements (see Ben-Akiva et al., 2007).

4.4. Data Requirements

For the logistics model presented, the following data are needed for step A (disaggregation from flows to firms):

1. The number of firms (or the local units for firms with multiple establishments) by commodity type and zone.
2. The turnover of these local units and/or the number of employees of these firms.

This information is required both at the production and the consumption end. Another requirement is the consumption pattern (in terms of the commodity classification used) of the firms by commodity type produced. We assume that each firm (local unit) will produce goods in only one commodity class, but it may consume goods from several commodity groups.

Step B (logistic decisions) requires information on the following items:

1. Data on individual shipments: sector of sender and receiver, origin and destination, value of the goods, modes and vehicle/vessel type and size used, cargo unit, shipment size/frequency, use of freight terminals (including intermodal terminals and marshaling yards), consolidation and DCs, ports and airports. Preferably this is transport chain information: which shipments go directly from P to C, which use the above intermediate points?
2. Data on where the freight terminals, consolidation and DCs, ports and airports are located;
3. Data on logistics costs: transport costs per km, terminal costs, handling and storage costs for all available alternatives.

Most crucial are the data on the shipments of individual firms (item 1 for step B above). The spatial detail needs to be that of the zones used in the model. Step C requires no extra information.

4.5. Estimation, Calibration and Validation

We distinguish between model estimation (which takes place on disaggregate data using formal statistical methods), model calibration (which takes place on aggregate data and may or may not involve formal statistical methods) and validation (which takes place after having done the assignment and involves a comparison with traffic count data). Figure 4.4 represents an estimation, calibration and validation process for the ADA model system for freight transport. The estimation and calibration data are shown above the boxes, while the validation data are below the boxes. The matrices of PC flows are usually partly based on observations and partly synthetic (model-generated). The data used in this process comes from a CFS, regional input–output systems, economic statistics from national accounts and foreign trade data. The logistics model is estimated on a CFS or similar disaggregate data, coupled with information on terminals and time and costs data from the networks.

The model application process is iterative (see Figure 4.4, middle and upper part): after assignment, the new generalized costs are used to adjust the PC matrices, etc. This gives rise to an *inner loop*, which functions as follows:

1. The PC models (e.g., MRIO models) provide initial PC matrices;
2. The logistics model transforms these into OD matrices, using transport cost provided by the network model, and adds empty vehicles;
3. The network model assigns the OD matrices (including empty vehicles) to the networks;
4. The network model and the logistics model provide transport and logistics costs matrices to the PC model;

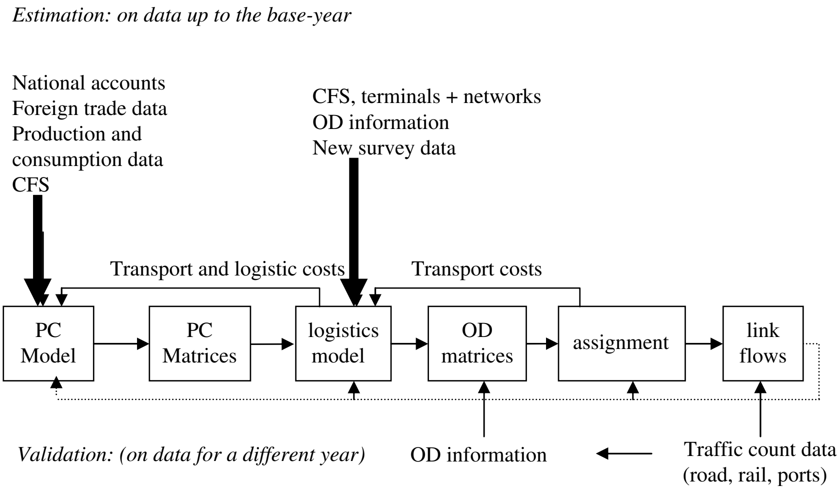

Figure 4.4: Estimation, Calibration and Validation of the Model Systems.

5. The PC model produces new base matrices on the basis of the new transport and logistics costs and provides these to the logistics model.

This loop continues until equilibrium is reached (in practice, until a preset maximum distance from equilibrium is reached). Estimation is not required within this inner loop. The inner loop addresses the adjustment of model variables (inputs and outputs), not model coefficients.

4.5.1. Estimation with Disagregate Data

Data on logistics choices of individual shipments are used in model estimation for step B. The model is based on the total annual logistics costs, such as Eq. (4.2). A random cost discrete choice model can be obtained by using total annual logistics costs as the observed component and by adding random cost components ε that follow specific statistical distributions. These random components account for omitted variables, measurement errors and such.

$$C_{mnql} = G_{mnql} + \varepsilon_{mnql} \tag{4.3}$$

where (dropping the subscripts *r* and *s* for the zones and *k* for the modes):
C_{mnql}: total logistic and transport cost
G_{mnql}: observed component of total transport and logistics costs
ε_{mnql}: random cost component.

Using Eq. (4.2) for G_{mnql} we get:

$$C_{mnql} = \beta_{0ql} + \beta_1 \cdot \frac{Q}{q} + X_{mnql} + J_{mnql} + \beta_2 \cdot j \cdot v \cdot Q + \beta_3 \cdot \frac{t_{mnl} \cdot v \cdot Q}{365} + (\beta_4 + \beta_5 \cdot v) \cdot \frac{q}{2} + a \cdot ((LT \cdot \sigma_Q^2) + (Q^2 \cdot \sigma_{LT}^2))^{1/2} + \varepsilon_{mnql} \quad (4.4)$$

where:
β_{0ql} = alternative-specific constant
β_1 = o
β_2 = d.g
β_3 = d (in transit)
β_4 = w
β_5 = d (warehousing)

In Eq. (4.4) we included a number of items, such as order costs, inventory costs and capital costs of goods during transit, in the coefficients to be estimated because the data on these items can be very difficult to obtain. As a result, the coefficients have specific logistical interpretations. We distinguish between the implied discount rate (d) of the inventory in transit (β_3) and of the inventory in the warehouse (β_5), because these need not be the same.

If we assume the extreme value distribution for ε, the model becomes logit. Nested logit is appropriate when some alternatives (e.g., road and rail transport) have a greater degree of substitution than other alternatives (e.g., road and sea transport). This is an empirical question, and statistical tests, particularly likelihood ratio tests, show whether or not the Nested Logit model is justified. A Logit Mixture Model provides additional flexibility and may be relevant for the logistics model, given the heterogeneity often found in freight transport. Heterogeneity can be captured in two ways:

- Two error components in ε:
  - one following the extreme value distribution;
  - and the other following, for instance, a multivariate normal distribution to allow for flexible correlation structures between alternatives.
- The coefficients in G (the β's) follow a distribution. This is the random coefficients model or tastes variation model. It may capture heterogeneous preferences in freight transport decision-making.

4.5.2. Calibration to Aggregate Data

In the absence of disaggregate estimation data, it is possible to use a deterministic logistics model. This model can be more easily calibrated with aggregate data (OD information).

Through this ADA model, we have developed a procedure to calibrate parameters in the cost function to available aggregate data, which was tested in Norway. For the purpose of calibrating the model to aggregate data, a number of calibration parameters (e.g., for implied discount rates such as β_3 and β_5) in Eq. (4.4), mode-specific constants, a constant for direct transport) were added to the cost function Eq. (4.2). Observed OD data by mode and commodity type for aggregate zones (e.g., 10×10 zones for a country) are used as calibration targets. The calibration parameter values that minimize the difference between the model outputs and the calibration targets were then determined in an iterative process.

4.5.3. *Validation to Traffic Count Data*

The validation process is depicted in Figure 4.4 (bottom part of the figure). After the assignment of the OD flows to the networks, the predicted link flows are compared to observed link flows from traffic counts (especially for road and possibly for rail and port throughput). Any large discrepancies that may arise require analysis. The parameters in all the models are then recalibrated, employing an iterative procedure. This process creates the *outer loop*. In the outer loop, or model calibration loop, model coefficients in all constituent submodels are adjusted to reach a good match with aggregate data.

4.6. Applications to Norway, Sweden, Flanders and Denmark

The ADA model was first specified in a project for Norway and Sweden (RAND Europe, Solving and INRO, 2004) to replace the existing multimodal network models for freight transport. Both countries have model components for deriving PC flows and for network assignment. Prototype versions (version 0) of the logistics model were developed and tested (RAND Europe, 2006; RAND Europe and Sitma, 2005, 2006). There are two different models-one that is part of the Norwegian national freight model system, and one for the Swedish national freight transport forecasting system – but both have the same structure. The Norwegian and Swedish models differ in terms of the zoning system, the modes and vehicle types used, the commodity classification, the construction of the PC matrices, the cost functions and the consolidation rules.

In 2006/2007, version 1 models for Norway and Sweden were constructed. The logistics cost minimization in this improved deterministic model takes place in two steps. In the first step (transport chain generation), the optimal transshipment locations are determined for each type of transport chain (e.g., road-sea-road) between each origin and destination zone. The transport chain generation program determines the optimal transfer locations on the basis of the set of all possible multimodal transshipment locations (which is exogenous input). These terminals are coded as separate nodes and the program uses unimodal network information on times and distances between all the centroids and all the nodes for all available submodes to find the paths, which minimize the total logistics costs.

In the second step (transport chain choice), shipment size and transport chain (number of legs, selection of modes and vehicle types) are determined by enumerating all available options for a specific firm-to-firm flow and selecting the one with the lowest logistics costs.

For Norway, this model uses all firm-to-firm flows based on data on firms by number of employees and municipality. No expansion is needed to determine the population of all goods flows in Norway. There is also a SCGE model (PINGO), that can be used to produce forecasts of PC matrices (Hansen, 2010).

In Sweden, the logistics model is combined with procedures for estimating PC matrices, which are based on the National Commodity Flow Survey and mainly use gravity type models (base matrices). A sample of firm-to-firm flows (for different size classes) is used for application of the disaggregate logistics choices, after which, an expansion procedure needs to be used to arrive at population totals.

For both countries, the logistics model produces OD matrices (e.g., in terms of vehicles and tons) that are assigned to the road, rail, sea and air networks. The Swedish model allows for consolidation of shipments within the same commodity type; in the Norwegian model, consolidation is possible within larger segments of the goods market (combinations of commodity types).

The Norwegian model has been used in many applications (including a national transport plan with long-run forecasts and various sensitivity analyses, corridor analysis, various cost-benefit analyses, the Norwegian Climate project; see Kleven, 2011). An example of long-run forecasts from this model is in Figure 4.5.

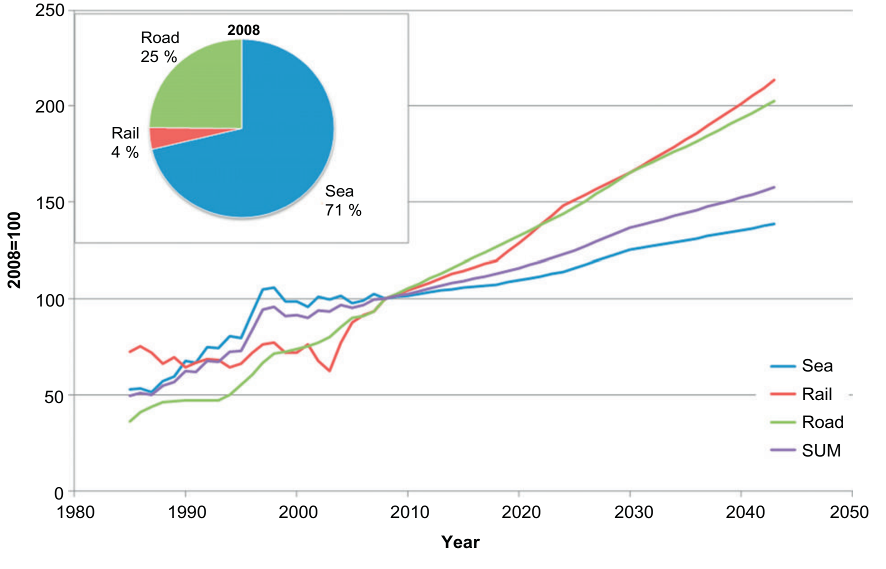

Figure 4.5: Realizations (1980–2008) and Forecasts (2008–2050) by Mode (2008 = 100) in the Norwegian National Transport Plan (Kleven, 2011).

The Swedish model has so far been used in a limited number of practical studies (changing the requirements for maritime fuels, impact of rail user charges, co-modality project, Sweden-Ruhr Area corridor study, see Vierth, 2011).

The freight transport model for the Mobility Masterplan of Flanders (the northern half of Belgium) also follows the ADA. The PC matrices come from an existing trade model, called Planet. Then comes a newly developed logistics model for the choice of shipment size and transport chain. This also is a deterministic model minimizing total logistics costs, that was calibrated using aggregate data on the mode shares for domestic transport, import and export of Flanders. Assignment for road transport is done jointly with the assignment of cars (de Jong et al., 2010). The ADA model for Flanders was used for several different scenarios for 2010–2040 and many sensitivity analyses. An interesting example of the outcomes of this model and the potential benefit of logistics modeling (with transport chain and shipment size optimization instead of just modal split) are the road transport costs elasticities. For the effect on tons by road, these elasticities are around 0, but for vehicle kilometers by road they are around –1. This difference in elasticities is mainly due to:

- Shifts form direct road transport to transport chains like road-rail-road, which count twice for road tonnage, but which have much shorter road distances than direct road transport;
- Increases in shipment size.

A model system similar to that in Norway and Sweden is now being built for Denmark (Hansen, 2011). A key change with regards to the Norwegian and Swedish models is that the Danish model will use a pivot-point procedure on the truck matrix (and similarly for sea and rail transport). This means that the logistics model will only deliver *changes* in the OD matrices between a base and a future year. These changes will then be applied to a base year truck OD matrix, which will be based to the maximum possible degree on observed data (e.g., traffic counts, surveys), making the overall model less synthetic and better empirically founded. The Danish model will make use of disaggregate data (partly from Stated Preference interviews in the Fehmarn Belt corridor).

The new Transtools 3 model (for the European Commission; see the chapter in this book by Nielsen et al.), also plans to use an approach for freight transport and logistics based on the experiences in Norway and Sweden (and Flanders, Denmark), in combination with base year OD matrices by mode, derived independently.

4.7. Conclusion

We presented an ADA model system for international, national or regional freight transport, including: aggregate models for the determination of goods flows between production and consumption zones (such as input–output models); disaggregate models for logistics decisions, including shipment size, number of legs in the transport chain, use of consolidation and DCs and mode; vehicle type and loading unit for each leg; and assignment of the aggregate OD vehicle flows to the networks. The ADA model is specified at the disaggregate level (individual shipments) where disaggregate modeling is feasible and most attractive, but at the aggregate level where disaggregate models are not possible or not attractive because of constraints on available data, software and runtime and for ease of interpretation.

When the middle part of a model system is disaggregate, and the parts that come before and after it are aggregate, additional disaggregation and aggregation steps become necessary. There are many ways to do a disaggregation and the problem has many feasible disaggregate outcomes. Sometimes data are available to determine which solution is most likely, but more often assumptions need to be made as the basis for disaggregation. The aggregation process however will have only one solution (though there can be different dimensions to represent the aggregate outcomes, e.g., in tons or in vehicles). It is important to do the disaggregation in such a way as to preserve the original PC zone-to-zone totals. Estimation of the disaggregate logistics model described in this chapter requires data on individual shipments, which in most countries are not available to researchers. However, a calibration of such a micro-level model to data at a more aggregate level is also feasible.

In the longer run, an aggregate-disaggregate-disaggregate (ADD model), with assignment of each individual vehicle to the networks, may be possible, though it would likely incorporate a mesoscopic, not a microscopic, assignment. Modeling the international and interregional trade patterns at a disaggregate level (DDD model) may also be possible in the longer run (through choice of supplier by receivers, or the other way around). However, given the nature of the data that are available for this step (trade data from custom records, production and consumption by sector; at best, multiregional input/output data), aggregate models for PC flows can be expected to remain mainstream.

Acknowledgments

This chapter is, to a large extent, based on a series of projects carried out for the Work Group for Transport Analysis in the Norwegian national transport plan, the Samgods group in Sweden, the Government of the Flanders region in Belgium and the Technical University of Denmark (DTU). Organizations that were involved in these projects include Significance, RAND Europe, SITMA, IMOB University Hasselt, DTU, Solving International, Solving Bohlin & Strömberg, and INRO Canada. The authors would like to thank Inge Vierth of VTI/CTS for comments on an earlier version.

References

Bates, J. J. (2011). Freight modelling and policy analysis in The Unites Kingdom. Presentation at the *CTS-Seminar on European and National Freight Demand Models*, Stockholm.

Ben-Akiva, M. E., Bottom, J., Gao, S., Koutsopoulos, H., & Wen, Y. (2007). Towards disaggregate dynamic travel forecasting models. *Tsinghua Science and Technology Journal*, *12*(2), 115–130.

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

Bröcker, J. (1998). Operational spatial computable general equilibrium modeling. *Annals of Regional Science*, *32*, 367–387.

Cascetta, E. (2001). *Transportation systems engineering: theory and methods*. Boston, MA: Kluwer Academic Publishers.

Hansen, C. O. (2011). Freight modelling and policy analysis in Denmark. Presentation at the *CTS-Seminar on European and National Freight Demand Models*, Stockholm.

Hansen, W. (2010). Developing a new spatial computable general equilibrium for Norway. Paper presented at European Transport Conference 2010, Glasgow.

Holguín-Veras, J., Xu, N., de Jong, G. C., & Maurer, H. (2011). An experimental economics investigation of shipper-carrier interactions on the choice of mode and shipment size in freight transport. *Networks and Spatial Economics*, *11*, 509–532.

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

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

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

Ivanova, O., Heyndrickx, C., Spitaels, K., Tavasszy, L., Manshanden, W., Snelder, M., & Koops, O. (2007). *RAEM version 3.0*. Leuven: Transport Mobility Leuven.

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

De Jong, G. C., & Ben-Akiva, M. E. (2007). A micro-simulation model of shipment size and transport chain choice in Norway and Sweden. Paper submitted for publication in *Transportation Research B*.

De Jong, G. C., Baak, J., Ruijs, K., Pieters, M., Bellemans, T., Janssens, D., & Wets, G. (2010). Passenger and freight transport in Flanders 2010–2040 under three scenario's. Paper presented at the European Transport Conference 2010, Glasgow.

De Jong, G. C., Vierth, I., Tavasszy, L., & Ben-Akiva, M. E. (2012). Recent developments in national and international freight transport models within Europe. Paper accepted for publication in *Transportation*.

Kleven, O. (2011). Freight modelling and policy analyses in Norway. Presentation at the *CTS-Seminar on European and National Freight Demand Models*, Stockholm.

Liedtke, G. (2005). *An actor-based approach to commodity transport modeling*. Ph.D. thesis, University of Karlsruhe, Karlsruhe.

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

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

RAND Europe. (2006). *Technical report on the further development of a logistics module in the Norwegian and Swedish national freight model systems*. Report for the NTP and Samgods group, PM-2216-SIKA, RAND Europe, Leiden.

RAND Europe and SITMA. (2005). The development of a logistics module in the Norwegian and Swedish national freight model systems, Deliverable 1b: further model specification. Report for the NTP and Samgods group, PM-1968-SIKA RAND Europe, Leiden.

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

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

Samimi, A., Mohammadian, A., & Kawamura, K. (2010). "Freight Demand Microsimulation in the U.S.", *World Conference on Transport Research* (WCTR), Lisbon.

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

Tavasszy, L. A., Ruijgrok, K., & Davydenko, I. (2012). Incorparating logistics in freight transportation models: State of the art and research opportunities. *Transport Reviews: a transnational transdisciplinary journal,* doi:10.1080/01441647.2011.644640

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

Tavasszy, L. A., Thissen, M. J. P. M., Muskens, A. C., & Oosterhaven, J. (2002). Pitfalls and solutions in the application of spatial computable general equilibrium models for transport appraisal. Paper presented at the 42nd Congress of the Regional Science Association, Dortmund.

TNO. (2008). Trans-Tools, Deliverable 6: Trans-Tools Final Report. Working Paper No. 6, Delft, TNO Inro (The Netherlands).

Vierth, I. (2011). Freight modelling and policy analyses in Sweden. Presentation at the *CTS-Seminar on European and National Freight Demand Models*, Stockholm.

Vold, A., & Jean-Hansen, V. (2007). *PINGO: A model for the prediction of regional and interregional freight transport in Norway*. TOI, Oslo.

Annex 1. Practical method to disaggregate zone-to-zone (z2z) flows to firm-to-firm (f2f) flows (as implemented in Norway)

Below, the procedure is explained for a hypothetical z2z relation. This example is for a commodity type *k*, but we drop the subscript *k* for convenience.

There are 400,000 tons going from zone r to s according to the PC matrices. We want to allocate this number to f2f relations within the zone pair rs.

We know (input data for the base year) that there are 10 firms sending k from r (possibly to all s). We also know (again from the input data) that there are 20 firms receiving k in s (possibly from all r).

So within rs there could be at most $10 \times 20 = 200$ *f2f relations.*

From logistics experts we got estimates for the number of receivers per sender, by commodity (a national average number for each commodity type). Let there be 30 receivers per sender for k. Suppose that for Norway as a whole there are 1,000 receivers for k (all zones s) and that we also know the number of senders of k (here: 500) from all zones r. So in total for k there should actually be $30 \times 500 = 15{,}000$ *relations. As a by-product this gives the implied number of senders per receiver:* $15{,}000/1{,}000 = 15$.

The maximum overall number of relations for k is $1{,}000 \times 500 = 500{,}000$. *So* $15{,}000/500{,}000 = 3\%$ *of the maximum number of relations materializes.*

In equation form:

$$Fraction = \frac{Receivers\ Per\ Sender \times Total\ Senders}{Total\ Receivers \times Total\ Senders} = \frac{Receivers\ Per\ Sender}{Total\ Receivers}$$

In which:

Receivers Per Sender: average number of receivers per sender (from logistics experts).

Total Receivers: number of receivers in all the zones in Norway (from business register data).

Total Senders = number of senders in all the zones in Norway (from business register data).

Now for the zonal pair rs we assume that this 3% fraction is applicable. With 200 potential f2f relations there should be 6 actual f2f relations:

Actual number of f2f relations from zone r to zone s = fraction × *(Senders × Receivers)*

In which:

Senders: number of senders in a zone r.

Receivers: number of receivers in zone s.

We now select these 6 mn relations at random from the 200 available by using proportionality to the product of the production volume of firm m and the consumption volume of firm n for the commodity in question. Then we can divide the 400,000 tons over the 6 relations proportionally to the share of a mn relation's product in the sum of *the products over all 6 mn relations. The sum of the allocated flows over the 6 relations will equal 400,000 tons (preservation of PC flow).*

Sometimes, the data we use on firms in the domestic zones (production and consumption files) do not include any producing or consumption firm in a zone for a commodity type for which there is z2z information in the PWC base matrices. In these cases we have generated a single artificial firm (sender or receiver) for that commodity in that zone.

Furthermore we use some extra rules to prevent getting too many small f2f flows. For export and import flows we have no information on firms at the foreign end. We assume that there will only be one sender per zone for import and only one receiver per zone for export.

Chapter 5

Dynamic Optimization and Differential Stackelberg Game Applied to Freight Transport

Terry L. Friesz, Amir H. Meimand and Bo Zhang

Abstract

In this chapter the dynamic shipper–carrier problem is modeled as a differential Stackelberg game. The shippers act as followers operating under the Cournot–Nash behavioral assumption while competing on the sale of a homogenous product in several markets. The followers choose their strategies simultaneously by maximizing their respective utility functions while assuming their competitors' strategies are fixed. The leader in this problem is a transportation company referred to as the carrier who seeks to maximize its objective function while considering the followers' reactions. We show that the differential Nash game describing the shippers' competition may be articulated as a differential variation inequality (DVI). The DVI is then reformulated as a nonlinear complimentarily problem (NCP) to characterize the follower's Nash–Cournot equilibrium by the implicit solution of a system of equations. This allows the shippers' response to the carrier's policy to be embedded into the carrier's problem as a set of nonlinear constraints. The penalty method is then used to reduce the number of explicit constraints and finite-dimensional time discretization is employed to approximate the model as a mathematical program which may be solved by the multistart global optimization scheme found in the off-the-shelf software package GAMS when used in conjunction with the commercial solver MINOS. We also present a small scale numerical example with two followers and four markets.

Keywords: Freight network models; Shipper–carrier models; network oligopoly; differential variational inequalities; differential stackelberg games; mathematical programs with equilibrium constraints

Freight Transport Modelling

ISBN: 978-1-78190-285-1

5.1. Introduction

Transportation network design analyzes how changes to the transportation system impact freight flows. Harker and Friesz (1986) defines three general approaches used to predict freight flows. These include the econometric model, the spatial price equilibrium model, and the freight network equilibrium model. Our main interest in this chapter lies in presenting a computable model for the latter approach.

The shipper–carrier model studies the interaction between two groups of players: shippers, who are decision makers that ship a commodity to markets within a network, and carriers, who are decision makers that provide transportation services to the shippers. Since the shippers' demands for transportation will generally vary with the quality of service provided by carriers, and the carriers' ability to provide service will be impacted by the scale of services, it is clear that a theoretically precise model must capture the behavior of both types of decision makers simultaneously.

The static shipper–carrier problem has been studied extensively in the literature. The first model considering multiple agents was Friesz, Tobin, and Gottfried (1981). The model considered both shippers and carriers explicitly. This model was extended to include simultaneous shipper–carrier loading (Friesz, Gottfried, & Morlok, 1986). Based on this model, Harker and Friesz (1986) developed the spatial price equilibrium model. Moreover, Fernandez, De Cea, and Soto (2003) presented a model to consider a multimodal shipper who makes shipment decisions based on each mode's expected level of service. They discussed the sufficient conditions for the existence and uniqueness of a solution. Xiao and Yang (2007) studied a partially noncooperative game among shippers, carriers, and infrastructure companies who all seek to maximize their respective profits. In their model, the shipper acts as a noncooperative agent while the carriers and infrastructure companies behave cooperatively. They showed that by using a nonlinear tariff, equilibrium flows can also maximize total system profit. Moreover, Mutlu and Cetinkaya (2011) modeled the shipper–carrier problem using two specific channel structures: the centralized channel, in which the goal is to maximize the total channel profit, and the decentralized channel, in which the shipper and carriers selfishly choose their own policies. Shah and Brueckner (2012) developed an analytical model in which carriers compete by setting the prices and frequencies of shipments. They also considered vehicle carrying capacities, delivery time, and brand loyalty of shippers. The model also provided a possible explanation for the presence of excess capacity in the freight industry as an equilibrium choice made by carriers in the presence of a minimum vehicle capacity constraint.

There are comparatively fewer studies of the dynamic shipper–carrier model. Agrawal and Ziliaskopoulos (2006) presented a dynamic model which reaches equilibrium when no shipper can reduce their cost by changing the carrier of any shipment. An algorithm based on an iterative variational inequality is proposed which uses the market equilibrium as the feedback from the carriers to the shippers. Unnikrishnan, Valsaraj, and Waller (2009) presented a dynamic model to determine the optimal storage capacity at transshipment nodes when demand is stochastic. A two-stage linear programming approach is developed where, in the first stage, the shipper decides the optimal capacity at transshipment nodes. In the second stage, the shipper chooses a routing strategy based on the realized demand.

This chapter will formulate the dynamic shipper–carrier problem to find a Stackelberg–Cournot–Nash equilibria by extending the conceptual framework proposed by Fisk (1983). Stackelberg–Cournot–Nash equilibrium were first introduced by Sherali, Sosyter, and Murphy (1983). In Stackelberg–Cournot–Nash equilibrium, there are $N + 1$ players, N of which are followers competing in a Nash game. Each follower seeks to maximize its objective function under the assumption that the other players' strategies are fixed. The $(N + 1)$th player is the leader and optimizes its objective function while taking into account the reaction of the other N players.

In this chapter the shippers are suppliers who provide a homogenous good; they form an oligopoly with respect to the market for their outputs. As such they are agents in a noncooperative differential Nash game. Furthermore they procure transportation services from a single carrier, who is a Stackelberg leader. The theory of oligopolistic markets enjoys a vast literature, examples of which are Greenhut, Norman, and Hung (1987) and Greenhut and Lane (1989). Chamberlin (1993) introduced a different type of competition called monopolistic competition in which firms supply similar but not identical products which are allowed to be substitutes.

The aim of this chapter is to study a Stackelberg game between one leader and several followers competing under the Cournot–Nash assumption. Intriligator (1971) has presented a mathematical analysis of Stackelberg–Cournot duopoly in the case of linear cost and demand functions. He also derived a closed-form equilibrium solution. Moreover, Miller, Tobin, Friesz (1992) formulated a facility location model for firms competing on a discrete network where the locating firm acts as the leader and the other firms are followers. A hierarchical mathematical programming approach combined with sensitivity analysis is proposed to solve the problem. Miller, Tobin, Friesz, and Kwon (2007) extended the previous model of Miller et al. (1992) to the dynamic model in which the locating firm is an "entering" firm competing with several other already established oligopolistic competitors. Tobin (1992) employed the theory of sensitivity analysis to study the reaction of the Cournot–Nash players to a Stackelberg player. He also investigated the effect of this reaction on the uniqueness of the Stackelberg–Cournot–Nash equilibrium.

In this chapter the dynamic shipper–carrier problem is modeled as a differential Stackelberg game. The shippers act as followers operating under the Cournot–Nash behavioral assumption while competing on the sale of a homogenous product in several markets. The leader in this problem is a transportation company referred to as the carrier who seeks to maximize its objective function while considering the followers' reactions. We show that the differential Nash game describing the shippers' competition may be articulated as a differential variation inequality (DVI). The DVI is then reformulated as a nonlinear complimentarily problem (NCP) to characterize the follower's Cournot–Nash equilibrium by the implicit solution of a system of equations. This allows the shippers' response to the carrier's policy to be embedded into the carrier's problem as a set of nonlinear constraints. The penalty function method is then used to reduce the number of explicit constraints and finite-dimensional time discretization is employed to approximate the model as a mathematical program which may be solved by the multi-start global optimization scheme found in the off-the-shelf software package GAMS when used in conjunction with the commercial solver MINOS. We also present a small scale numerical example with two followers and four markets.

5.2. The Notion of Nash and Stackelberg Equilibria

Game theory provides a framework for modeling the interaction between groups of players whose utility functions and set of feasible strategies are related. There are several behavioral assumptions one may employ in modeling a game. The aim of this section is to present two of these behavioral models: namely, Nash (1951) and Stackelberg (1952) games. Both Nash and Stackelberg games feature noncooperative equilibrium solutions. In such a solution, no player can improve their utility function by deviating from their equilibrium strategy (Simaan, 1977). Many real-world problems can be formulated using these behavioral models, including several applications in supply chain, logistics, and transportation (Dafermos & Nagurney, 1987; Henderson & Quandt, 1980; Friesz, Rigdon, & Mookherjee, 2006).

5.2.1. Nash Equilibrium

Nash (1951) generalized the concept of equilibrium for the behavioral model consisting of N players who cannot improve their own self-interest by deviating from their Nash equilibrium strategy, given that the other players use their equilibrium strategies (Friedman, 1979). Suppose there are N players in a game and each player $i \in [1,N] \subseteq I_{++}$ chooses a feasible strategy tuple x^i from the strategy set Ω_i to maximize the utility function $\Theta_i : \Omega_i \to \Re^1$, which will generally depend on other players' strategies. Every player $i \in [1,N]$ solves his/her best response problem:

$$\begin{array}{ll} \max & \Theta(x^i; x^{-i}) \\ \text{subject to}: & x^i \in \Omega_i \end{array} \tag{5.1}$$

Note that in problem (5.1) we use the notation:

$$x^{-i} = (x^j, j \neq i) \quad i \in [1, N]$$

to refer to the tuple of strategies of players other than i. It is assumed that the non-own tuple x^{-i} is known to player i. A Nash equilibrium is the vector of strategies

$$x = (x^i : i \in [1, N])$$

formed from player-specific tuples such that each x^i solves the mathematical program (5.1). We denote a Nash equilibrium by $NE(\Theta, \Omega)$ (Friesz, 2010).

5.2.2. *Stackelberg Equilibrium*

In the Nash game discussed previously all players act simultaneously. However, in a Stackelberg game, two players labeled as the leader and the follower make decisions sequentially in a noncooperative fashion (Stackelberg, 1952). The leader, who is the first mover, chooses his/her best policy x from the strategy set Ω_L to maximize his/her utility function $\Theta_L : \Omega_L \to \Re^1$ assuming the follower acts rationally. Then, the follower, who is the second mover, determines his/her best policy $y \in \Omega_F$ to maximize his/her utility function $\Theta_F : \Omega_F \to \Re^1$ in response to the policy of leader. A Stackelberg equilibrium is obtained when the leader solves his/her problem taking into account the best response of the follower.

Let x be the vector of the leader's strategy and y be the follower's where $y = \phi(x) : \Omega_F \to \Omega_L$ is the follower's best response for a given x. The leader's problem is:

$$\begin{array}{ll} \max\limits_{x} & \Theta_L(x, \phi(x)) \\ \text{subject to :} & \end{array} \tag{5.2}$$

$$x \in \Omega_L \tag{5.3}$$

or alternatively:

$$\begin{array}{ll} \max\limits_{x} & \Theta_L(x, \phi(x)) \\ \text{subject to :} & \end{array} \tag{5.4}$$

$$y = \phi(x) \tag{5.5}$$

$$x \in \Omega_F \tag{5.6}$$

$$y \in \Omega_L \tag{5.7}$$

Usually it is not possible to obtain $\phi(x)$ explicitly and the constraints (5.5) will take an alternative form which incorporates this relationship implicitly (Fisk, 1983). We denote a Stackelberg equilibrium by $SE(\Theta_L, \Theta_F, \Omega_L, \Omega_F)$.

5.3. Alternative Formulations for Nash Game

In this section we show that a differential Nash game may be articulated as a differential variational inequality (DVI) with clearly distinguishable states and control. The control variables of the DVI are determined not by extremizing behavior of a single player but rather by the Cournot–Nash competition in which each player participate. We also show that the DVI representation of differential Nash game may be reformulated as a functional nonlinear complementarity problem, in an appropriate Hilbert space, to form the foundation for obtaining the explicit form of follower's best response when the leader's policy is given.

To articulate a differential Nash game as a differential variational inequality (DVI) we introduce the Hilbert space which is a complete vector space with well-defined norm and well-defined inner product that induces a norm. The specific function spaces we employ are those that allow optimal control problem to be analyzed as infinite dimensional mathematical programming, and are chosen setting of an infinite dimensional variational inequality.

We begin by following notation:

$$
\begin{aligned}
u^i &\in (\mathcal{L}^2[t_0, t_f])^{m_i} \\
x^i &\in (\mathcal{H}^1[t_0, t_f])^{n_i} \\
x_0^i &\in \mathcal{R}^{n_i} \\
m &= m_1 + m_2 + \ldots + m_N \\
n &= n_1 + n_2 + \ldots + n_N \\
\Theta^i &: (\mathcal{H}^1[t_0, t_f])^n \times (\mathcal{L}^2[t_0, t_f])^m \times \mathcal{R}_+^1 \to (\mathcal{L}^2[t_0, t_f])^m \\
f^i &: (\mathcal{H}^1[t_0, t_f])^n \times (\mathcal{L}^2[t_0, t_f])^m \times \mathcal{R}_+^1 \to (\mathcal{L}^2[t_0, t_f])^n \\
\Gamma^i &: (\mathcal{H}^1[t_0, t_f])^n \times \mathcal{R}_+^1 \to (\mathcal{H}^1[t_0, t_f])^r
\end{aligned}
$$

Note that $(\mathcal{L}_+^2[t_0, t_f])^m$ is the m–fold product of the space non-negative of square-integrable functions $(\mathcal{L}_+^2[t_0, t_f])$ with inner product while $(\mathcal{H}^1[t_0, t_f])^n$ is the n–fold product of the Sobolev space $(\mathcal{H}^1[t_0, t_f])$.

Now we consider a differential Nash game in which there are N agents and every agent $i \in [1, N]$ seeks to solve the optimal control problem:

$$
\min \ J_i(u^i; u^{-i}) = \Gamma^i[x^i(t_f), t_f] + \int_{t_0}^{t_f} \Theta^i(x^i, u^i(t), x^{-i}, u^{-i}, t)dt \tag{5.8}
$$

subject to:

$$
\frac{dx^i}{dt} = f^i(x^i, u^i, t) \tag{5.9}
$$

$$
x^i(t_0) = x_0^i \tag{5.10}
$$

$$
Z^i[x^i(t_f), t_f] = 0 \tag{5.11}
$$

$$
u^i \in \Omega_i \tag{5.12}
$$

for each fixed yet arbitrary non-own control tuple:

$$
u^{-i} = (u^j : j \neq i)
$$

where u^i and x^i are control and state variables respectively. Also, (5.9) is the state dynamic, (5.10) is the initial condition, and (5.11) is the terminal condition for state variable. Moreover, Ω_i is the set of feasible control.

A differential Nash equilibrium denoted by $DNE(\Theta, \Omega)$, is a tuple of strategies $u = (u^i : i = 1, 2, \ldots, N)$ such that each u^i solves the optimal control problem (5.8)–(5.12) assuming u^{-i} is fixed.

5.3.1. DVI Formulation

Differential variational inequalities (DVI) were first formally introduced and comprehensively discussed by Pang and Stewart (2008) as a unifying mathematical framework for a host of dynamic problems involving inequalities and discontinuities. Some applications of the DVI formulation of a dynamic network economy can be found in Brander and Zhang (1993), Nagurney, Dupuis, and Zhang (1994), Wie and Tobin (1997), Friesz et al. (2006), and Markovich (2008). To present the differential variational inequality formulation we use the following notations:

$$y^i = \begin{pmatrix} x^i \\ \lambda^i \end{pmatrix} \quad \text{and} \quad y^{-i} = \begin{pmatrix} x^{-i} \\ \lambda^{-i} \end{pmatrix}$$

where λ^i is the adjoint variables for player i obeying:

$$\frac{d\lambda^i}{dt} = -\nabla_{x^i} H_i(x^i, u^i, \lambda^i, t; x^{-i}, u^{-i})$$

when $H_i(x^i, u^i, \lambda^i, t; x^{-i}, u^{-i})$ is the Hamiltonian associated with optimal control problem (5.8)–(5.12):

$$H_i(x^i, u^i, \lambda^i, t; x^{-i}, u^{-i}) = \Theta^i(x, u, t) + (\lambda^i)^T f^i(x, u, t)$$

We define:

$$g^i = \begin{pmatrix} f^i \\ -\nabla_{x^i} H_i(x^i, u^i, \lambda^i, t; x^{-i}, u^{-i}) \end{pmatrix}$$

$$Z^i[x(t_f), t_f] = \begin{pmatrix} \Gamma^i[x(t_f), t_f] \\ \lambda^i(t_f) - \dfrac{\partial \Gamma^i[x(t_f), t_f]}{\partial x(t_f)} \end{pmatrix}$$

for each $i \in [1,N]$, so that

$$y = (y^i : i = 1, 2, \ldots, N)$$
$$g = (g^i : i = 1, 2, \ldots, N)$$
$$Z = (Z^i : i = 1, 2, \ldots, N)$$

Also,

$$y(t_0) = y_0 = \begin{pmatrix} x(t_0) \\ \lambda(t_0) \ \text{free} \end{pmatrix}$$

In addition we define:

$$G^i(y^i, u^i, t; y^{-i}, u^{-i}) = \nabla_{u^i} H_i(x^i, u^i, \lambda^i, t; x^{-i}, u^{-i})$$
$$G = (G^i : i = 1, \ldots, N)$$

Now, we can show the equivalence of differential variational inequality and differential Nash equilibrium by Theorem 5.1:

Theorem 5.1. *(Differential variational inequality equivalent to differential Nash equilibrium). The differential Nash equilibrium $DNE(\Theta, \Omega)$ is equivalent to the following differential variational inequality, denoted by $DVI(F, f, U, x_0, t_0, t_f)$, when $f^i(x^i, u^i, t)$ and $\Theta_i(x^i, u^i, t)$ are convex and continuously differentiable with respect to (x^i, u^i) for all fixed non-own tuples (x^{-i}, u^{-i}) for all $i \in [1, N]$.*

$$\textit{find} \ \ u^* \in \Omega \tag{5.13}$$

such that

$$\int [G(y(u^*, t), u^*, t)]^T (v - u^*) dt \geq 0 \quad \forall v \in \Omega \tag{5.14}$$

where

$$y(u, t) = \arg\left\{ \frac{dy}{dt} = g(y, u, t), y(t_0) = y_0, Z[y, y(t_f), t_f] = 0 \right\} \tag{5.15}$$

Proof. See Friesz (2010). ■

5.3.2. Nonlinear Complementary Formulation

A DVI is an abstract form of an infinite set of constraints. Several numerical methods to solve DVI exist in the literature, including Gap function, Nonlinear Complementarity, and Fixed-Point method (Friesz, 2010).

We are interested in converting a DVI into a finite set of equations. This conversion is possible when all constraints on the control set U are linear, that is,

$$U = \{u \geq 0 : Au \leq b\} \tag{5.16}$$

where $b \in \Re^l$ is a constant vector and

$$A = (a_{ij})$$

is a constant $l \times m$ matrix.

We restate $DVI(F, f, U, x_0, t_0, t_f)$ by converting it into a nonlinear complementarity problem in infinite dimensions. That conversion is based on the Kuhn-Tucker conditions for each player's problem, that is,

$$\nabla_u H(x, u, \lambda, t) + \sum_{j=1}^{m} \rho_j \nabla_u(-u_j) + \sum_{j=1}^{l} \xi_j \nabla_u (Au - b)_j = 0 \tag{5.17}$$

$$\rho_j u_j = 0 \tag{5.18}$$

$$\rho_j \geq 0 \tag{5.19}$$

$$\xi_j (Au - b)_j = 0 \tag{5.20}$$

$$\xi_j \geq 0 \tag{5.21}$$

when (ρ_j, ξ_j) is a pair of dual variables associated with constraints (5.16) for all $j \in [1, m]$. Also for each $i = 1,2, \ldots m$,

$$F_i(x, u, t) + \sum_{j=1}^{n} \lambda_j \frac{\partial}{\partial u_i} f_i(x, u, t) + \sum_{j=1}^{l} \xi_j \frac{\partial}{\partial u_i} (Au - b)_j = \rho_i \geq 0 \tag{5.22}$$

Since $\rho_i u_i = 0$ due to complementary slackness and $(\partial/\partial u_i)(Au - b)_j = a_{ij}$, we have

$$\left[F_i(x, u, t) + \sum_{j=1}^{n} \lambda_j \tfrac{\partial}{\partial u_i} f_i(x, u, t) + \sum_{j=1}^{l} a_{ij} \xi_j \right] u_j = \rho_j u_j = 0 \tag{5.23}$$

for each $i = 1, 2, \ldots m$. Thus, we arrive at the following functional nonlinear complementarity problem:

$$\langle G(z, t), z \rangle = 0 \tag{5.24}$$

$$G(z, t) \geq 0 \tag{5.25}$$

$$z \geq 0 \tag{5.26}$$

where

$$z \in (\mathcal{L}^2[t_0, t_f])^{2m+l}$$
$$G : (\mathcal{L}^2[t_0, t_f])^{2m+l} \times \Re^1_+ \rightarrow (\mathcal{L}^2[t_0, t_f])^{2m+l}$$

and

$$G = \begin{pmatrix} F(x,u,t) + [\nabla_u f(x,u,t)]^T.\lambda + [\nabla_u(Au-b)]^T.\xi \\ Au-b \\ u \end{pmatrix}, z = \begin{pmatrix} u \\ \xi \\ \rho \end{pmatrix}$$

for which it is understood that $x(\cdot)$ and $\lambda(\cdot)$ are operators obeying (5.15). Note that F(x, u, t) and $f(x, u, t)$ are linear in $u(\cdot)$, the reformulation (5.24), (5.25), and (5.26) yields a linear complementarity problem.

5.4. Penalty Method for Solving Stackelberg Game

In Stackelberg games, leaders solve their problem by considering the followers' responses as a set of constraints (Fisk, 1986). If there is only one leader, optimal solutions are found by solving an optimal control problem with nonlinear constraints. In other words, the lower level problem can be formulated as a NCP; then the Stackelberg problem is converted to a single optimal control with nonlinear constraints. One possible method for solving such a problem is the penalty function method.

5.4.1. Penalty Function Method

A penalty function in nonlinear programming is a function that is appended to the objective function and assigns a penalty when a given constraint is violated. This reduces the number of explicit constraints and can be used to handle particularly problematic constraints, such as nonlinear constraints or constraints that destroy special structure. When we want to convert the Stackelberg game, which is bi-level mathematical program, to a single level program, we will encounter nonlinear constraints. These nonlinear constraints emerge from the NCP formulation of the lower level problem. Our main goal in this section is to employ the penalty function method to attach the nonlinear constraints to the objective function and produce a nonlinear program with linear constraints. The NCP (5.24), (5.25), and (5.26) lead directly to the functional mathematical program:

$$\begin{aligned} &\min \qquad J = \langle G(z,t), z \rangle \equiv \sum_{i=1}^{m} \int G_i(z,t) z dt \\ &\text{subject to :} \\ &\qquad G(z,t) \geq 0 \\ &\qquad z \geq 0 \end{aligned} \tag{5.27}$$

Considering problem (5.4)–(5.7), the Stackelberg game can be reformulated as:

$$
\begin{aligned}
&\max_x \qquad \Theta_L(x,z) - \sum_{i=1}^{m} \mu_i \phi_i(z,t) \\
&\text{subject to}: \\
&\qquad \phi_i(z,t) = \int_{t_0}^{t_f} G_i(z,t) z_i dt \qquad \forall i = 1, \ldots, m \\
&\qquad G(z,t) \geq 0 \\
&\qquad z \geq 0 \\
&\qquad x \in \Omega_F
\end{aligned}
\tag{5.28}
$$

where $\mu_i \in \Re^1_{++}$ is a large number called penalty parameter. When μ_i approaches to infinity we would expect the optimal solution of the above problem to enforce $\phi_i(z,t)$ close to zero, because otherwise a large penalty $\mu_i \phi_i(z,t)$ will be incurred. Most algorithm using penalty functions employ a sequence of increasing penalty parameters. With each new value penalty parameter, an optimization technique is used, starting from the optimal solution obtained for the parameter value chosen previously. We present below the algorithm composed of three main steps to solve Eq. (5.28):

Step 0. *Initialization*: Set $k = 0$ and specify (x^0, z^0), an initial solution. Also set values of μ_i^0 for all $i \in [1, m]$.

Step 1. *Solve the problem involving penalty function*: Form and solve the problem (5.28) and call the solution (x^{k+1}, z^{k+1}).

Step 2. *Stopping criteria*: If $\varepsilon > 0$ is a pre-set scalar tolerance and

$$\sum_{i=1}^{m} \mu_i^{k+1} \phi_i(z^{k+1}, t) \leq \varepsilon$$

then stop and declare the optimal solution is

$$(x^*, z^*) \approx (x^{k+1}, z^{k+1})$$

Otherwise set

$$\mu_i^{k+1} = \beta \mu_i^k$$

for all $i \in [1, m]$ where β is a scalar and $\beta > 1$ and go to Step 1.

5.5. The Shipper–Carrier Problem

In this section we present the application of differential games to a major class of predictive intercity freight models: freight network equilibrium models. This class of models focuses on the interaction between shippers and carriers (Harker, 1984). Specifically, in this section we present a model of a Stackelberg game where the shippers behave as followers who compete with each other to sell a homogeneous commodity in several markets and make decisions about where to sell and how to ship commodity between Origin and Destination (OD) pairs and a single transportation company referred to as the carrier is the leader in this model. The carrier generally operates a fleet of vehicles and receives requests from shippers to transport the commodity between OD pairs.

5.5.1. Notations

We employ the notation used in Friesz et al. (2006) augmented to handle temporal considerations. Time in general is denoted by the scalar $t \in \Re^1_+$, and the initial and final times are denoted by $t_0 \in \Re^1_+$, and $t_f \in \Re^1_+$, respectively, with $t_0 < t_f$ so that $t \in [t_0, t_f]$. There are several sets important to articulating the shipper–carrier model on a network; these are as follow: $\mathcal{S}$ for shippers, $\mathcal{A}$ for directed arcs, $\mathcal{P}$ for paths, $\mathcal{N}$ for nodes, and $\mathcal{W}$ for OD pairs. It is meaningful by using the following subscripts or superscripts s for a specific shipper, i for a specific node, p for a specific path and w for a specific OD pair. The carrier decides about the transportation price for OD pairs $(i,j) \in \mathcal{W}$, Ψ_{ij} as well as arc flow, f_a for all $a \in \mathcal{A}$, and path flow, h_p for all $p \in \mathcal{P}$. Each shipper controls the allocation of available commodity to sell, D^f and shipping flows, S^f. Inventories, I^f are state variables determined by the controls.

5.5.2. The Carrier's Objective Functional and Constraints

The carrier tries to maximize its net profit $\Phi(\Psi_{ij}, f_a; v^*_{ij})$ which is the difference between the carrier's revenue and transportation delay cost:

$$\max \ \Phi_c(\Psi_{ij}, f_a; v^*_{ij}) = \int_{t_0}^{t_f} e^{-\rho t} \left\{ \sum_{s \in S} \sum_{(i,j)} \Psi_{ij}(t) v^{s*}_{ij}(t) - \sum_{a \in A} D_a(t) f_a(t) \right\} dt \tag{5.29}$$

where $\rho \in \Re^1_{++}$ is a constant nominal rate of discount, $\Psi_{ij}(t) \in \Re^1_{++}$ denotes the transportation price per unit of flow v_{ij} for OD pair $(i,j) \in \mathcal{W}_s$ at time $t \in [t_0, t_f]$, $v^{s*}_{ij}(t)$ is the shipping flow for OD pair $(i,j) \in \mathcal{W}_s$ determined by the shipper s. We use the notation $f_a(t)$ to denote the flow of commodity on arc $a \in \mathcal{A}$. We also assume the arc travel time function $D_a(t)$ is given. Moreover, the flow of commodity on path $p \in P$ traversing arc a at time t will be denoted by $h_p(t)$. We employ the usual arc path incidence matrix where:

$$\delta_{ap} = \begin{cases} 1 & \text{if arc } a \text{ belongs to path } p \\ 0 & \text{otherwise} \end{cases}$$

So the arc flow is given by:

$$f_a(t) = \sum_{p \in P_{ij}} \delta_{ap} h_p(t) \quad \forall a \in A, \forall (i,j) \in \mathcal{W}$$

Also, the arc flow is nonnegative and bounded from above, that is,

$$f_a(t) \leq B_a, \ \ \forall a \in \mathcal{A}$$

The total flow on paths connecting OD pair (i, j) must be equal to the total shipping flow for the same OD pair:

$$\sum_{p \in P_{ij}} \int_{t_0}^{t_f} h_p(t) = \sum_{s \in S} \int_{t_0}^{t_f} v_{ij}^{s*}(t) \quad \forall (i,j) \in \mathcal{W}$$

There are also positive upper and lower bounds, based on market regulation on transportation price charged by carrier. Thus, we have:

$$M_{ij} \leq \Psi_{ij}(t) \leq T_{ij}$$

where $M_{ij} \in \Re^1_{++}$ and $T_{ij} \in \Re^1_{++}$ are known constants.

5.5.3. *The Shippers' Objective Functional, Dynamics, and Constraints*

Each shipper has the objective of maximizing net profit expressed as revenue less cost and takes the form of an operator acting on allocations of available commodity to sell and shipment flow for OD pairs. For each shipper $s \in S$, net profit is given by the following functional:

$$\max \ \Phi_s(d_i^s, v_{ij}^s; d_i^{s-}, v_{ij}^{s-}) = \int_{t_0}^{t_f} e^{-\rho t} \left\{ \sum_{i \in N} \pi_i(t, \sum_{g \in \mathcal{S}} d_i^g) d_i^s(t) - \sum_{(i,j) \in W} \Psi_{ij}^*(t) v_{ij}^s(t) - \sum_{i \in \mathcal{N}} \psi_i^s(I_i^s, t) \right\} dt \tag{5.30}$$

again $\rho \in \Re^1_{++}$ is a constant nominal discount rate, ψ_i^s is shipper s's inventory cost at node i, and I_i^s is the inventory/backorder of shipper s at node i. In Eq. (5.30), d_i^s is the quantity of commodity of shipper $s \in \mathcal{S}$ provided for sale at node $i \in \mathcal{N}$. Also,

$$Z_s[I^s(t_f), t_f]$$

is the liquidation value of inventory remaining at the terminal time, where $I^s = (I_i^s : i \in \mathcal{N}_f)$. Because our formulation is in terms of flows, it is convenient to employ the inverse demand functions $\pi_i(t, d_i)$ where

$$d_i(t) = \sum_{g \in \mathcal{S}} d_i^g(t)$$

is the total quantity of commodity for sale at node i. The reader should note that $\Phi_s(d^s, v^s; d^{-s}, v^{-s})$ is a functional that is determined by the controls $d^s(t)$ and $v^s(t)$ of shipper s when other shippers' decisions $d^{-s} \equiv (d^{s'} : s' \neq s)$, $v^{-s} \equiv (v^{s'} : s' \neq s)$ are taken to be exogenous data to shipper s. Further, $\Psi_{ij}(t)$ for OD pair $(i,j) \in \mathcal{W}$ is announced by the leader. The first term of the objective functional $\Phi_s(d^s, v^s; d^{-s}, v^{-s})$ in expression (5.30) is the shipper's revenue, the second term is the shipper's cost of transportation, and the third term is the shipper's inventory holding cost.

We next impose the terminal time inventory constraints:

$$I_i^s(t_f) = \tilde{K}_i^s \quad \forall s \in \mathcal{S}, i \in \mathcal{N}_s \tag{5.31}$$

where $\tilde{K}_i^s \in \Re^1_{++}$ are exogenous. The controls are nonnegative and bounded from above, that is,

$$0 \leq d^s(t) \leq D^s \tag{5.32}$$

$$0 \leq v^s(t) \leq V^s \tag{5.33}$$

where

$$D^s \in \Re^{|S|}_{++}, V^s \in \Re^{|S|}_{++}$$

Constraints (5.32) and (5.33) are recognized as pure control constraints, while Eq. (5.31) is terminal condition for the state space variables. Naturally

$$\Omega_s = \left\{(d^s, v^s) : \text{Eqs. } (32), (33)\right\}$$

is the set of feasible controls.

Shipper s solves an optimal control problem to determine its quantity of commodity for sale d^s and shipping demand v^s by maximizing its profit functional $\Phi_s(d^s, v^s; d^{-s}, v^{-s})$ subject to inventory dynamics expressed as flow balance equations. The inventory dynamics for shipper $s \in \mathcal{S}$, expressing simple flow conservation, obey

$$\frac{dI_i^s}{dt} = \sum_{j \in N^+(i)} v_{ji}^s(t) - \sum_{j \in N^-(i)} v_{ij}^s(t) - d_i^s(t) \quad \text{for all } i \in N \tag{5.34}$$

$$0 \le d_i^s(t) \le D_i^s, 0 \le v_{ij}^s(t) \le V_{ij}^s \quad \text{for all } i \in N, \text{for all } (i,j) \in W$$

$$I_i^s(t_0) = K_i^s \quad \forall\, i \in \mathcal{N}_s \tag{5.35}$$

$$I_i^s(t_f) = \underset{i}{\tilde{K}} \quad \forall\, i \in \mathcal{N}_s \tag{5.36}$$

where $K_i^s \in \Re_{++}^1$ and $\tilde{K}_i^s \in \Re_+^1$ are exogenous. In addition to the terminal time inventory (state) constraints (5.31), the model is general enough to handle inventory constraints over the entire planning horizon $[t_0, t_f]$. For instance, non-negativity of the inventory (state) variables could be imposed to restrict shippers from taking backorders.

We now state shipper *s*'s problem. With v^{-s} and d^{-s} as exogenous inputs, compute v^s and d^s (thereby finding I^s) by solving the following extremal problem:

$$\left.\begin{array}{ll} \max & \Phi_s(d^s, v^s; d^{-s}, v^{-s}) \\ \text{Subject to}: & (d^s, v^s) \in \Omega_s \end{array}\right\} \forall s \in \mathcal{S} \tag{5.37}$$

where

$$\Omega_s = \{(d^s, v^s) \ (\text{Eqs. } (5.32) - (5.36))\}$$

also for all $s \in \mathcal{S}$. That is, each shipper is a Nash agent that knows and considers the current decisions of other shippers to make its own noncooperative decisions. As such, Eq. (5.37) is a differential Nash game. Moreover, each shipper's best response problem (5.37) is a continuous time optimal control problem.

5.5.4. DVI Formulation of the Lower Level Problem

To continue our discussion of the shipper problem, we assume the Nash game expressed above is regular in the sense of the following definition:

Definition 5.1. *The dynamic shipper competition introduced above will be considered regular if: (i) the state operator $I(v, d)$ exists and is unique while each of its components is continuous and G-differentiable, (ii) the inverse demand and inventory cost functions are continuously differentiable with respect to controls and states; (iii) for each $s \in \mathcal{S}$, the composite terminal cost function*

$$Z_s[I^s(t_f), t_f] + \sum_{i \in \mathcal{N}s} \gamma_i^s[\tilde{K}_i^s - I_i^s(t_f)]$$

is continuously differentiable with respect to $I_i^s(t_f)$ for all $i \in \mathcal{N}_s$.

In the above definition, each γ_i^s is a constant dual variable that prices out the terminal constraint on inventory.

We also note that Eq. (5.37) is an optimal control problem with fixed terminal time. Its Hamiltonian is:

$$H_s(d^s, v^s, I^s, \lambda^s; d^{s-}, v^{s-}) = \Phi_s(d^s, v^s, I^s, \lambda^s; d^{s-}, v^{s-}) + \Gamma_s(d^s, v^s, I^s, \lambda^s) \tag{5.38}$$

where

$$\Gamma_s(v^s, d^s, I^s, \lambda^s) = \sum_{i \in \mathcal{N}s} \lambda_i^s(t) \left(\sum_{j \in N^+(i)} v_{ji}^s(t) - \sum_{j \in N^-(i)} v_{ij}^s(t) - d_i^s(t) \right) \tag{5.39}$$

and $\lambda_i^s(t) \in \Re_+^1$ is the adjoint variable corresponding to the dynamics of shipper s at node i where $\lambda^s \in (\mathcal{H}^1[t_0, t_f])^{|\mathcal{N}|}$. Clearly Φ_s is the instantaneous profit. To interpret Ψ_s, we need to understand the relevant dynamic shadow benefits and shadow costs. To that end, recall that, along an optimal trajectory, adjoint variables obey

$$\lambda_i^s = \frac{\partial J_s}{\partial I_i^s}$$

Consequently,

$$\Gamma_s = \sum_{i \in \mathcal{N}_s} \frac{\partial J_s}{\partial I_i^s} \frac{dI_i^s}{dt}$$

which is recognized as the shadow value of dynamic benefits arising from current inventory held; it can be either a cost or a benefit, depending on its sign.

Due to the regularity conditions, the Pontryagin maximum principle (Sethi and Thompson, 2000) can be applied here and consequently we have the nonlinear program

$$\max \ H_s \quad \text{subject to}: 0 \leq (v^s, d^s) \leq (V^s, D^s)$$

for every shipper $s \in \mathcal{S}$ for all time $t \in [t_0, t_f]$. Consequently, since the feasible set is convex, the finite-dimensional variational inequality principle from the necessary conditions requires any optimal solution to satisfy:

$$\frac{\partial H_s^*}{\partial d_i^s} (d_i^s - d_i^{s*}) \leq 0 \tag{5.40}$$

$$\frac{\partial H_v^*}{\partial v_{ij}^s} \left(v_{ij}^s - v_{ij}^{s*} \right) \leq 0 \tag{5.41}$$

for every $s \in \mathcal{S}$ at every time $t \in [t_0, t_f]$. According to Theorem 5.2 which is stated below, the following variational inequality has solutions that are differential Nash equilibria for a noncooperative game in which individual shippers maximize net profits in light of current information about their competitors:

$$\text{find } \ (v^{s*}, d^{f*}) \in \Omega \ \text{ such that}$$

$$\sum_{s\in\mathcal{S}}\int_{t_0}^{t_f}\left[\sum_{i\in\mathcal{N}_s}\frac{\partial\Phi_s^*}{\partial v_i^s}(v_i^s-v_i^{s*})+\sum_{i\in\mathcal{N}_s}\frac{\partial\Phi_s^*}{\partial d_i^s}(d_i^s-d_i^{s*})\right]dt\leq 0 \quad \forall(d,s)\in\Omega \tag{5.42}$$

where

$$\Phi_s^*=\Phi_s(v^{s*},d^{s*};v^{-s},d^{-s};t) \tag{5.43}$$

$$\Omega=\prod_{s\in\mathcal{S}}\Omega_s \tag{5.44}$$

We note that Eq. (5.42) is a differential variational inequality expressing the differential Nash game that is our present interest.

The issue of immediate concern is to formally demonstrate that solutions of Eq. (5.42) are differential Nash equilibria. In fact, we state and prove the following result:

Theorem 5.2. *(Differential variational inequality formulation of dynamic shipper competition.) Any solution of Eq. (5.42) is a solution of the dynamic shipper competition problem when regularity in the sense of Definition 5.1 holds.*

Proof. We follow the presentation in Friesz (2010) and begin by noting that Eq. (5.42) is equivalent to the following optimal control problem:

$$\max\ G(v,d)=\sum_{s\in\mathcal{S}}\int_{t_0}^{t_f}\left[\sum_{i\in\mathcal{N}_s}\frac{\partial\Phi_s^*}{\partial d_i^s}d_i^s+\sum_{(i,j)\in\mathcal{W}}\frac{\partial\Phi_s^*}{\partial v_{ij}^s}v_{ij}^s\right]dt$$

$$\text{subject to}:\quad (5.33)-(5.36)$$

where it is essential to recognize that $G(v, d)$ is a linear functional that assumes knowledge of the solution to our Nash game; as such, $G(v, d)$ is a mathematical construct for use in analysis and has no meaning as a computational device. The augmented Hamiltonian for this artificial optimal control problem is

$$H_0=\sum_{s\in\mathcal{S}}\left[\sum_{i\in\mathcal{N}_s}\frac{\partial\Phi_s^*}{\partial d_i^s}d_i^s+\sum_{(i,j)\in\mathcal{W}}\frac{\partial\Phi_s^*}{\partial v_{ij}^s}v_{ij}^s\right]+\sum_{s\in\mathcal{S}}\Psi_s$$

Applying maximum principal, we have

$$\max H_0\quad\text{subject to}:\quad (V^s,D^s)\geq(v^s,d^s)\geq 0$$

The corresponding necessary conditions for this mathematical program are identical to Eqs. (5.30) through (5.36), since

$$\frac{\partial H_0^*}{\partial d_i^s}=\frac{\partial\Phi_s^*}{\partial d_i^s}+\frac{\partial\Psi_s^*}{\partial d_i^s}=\frac{\partial H_s^*}{\partial d_i^s}$$

$$\frac{\partial H_0^*}{\partial v_{ij}^s} = \frac{\partial \Phi_s^*}{\partial v_{ij}^s} + \frac{\partial \Psi_s^*}{\partial v_{ij}^s} = \frac{\partial H_s^*}{\partial v_{ij}^s}$$

where

$$H_0^* = \sum_{s \in \mathcal{S}} \left[\sum_{i \in \mathcal{N}_s} \frac{\partial \Phi_s^*}{\partial d_i^s} d_i^{s*} + \sum_{(i,j) \in \mathcal{W}} \frac{\partial \Phi_s^*}{\partial v_{ij}^s} v_{ij}^{s*} \right] + \sum_{s \in \mathcal{S}} \Psi_s^*$$

and

$$\Psi_s^* = \Psi_s(d^{s*}, v^{s*}, I^{s*}, \lambda^{s*})$$

We next note that the following existence result holds: ■

Theorem 5.3. *(Existence of dynamic shipper equilibrium.) When the variational inequality Eq. (5.42) is regular in the sense of Definition 5.1 and Ω is convex and compact, there exists a solution to the dynamic shipper competition problem.*

Proof. By the assumption of regularity, $I(v, d)$ is well defined and continuous. Thus, $\Phi \equiv (\Phi_s : s \in S)$ is continuous in I. Also, by regularity, we know Ω is convex and compact. Consequently, by Theorem 6.3 in Friesz (2010), Eq. (5.42) has a solution. As a result of Theorem 5.2, there exists a solution to the dynamic shipper competition problem. ■

5.5.5. Single Level Optimization Problem

To express the Stackelberg model as a single level mathematical program, the shippers' best response problem might be stated implicitly and included in the model as constraints (Fisk, 1986). Lets define:

$$G = \begin{pmatrix} \nabla_d H(v, d, I, \lambda) + [\nabla_d (d - d_{\max})]^T . \xi \\ \nabla_v H(v, d, I, \lambda) + [\nabla_v (v - v_{\max})]^T . \xi \\ d - d_{\max} \\ v - v_{\max} \\ d \\ v \end{pmatrix} \quad \text{and} \quad z = \begin{pmatrix} d \\ v \\ \xi_v \\ \xi_d \\ \rho_d \\ \rho_v \end{pmatrix} \tag{5.45}$$

So, the DVI (5.42) stating the condition for Nash equilibrium between shippers can be expressed in the form of the following NCP problem:

$$\langle G(z, t), z \rangle = 0 \tag{5.46}$$

$$G(z,t) \geq 0 \tag{5.47}$$

$$z \geq 0 \tag{5.48}$$

We may now formulate the Stackelberg game between the shippers and the carrier as the following single level optimization problem employing the penalty function:

$$\max \quad J^c(\Psi_{ij}, f_a, v_{ij}, d) = \int_{t_0}^{t_f} e^{-\rho t}\left\{\sum_{s\in S}\sum_{(i,j)}\Psi_{ij}(t)v_{ij}^s(t) - \sum_{a\in A} D_a(t)f_a(t)\right\}dt$$
$$-\sum_{i=1}^{m}\mu_i\int_{t_0}^{t_f} G_i(z,t)z_i dt \tag{5.49}$$

$$\sum_{p\in P_{ij}}\int_{t_0}^{t_f} h_p(t) = \sum_{s\in S}\int_{t_0}^{t_f} v_{ij}^s(t) \qquad \forall (i,j)\in \mathcal{W} \tag{5.50}$$

$$f_a(t) = \sum_{p\in P_{ij}} \delta_{ap}h_p(t) \quad \forall a\in A, \forall (i,j)\in \mathcal{W} \tag{5.51}$$

$$G(z,t) \geq 0, \quad z \geq 0 \tag{5.52}$$

$$M_{ij} \leq \Psi_{ij}(t) \leq N_{ij} \qquad \forall (i,j)\in \mathcal{W}, \; f_a(t) \leq B_a \quad \forall a \in \mathcal{A} \tag{5.53}$$

Problem (5.49)–(5.53) can be solved by the algorithm described in Section 5.4. Moreover, since there are no time shifts and all constraints are linear in the formulation proposed above, time discretization and finite-dimensional mathematical programming is especially appealing to solve the mathematical program in step 1 of each iteration of the penalty function method discussed in Section 5.4.1. If we define the discrete instant of time $t_k = t_0 + k\Delta t$, where Δt is the time step employed, $t_N = t_f$, and

$$N = \frac{t_f - t_0}{\Delta t}$$

is the number of discretizations, then a finite dimensional mathematical program is created. This forms a computable approximate solution if the objective function is convex.

5.5.6. *Numerical Example*

Let us consider a network consisting five arcs, four nodes. Two shippers offers identical commodity at every node. Figure 5.1 illustrates this network:

The time interval of interest is [0, 20]; that is $t_0 = 0$, $t_f = 20$. In this example, shippers 1 and 2 are interested in transporting their commodity between OD pairs

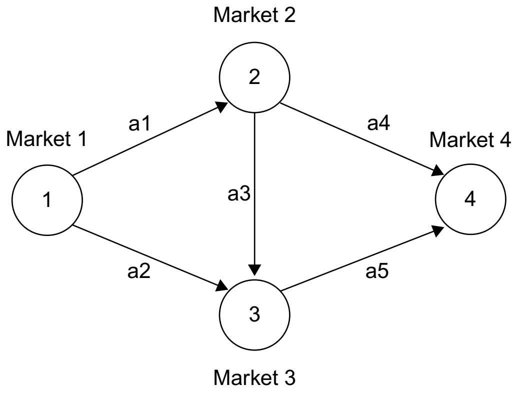

Figure 5.1: Network of Five Arcs 4 Nodes.

within sets $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively, where $\mathcal{W}_1 = \{(1,2),(1,3),(2,4),(3,4)\}$ and $\mathcal{W}_2 = \{(1,2),(2,3),(2,4),(3,4)\}$.

The initial inventory at each node for each shipper is:

$$\begin{array}{llll} I_1^1(0) = 250 & I_2^1(0) = 70 & I_3^1(0) = 10 & I_4^1(0) = 10 \\ I_1^2(0) = 150 & I_2^2(0) = 100 & I_3^2(0) = 50 & I_4^2(0) = 50 \end{array} \tag{5.54}$$

In addition, we impose the condition that no backordering is allowed by any shipper at any node at the terminal time t_f. That is,

$$I_i^s(20) = 15 \quad \forall s \in \mathcal{S}, \ \forall i \in \mathcal{N}_s \tag{5.55}$$

The inventory dynamics are the flow balance equations:

$$\begin{aligned} \frac{dI_1^1}{dt} &= -v_{12}^1 - v_{13}^1 - d_1^1 \\ \frac{dI_2^1}{dt} &= v_{12}^1 - v_{24}^1 - d_2^1 \\ \frac{dI_3^1}{dt} &= v_{13}^1 - v_{34}^1 - d_3^1 \\ \frac{dI_4^1}{dt} &= v_{24}^1 + v_{34}^1 - d_4^1 \\ \frac{dI_1^2}{dt} &= -v_{12}^2 - d_1^2 \\ \frac{dI_2^2}{dt} &= v_{12}^2 - v_{24}^2 - v_{23}^2 - d_2^2 \end{aligned} \tag{5.56}$$

$$\frac{dI_3^2}{dt} = v_{23}^2 - v_{34}^2 - d_3^2$$

$$\frac{dI_4^2}{dt} = v_{24}^2 + v_{34}^2 - d_4^2$$

We assume the inverse demand functions at each node i take the following form:

$$\pi_i(d_i, t) = \alpha_i - \beta_i(d_i)^m \tag{5.57}$$

where $m \in \Re_{++}^1$ is a constant. Also, $\alpha_i \in \Re_{++}^1$ and $\beta_i \in \Re_{++}^1$ for all i are constants. We assume that holding costs are quadratic and of the form:

$$\psi_i^s = \frac{1}{2}\eta_i^s(I_i^s)^2 \quad \forall s \in \mathcal{S}, \quad \forall i \in \mathcal{N}_s \tag{5.58}$$

where $\eta_i^s \in \Re_{++}^1$ are constants, again for all i and s. Assuming the quadratic holding cost function is quite natural since opportunity costs usually increase more than proportionally to the amount of capital left inactive in cash (Baccarin, 2002).

The relationships between arc and path variables are summarized in the Table 5.1: Hence the relevant arc-path incidence matrix is

$$\Delta_p = (\delta_{ap}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and the travel time function for each arc is:

$$D_{a_i} = A_{a_i} + B_{a_i}(f_{a_i})^n \; i = 1, 2, 3, 4, 5 \tag{5.59}$$

Table 5.1: Path-arc

Path	**Arc sequence**
P_1	a_1
P_2	a_4
P_3	a_3, a_5
P_4	a_3
P_5	a_2
P_6	a_1, a_3
P_7	a_5

where, $A_a \in \Re^1_{++}$ and Also, $B_a \in \Re^1_{++}$ are known constants. We impose the following vectors of bounds on the control variables:

$$D^s = 25 \; \forall s \in S$$

$$V^s = 15 \; \forall s \in S$$

$$T_{ij} = 200 \; \forall (i,j) \in \mathcal{W}$$

$$M_{ij} = 800 \; \forall (i,j) \in \mathcal{W}$$

Each shipper's instantaneous profit function is found by substituting Eqs. (5.57) and (5.58) into Eq. (5.30). Also, the carrier's instantaneous profit function is found by substituting Eq. (5.59) into Eq. (5.29). A discrete-time approximation of the corresponding problem (5.49)–(5.53) is created using are $N = 21$ equal time steps. The resulting nonlinear program is solved using GAMS with the MINOS solver. The numerical values of the models parameters are presented in Table 5.2:

5.5.7. *Computational Performance and Numerical Results*

As mentioned earlier, the .nite-dimensional time discretization method is employed to solve this problem. The model is solved by GAMS in conjunction with the MINOS solver. The optimal solution is reported after 108 iterations of GAMS and the run time is less than 1 minute using a desktop computer with a single Intel Pentium processor and 1 GB RAM.

In this section we provide numerical results in graphical form. The net present value (NPV) of revenue generated by the end of planning horizon are shown in Table 5.3.

Table 5.2: Model's parameters

Parameter	**Value**	**Parameter**	**Value**	**Parameter**	**Value**
A_1	1	B_1, B_2	0.5	η_1^1	4
A_2	1	B_3	0.9	$\eta_2^1, \eta_1^2, \eta_2^2$	1
A_3	2	B_4, B_5	0.5	η_3^1, η_3^2	2
A_4	1	β_1	12	η_4^1, η_4^2	0.5
A_5	1	β_2	16	ρ	0.01
α_1	2100	β_3	14	N	20
α_2	2200	β_4	18	t_0	1
α_3	2400	n	1	t_f	20
α_4	2500	m	1	Δ	1

Table 5.3: NPV of instantaneous revenues

Player	NPV of revenue
Carrier (Leader)	3.58407×10^5
Shipper 1 (Follower)	2.33761×10^5
Shipper 2 (Follower)	1.26822×10^5

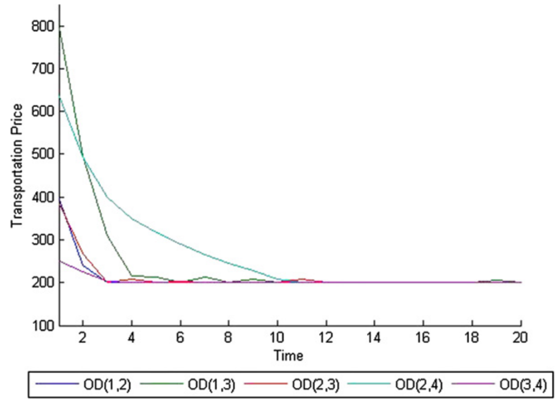

Figure 5.2: Price Trajectories for Transportation between OD Pairs.

Figure 5.2 shows the transportation price trajectories for the OD pairs set by the carrier, and Figure 5.3 depicts the shippers' response to the transportation price. We may observe the sole carrier adopts a decreasing-price policy over time (such a finding is not intuitive) when there are upper and lower bounds for prices.

We also observe since both shippers must meet their terminal inventory condition, the carrier takes advantage of this fact that they have less inventory in nodes 3 and 4 and have to transport some products to these nodes. As a result we observe the two highest transportation cost for OD pairs (1; 2) and (2; 4).

Inventory trajectories are plotted in Figure 5.4, and show that none of the shippers adopt a policy of backordering; this backordering behavior could be represented by a negative inventory level. Figure 5.3 presents the sales of commodity at different markets over time. Each shipper follows a different sales pattern for each node.

Figure 5.4 illustrates that both shippers' inventory levels at node 3 and 4 increase till period four. This increase is caused by the commodity transported from node 1 to 3 and node 2 to 4 by shipper 1 and from node 2 to 3 and 2 to 4 by shipper 2. All the while, inventory levels at node 3 and 4 decrease.

Next we will highlight an interesting application of model presented above. In particular, an oligopolist may conduct numerical experiments of alternative configuration of inventory holding sites. The oligopolist may also decline demand at specific nodes.

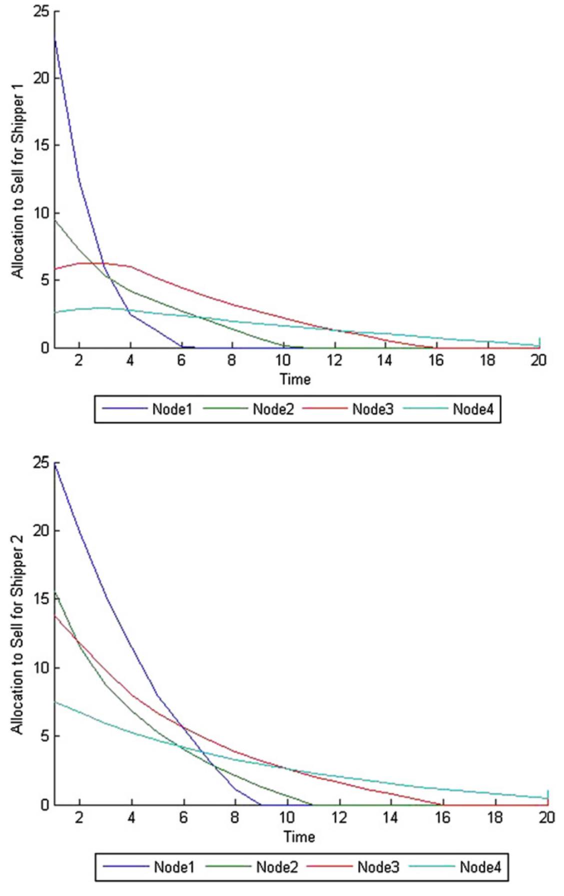

Figure 5.3: Allocation of Shippers' Commodity to Different Nodes.

In this numerical example, due to the higher commodity price at node 2 in comparison to node 1, as reflected in Figure 5.5, and low unit transportation cost movement between for OD pair (1; 2), shippers prefer to transport the commodity from node 1 to 2 rather than selling it at node 1. This flow circumstance may be observed in Figure 5.6.

5.6. Conclusion and Future Work

In this chapter we have provided a foundation for the game-theoretic model of Stackelberg-Cournot equilibrium and its application to the dynamic shipper–carrier freight assignment problem. We have shown that the Stackelberg game can be

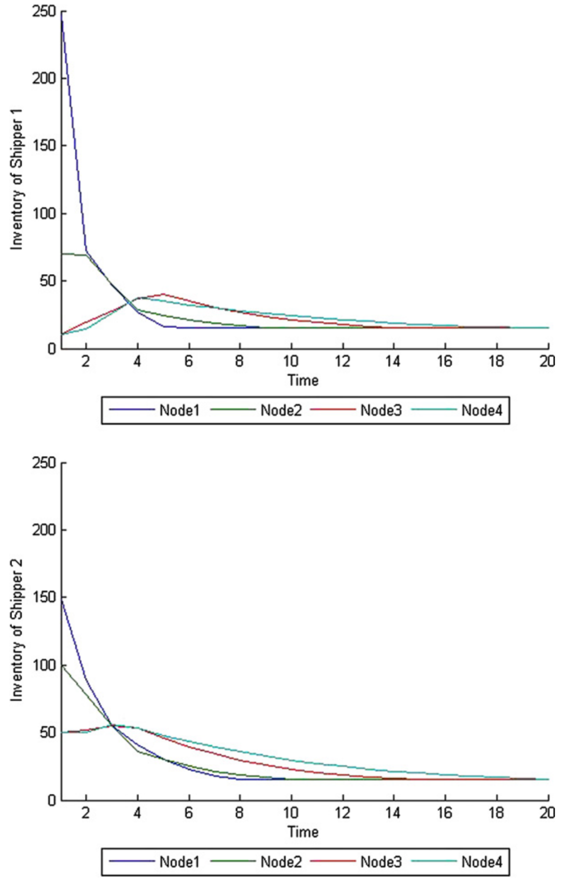

Figure 5.4: Inventory Trajectories of Shippers at Different Nodes.

formulated as a single level optimization problem by articulating the lower level problem as a differential variational inequality and reformulating it as a NCP. A discrete-time approximation method has been proposed to solve the single level optimal control problem. One advantage of our formulation is that its solution methodology takes advantage of well-documented and available optimization techniques. In particular, one may employ a standard nonlinear mathematical programming package to obtain a solution to the dynamic shipper–carrier freight assignment model.

We have presented an example that suggests this functional mathematical program may be a practical means of solving dynamic Stackelberg-Cournot games. The numerical example illustrates that the proposed model can successfully capture the dynamic pricing behavior and predict the freight flows on the network. To extend the presented model, one may assume some uncertainty in commodity

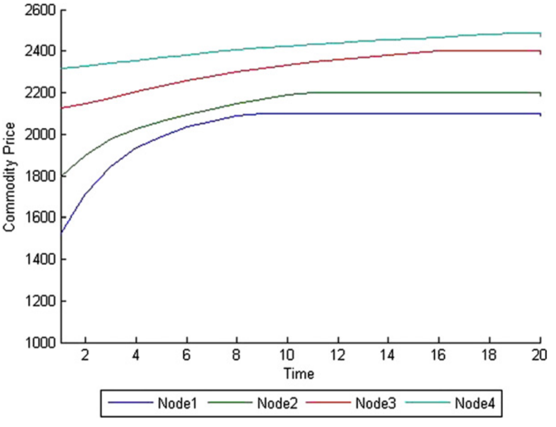

Figure 5.5: Market Price.

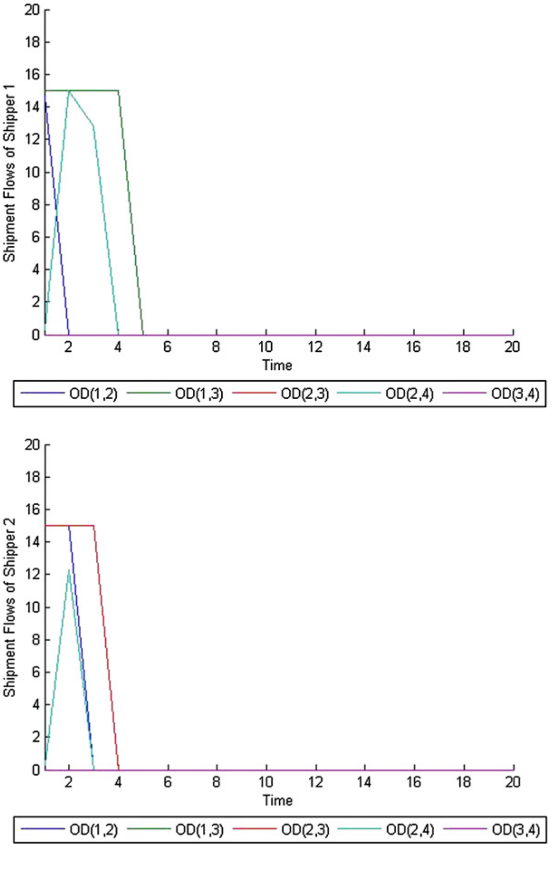

Figure 5.6: Shipment Trajectories of Commodity for Shippers.

price as well as in unit transportation cost and apply stochastic programming/robust optimization to handle the resulting problem.

Future research will focus on the extension of the model presented herein by introducing uncertainty on the players' observation. For example, it is likely that the leader does not have prefect information about the terminal condition of followers' state. In this case the MPEC problem becomes stochastic mathematical programs with equilibrium constraints (SMPEC).

Furthermore, it is possible to stochasticize the state dynamics using Brownian motion to represent uncertainty in the signals that determine the rate of change of state. The resulting stochastic dynamic Nash game will take the form of a stochastic differential variational inequality.

References

Agrawal, B. B., & Ziliaskopoulos, A. (2006). Shipper-Carrier dynamic freight assignment model using a variational inequality approach. *Journal of the Transportation Research Board*, *1966*, 60–70.

Baccarin, S. (2002). Optimal impulse control for cash management with quadratic holding-penalty costs. *Decisions in Economics and Finance*, *25*(1), 19–32.

Brander, J. A., & Zhang, A. (1993). Dynamic oligopoly behavior in the airline industry. *International Journal of Industrial Organization*, *11*(3), 407–435.

Chamberlin, E. H. (1993). *The theory of monopolistic competition*. Cambridge, MA: Harvard University Press.

Dafermos, S., & Nagurney, A. (1987). Oligopolistic and competitive behavior of spatially separated markets. *Regional Science and Urban Economics*, *17*(2), 245–254.

Fernandez, L., De Cea, J., & Soto, O. A. (2003). A multi-modal supply-demand equilibrium model for predicting intercity freight flows. *Transportation Research Part B*, *37*(7), 615–640.

Fisk, C. S. (1986). A conceptual framework for optimal transportation system planning with integrated supply and demand models. *Transportation Science*, *20*(1), 37–47.

Fisk, C. S. (1983). Game theory and transportation system modeling. *Transportation Research Part B*, *18*(5), 301–313.

Friedman, J. W. (1979). *Oligopoly and the theory of games*. New York, NY: North-Holland.

Friesz, T. L. (2010). *Dynamic optimization and differential games*. New York, NY: Springer.

Friesz, T. L., Rigdon, M. A., & Mookherjee, R. (2006). Differential variational inequalities and shipper dynamic oligopolistic network competition. *Transportation Research Part B*, *40*(6), 480–503.

Friesz, T. L., Gottfried, J. A., & Morlok, E. K. (1986). Sequential shipper-carrier network model for predicting freight flows. *Transportation Science*, *20*(2), 80–91.

Friesz, T. L., Tobin R. L., & Gottfried J. A. (1981). A new model and efficient computational procedure for determining freight network equilibrium. *Coal Transportation Modeling Workshop, Palo Alto*, December 3–4, 1981.

Greenhut, M., & Lane, W. J. (1989). Theory of oligopolistic: Competition. *The Manchester School*, *57*(3), 248–261.

Greenhut, M., Norman, G., & Hung, C. S. (1987). *The economics of imperfect competition: A spatial approach*. New York, NY: Cambridge University Press.

Harker, P. T. (1984). Variational inequality approach for the determination of oligopolistic market equilibrium. *Mathematical Programming*, *30*(1), 105–111.
Harker, P. T., & Friesz, T. L. (1986). Prediction of intercity freight flows I: Theory. *Transportation Research Part B*, *20*(2), 139–153.
Henderson, J. M., & Quandt, R. E. (1980). *Microeconomic theory: A mathematical approach* (3rd ed.). New York, NY: McGraw-Hill.
Intriligator, M. D. (1971). *Mathematical optimization and economic theory*. Englewood Cliffs, NJ: Prentice Hall.
Markovich, S. (2008). Snowball: A dynamic oligopoly model with indirect network effects. *Journal of Economic Dynamics and Control*, *32*(3), 909–938.
Miller, T., Tobin, R. T., & Friesz, T. L. (1992). Network facility-location models in Stackelberg-Nash-Cournot spatial competition. *The Journal of the RSIA*, *71*(3), 277–291.
Miller, T., Tobin, R. T., Friesz, T. L., & Kwon, C. (2007). Reaction function based dynamic location modeling in Stackelberg-Nash-Cournot competition. *Network and Spatial Economics*, *7*(1), 77–97.
Mutlu, F., & Cetinkaya, S. (2011). Coordination in retailer-carrier channels for long term planning. *International Journal of Production Economics*, *133*(1), 360–369.
Nagurney, A., Dupuis, P., & Zhang, D. (1994). A dynamical systems approach for network oligopolies and variational inequalities. *Annals of Operations Research*, *28*(3), 263–293.
Nash, J. (1951). Non-cooperative games. *The Annual of Mathematics*, *54*(2), 286–295.
Pang, J. S., & Stewart, D. E. (2008). Differential variational inequalities. *Mathematical Programming*, *113*(2), 345–424.
Sethi, S., & Thompson, G. L. (2000). *Optimal control theory* (2nd ed.). Boston, MA: Kluwer Academic Publishers.
Shah, N., & Brueckner, J. K. (2012). Price and frequency competition in freight transportation. *Transportation Research Part A*. Retrieved from http://dx.doi.org/10.1016/j.tra.2012.02.014
Sherali, H. D., Sosyter, A. L., & Murphy, F. H. (1983). Stackelberg-Nash-Cournot Equilibria: Characterizations and computations. *Operations Research*, *31*(2), 253–276.
Simaan, M. (1977). Equilibrium properties of the Nash and Stackelberg strategies. *Automatica*, *13*(6), 635–636.
Stackelberg, V. H. (1952). *The theory of market economy*. Oxford, UK: Oxford University Press.
Tobin, R. T. (1992). Uniqueness results and algorithm for Stackelberg-Cournot-Nash equilibrium. *Annals of Operations Research*, *34*(1), 21–36.
Unnikrishnan, A., Valsaraj, V., & Waller, S. T. (2009). Stochastic and dynamic shipper carrier network design problem. *Networks and Spatial Economics*, *9*(4), 525–550.
Wie, B. W., & Tobin, L. T. (1997). A dynamic spatial Cournot-Nash equilibrium model and an algorithm. *Computational Economics*, *10*(1), 15–45.
Xiao, F., & Yang, H. (2007). Three-player game-theoretic model over a freight transportation network. *Transportation Research Part C*, *15*(4), 209–217.

REGIONAL

Chapter 6

Discrete Choice Analysis of Shippers' Preferences

Moshe Ben-Akiva, Denis Bolduc and Jay Q. Park

Abstract

Freight choice models should focus on two objectives: quality of service and total logistics costs, both of which are important to shippers. This chapter examines discrete choice models that apply these objectives to enhance existing freight choice models. A total logistics cost model based on revealed preferences (RP) data is examined first. The model is then extended to include service quality. However, because service perceptions are unobservable, we use a factor analytic formulation to relate the service perceptions to perceptual indicators. These indicators are extracted and the extended model is then jointly estimated using RP data along with these indicators. Estimation efficiency of the extended model is further improved through the joint use of two data sets — one measuring RP, and one measuring stated preferences (SP). We find that the simultaneous use of both sets of data produces superior estimation results in the extended model. This demonstrates the significance of perceptions of service quality in explaining shippers' choices.

Keywords: Shippers' preferences; discrete choice models; service quality; total logistics costs; choice modelling

6.1. Introduction

The models presented in this chapter are designed to predict shippers' choice of mode more accurately. The choice model reflects competition among different modes by estimating choice probabilities as a function of relative utilities between available

Freight Transport Modelling

ISBN: 978-1-78190-285-1

modes. These utilities capture the differential effects of freight rate, logistics costs and service quality — all important components affecting mode choices.

The choice model is derived from the assumption that shippers seek to minimize total logistics costs. A basic version of our model is estimated using revealed preferences (RP) data collected in 1988 by a major U.S. railroad company. The survey included 166 transportation managers who each ship one of the following five commodities: paper, aluminum, pet food, plastics, and tires. This kind of data is expensive and usually proprietary. It was a unique opportunity to receive such a proprietary marketing survey from a railroad company for research.

After estimating a basic choice model using the RP data, we proceeded to several extensions. First, in order to account for heterogeneity across shippers, the discount term present in the total logistics cost specification is allowed to have a distribution. Then, a hybrid choice model formulation is used to introduce an unobserved variable, which measures perception of service quality, among the explanatory variables of the choice model. The hybrid choice model is a choice model that is augmented with a factor analytic submodel that relates the unobserved service quality variable to observed indicators. In the final section of the chapter, both the RP and the stated preferences (SP) data collected in the same survey are used to enhance the model estimation. The combination of RP and SP data resulted in more accurate parameter estimates.

6.2. Overall Approach

The starting point of this chapter is the assumption that logistics managers consider the total logistics costs when they select among freight modes.

6.2.1. Total Logistics Cost Model

Total logistics cost models assume that shippers minimize total costs, including transportation costs and inventory holding costs (see, e.g., Friedlaender, 1969; Meyer, Peck, Stenason, & Zwick, 1959). Several individual items interact in complex ways to determine total logistics costs. An attempt to minimize any single cost element can actually increase total cost. There are trade-offs among the different components of the total logistics cost. For instance, if transit time is long, shippers incur high inventory holding costs during the long in-transit time. If transit time is unpredictable, shippers have to hold high safety stocks which increase safety stock holding costs.

Logistics costs include components such as: order management, transport fee, transshipment charge, deterioration, loss and damage fee, cost of capital attached, inventory management, and penalty for unreliability. In order to model logistics choices we must address several complexities. Commodities are non homogenous with different requirements and characteristics. Shippers are heterogeneous. Firm characteristics such as annual revenue, truck availability, and railway siding are important. Shipment characteristics such as size, frequency, and goods typology (perishability, value) are also critical. Finally, the level of service attributes, (e.g., shipment time, cost, reliability) are key factors as well.

The framework for total logistics costs was used by Roberts, Brigham, and Miller (1976) and Roberts et al. (1976) to make normative recommendations of both modal selection and shipment size. Their recommendations were based on the calculation of the most cost-efficient transportation mode, and the assumption that total logistics costs can be precisely defined based on shipment characteristics (e.g., product price, demand projections, shipment distance, and predetermined discount rate). This deterministic approach is limited since it assumes that shippers only consider cost advantages and ignore nonprice attributes such as perceptions of service quality.

6.2.2. Discrete Choice Model

In many instances, normative recommendations provided by cost models do not coincide with actual choices. For example, it can be difficult to define and measure total logistics costs precisely. For such cases, a statistical approach that estimates a perceived discount rate and the importance of cost variables is required. The result is an approach that combines discrete choice methods with the minimization of total logistics costs in the same way that utility maximization is modeled for individuals' choice behavior (Ben-Akiva & Lerman, 1985). This approach was applied by Chiang (1979) to estimate the joint decision of modal selection and shipment size. Vieira (1992) estimated the value of service and market segmentation in modal selection. He compared the results of a choice model with the linear utility function of modal attributes and with a total logistics cost model. He found that the latter is superior in terms of data fit. This chapter extends the work of Chiang and Vieira.

6.3. The Basic Logistics Choice Model

6.3.1. Model Framework

The RP approach estimates demand response by analyzing actual choices in relation to actual situations. The conceptual framework of analyzing RP data is given in Figure 6.1.

Data on actual choices and choice situations may be acquired either through external monitoring or shipper surveys. While external observation is effective in collecting data on measurable service performance, a survey can reveal shippers' perceptions, a factor that may be as important as performance.

Since the RP data explain actual choice outcomes based on the attribute characteristics of alternatives in real choice situations, the attribute importance estimated in the RP approach is likely to predict future choices more accurately than

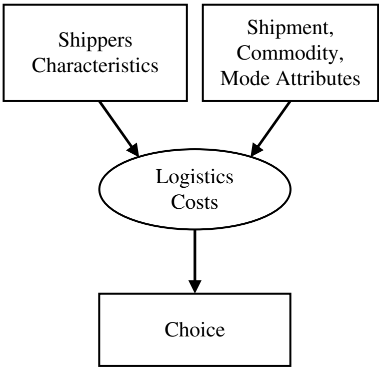

Figure 6.1: Basic logistics choice model.

hypothetical choices or stated importance. On the other hand, RP data are expensive and difficult to collect because researchers should observe or survey actual choices in addition to all alternative-specific attributes that might affect choice. In order to ensure a sufficient number of data points, researchers often collect data from many respondents or collect multiple observations of the same respondent. If some alternatives are rarely chosen, the sample size should be increased in order to collect adequate data on the seldom-chosen alternatives. As the sampling size increases, the collected data may contain measurement errors and suffer from nonresponses in important variables. Ideally, modal choices should be observed whenever a shipper sends a shipment. But such disaggregate observations are expensive and rarely available. Usually, only average statistics, such as modal shares among annual tonnage, are available. The shares represent the time-average statistics of a specific origin-destination pair for each shipper.

6.3.2. The Data

The empirical results presented in this chapter are based on a market research survey conducted by a major U.S. railroad company in 1988. As mentioned in the introduction, this kind of data is expensive and usually proprietary. The data set contains both RP and SP data. The survey included 166 transportation managers who each ship one of five commodities: paper, aluminum, pet food, plastics, and tires. Each shipper rated the performance of available transportation modes for their two largest corridors. All shippers ship at least $1 million worth of their respective commodity annually. Thus, they can be potential customers for both trucks and railroads. For detailed descriptions of the survey, refer to Vieira (1992).

We screened the data and eliminated 21 out of 166 respondents who did not report transit time or freight rate. Three alternatives were considered: truck, rail, and intermodal.

6.3.3. Notation

As mentioned above, to measure the movement of goods, only average statistics such as modal shares are available. Let i denote a mode and n represent a shipper. Our objective is to model the share s_{in} which denotes the percentage of the commodity shipped via mode i by shipper n among annual corridor shipments. We define X_n as a matrix with as many rows as there are alternatives where X_{in} corresponds to the ith row of X_n. $\boldsymbol{X}_{in}$ is a row vector that contains service attributes of mode i and characteristics of shippers n. We define $F(i|X_n;\ \beta)$ as a function that predicts the probability that shipper n selects mode i where $\boldsymbol{\beta}$ is a vector of unknown parameters. The number of alternatives is denoted as J.

6.3.4. Discrete Choice Model

Assume that a shipper selects the mode with the highest utility. We write the utility of mode i for shipper n as:

$$U_{in} = X_{in}\beta + \varepsilon_{in} \tag{6.1}$$

where ε_{in} denotes an error term that accounts for measurement errors. We assume that the explanatory variables enter the model through a linear relationship $X_{in}\beta$. The error terms, ε_{in}, $i = 1, \ldots, J$ are assumed to be independently and identically distributed (i.i.d.) standard normal random variables. We chose to apply a probit specification instead of a logit specification in the calculation of choice probabilities. Probit specifications are more flexible when it comes to implementing the model extensions that we develop later on in the chapter.

To derive the form of the $F(i|X_n;\ \beta)$ function, consider the following situation with three modes. The utilities are written as:

$$\begin{aligned} U_{1n} &= X_{1n}\beta + \varepsilon_{1n} \\ U_{2n} &= X_{2n}\beta + \varepsilon_{2n} \\ U_{3n} &= X_{3n}\beta + \varepsilon_{3n} \end{aligned}$$

The probability that mode 1 is selected by shipper n corresponds to:

$$\begin{aligned} F(1|X_n;\beta) &= \Pr(U_{1n} > U_{2n}, U_{1n} > U_{3n}) \\ &= \Pr(\varepsilon_{2n} < (X_{1n}\beta - X_{2n}\beta) + \varepsilon_{1n}, \varepsilon_{3n} < (X_{1n}\beta - X_{3n}\beta) + \varepsilon_{1n}) \end{aligned}$$

Given the assumption regarding the error terms, and through a conditioning onε_{1n}, we get:

$$\begin{aligned} F(1|X_n;\beta) = \textstyle\int_{-\infty}^{\infty} & \Pr(\varepsilon_{2n} < (X_{1n}\beta - X_{2n}\beta) + \varepsilon_{1n}|\varepsilon_{1n}) \bullet \\ & \Pr(\varepsilon_{3n} < (X_{1n}\beta - X_{3n}\beta) + \varepsilon_{1n}|\varepsilon_{1n}) f(\varepsilon_{1n}) d\varepsilon_{1n} \end{aligned} \tag{6.2}$$

The normality assumption, leads to the following form for the $F(i|X_n;\beta)$ function:

$$F(i|X_n;\beta) = \int_{-\infty}^{\infty} \left[\prod_{j \neq i} \Phi(X_{in}\beta - X_{jn}\beta + \varepsilon) \right] \phi(\varepsilon) d\varepsilon \tag{6.3}$$

where ε is a standard normal random term, ϕ denotes the standard normal density and Φ denotes the standard normal CDF. The estimation results presented in this chapter all result from the probit specification defined in Eq. (6.3).

Since we only have access to modal shares by the shippers, Park (1995) concluded that an estimator with desirable properties is obtained by maximizing the following function:

$$\text{Max} \;\; Q_N(s;\beta) = \sum_{n=1}^{N} \sum_{j=1}^{J} s_{jn} \ln F(j|X_n;\beta) \tag{6.4}$$

where the inner sum is taken over all alternatives and the outer sum is taken over the observations. In terms of statistical properties, this estimator, known as minimum discrimination information (MDI), produces estimates that are consistent and asymptotically normally distributed. In practice, Gensch and Soofi (1992) reported that MDI performs better than least squares estimation (i.e. Berkson's method) in terms of the information discrepancy and absolute error in holdout samples. The model specification that enters the choice probabilities $F(i|X_n;\beta)$ is developed in the following section.

6.3.5. Model Specification and Estimation

We write the utility of mode i for shipper n as:

$$\begin{aligned} U_{in} &= \mu(\text{logistics cost}_{in}) + \varepsilon_{in}, \\ \text{logistics cost}_{in} &= C_{in} + W_{in}\theta + r(T_{in} + Z_{in}\gamma) \end{aligned} \tag{6.5}$$

where μ is a negative scale parameter and the ε_{in} terms are i.i.d. standard normal. This leads to the choice probabilities in Eq. (6.3). The specification in Eq. (6.5) describes the total logistics cost of mode i for shipper n, where C_{in} corresponds to the transportation cost; $\boldsymbol{W}_{in}$ is a row vector that contains the mode-specific constant, mode-specific variables that capture the effects of ordering costs, on-time delivery and equipment availability; T_{in} is the daily value of in-transit stocks; Z_{in} includes discount rate related variables such as loss and damage costs and stock-out costs; $\boldsymbol{\theta}$ and γ are vectors of unknown coefficients; and r is an unknown parameter that represents the discount rate perceived by the shipper. Variables used in the model are described in Table 6.1.

Table 6.1: Variables used in the models

Dependent variable
Share of a transportation mode *i* by shipper *n* among annual corridor shipments (%)
Independent Variables
Generic:
Average price per ton ($ of 1988)
Average annual tonnage in the corridor (ton)
Latest acceptable delivery hours behind the schedule (hours)
Earliest acceptable delivery hours before the schedule (hours)
Mode specific:
Average freight rate per ton per mile (cents)
Average line-haul length between origin and destination (mile)
Average transit time (days)
Average shipment size in the corridor when a shipper *n* uses mode *i* (tons)
Percentage of shipments being lost or damaged (%)
Percentage of times that a carrier delivered shipments within acceptable time frame (%)
Equipment usability

Using the MDI estimation method, a basic choice model was estimated using the probit specification of Eq. (6.3) and the utility specification of Eq. (6.5). Estimation results are presented in Table 6.2. All coefficients have the correct signs. Rail is used as the basic mode. Monetary values are all expressed in terms of 1988 dollars.

All parameters are scaled by the coefficient of transportation costs, and, thus, need to be interpreted relative to transportation costs. The model specification explains the unit total logistics costs. The truck-specific constant of -0.138 suggests that, if all conditions are equal, a shipper will perceive smaller total logistics costs (by $13.8/ton) than those of rail, and will prefer truck.

The intermodal-specific constant of 1.635 suggests that, if all conditions are equal, a shipper will perceive higher total logistics costs (by $163.5/ton) than those for rail, and will shun intermodal service. This high penalty on intermodal usage may be due to the effects of omitted variables, such as perceptions of ease-of-use or riskiness of using unfamiliar service. The data set was collected in 1988 when intermodal service was still unfamiliar to most shippers and railroads did not market the service aggressively.

The ratio of value/ton in the truck utility can capture the effects of order costs. The intermodal-specific coefficient of this variable was insignificant. The coefficient estimate of -0.098 suggests that shippers will prefer truck when product value is high or when corridor tonnage is small. The truck-specific variable for shipment distance has an estimated coefficient of 0.372. This suggests that shippers will shun truck when the distance is long. Every 100 miles will increase total logistics costs of using truck by $37.2/ton, relative to rail and intermodal. Two variables that represent

Table 6.2: Estimation results of the basic model

Attributes	**Basic model**	
	Estimate	***t*-statistics**
Truck-specific constant	– 0.138	– 0.478
Intermodal-specific constant	1.635	2.924
Value/ton (truck specific)	– 0.098	– 2.131
Distance (truck specific)	0.372	1.945
Delivery time reliability	– 0.811	– 0.918
Equipment usability	– 4.346	– 4.032
Discount rate	0.456	0.782
In-transit stock holding cost	1	N.A.
Safety stock costs	0.169	1.948
Loss and damage	1.645	0.306
Scale (transportation cost)	– 1.036	– 4.831
Objective function at convergence	– 142.77	
ρ^2	0.552	
Adjusted ρ^2	0.52	

quantifiable service quality — delivery time reliability and equipment usability — are also employed in the model. They are treated as engineering variables in this specification, but may also be treated as perceptual indicators. This would allow for the application of the latent factor methodology (presented in a subsequent section of this chapter).

The estimated coefficients suggest that shippers are willing to pay \$81.1/ton for perfect on-time delivery and \$434.6/ton for perfect equipment availability. This model also estimates the discount rate that shippers use when they evaluate the costs of holding in-transit stocks, stock-out, and loss and damage. Since all coefficients of in-transit stock holding costs, stock-out costs, and loss and damage costs represent the implicit discount rate, it is difficult to judge the accuracy of the discount rate. The coefficient of in-transit stock suggests 45.6%, while that of stock-out costs suggests 7.7% (0.456×0.169), and that of damage costs suggests 75%. In our specification, the coefficient of in-transit stock is assumed to be the correct rate. The level of stock-out costs vary widely among shippers and its coefficient is not significant. Moreover, its coefficient may represent indirect effects (e.g., smoothing production schedule, volume purchasing to exploit special promotion, etc.) which cannot be included into total logistics costs. The coefficient of loss and damage costs should be interpreted as an indication that shippers may have to wait much longer than 100 days after the shipping date (e.g., 164.5 days) before they receive a full refund.

It is troublesome that our estimate (e.g., 45%) is much higher than a typical cost of capital (e.g., 15–20%). Three explanations are proposed. First, the discrepancy may arise from the shippers' attempts to control inventory costs by setting the internal discount rate much higher than external costs. If this is the case, most normative approaches that first calculate total logistics costs based on the surveyed discount rate and then recommend the mode based on minimum total logistics costs would not predict actual choices correctly. True rates need to be found by treating surveyed rates as indicators and by estimating a structural equation. Second, the discount rate may be overestimated because the heterogeneity of discount rates among shippers was not properly incorporated. For example, if our data included a few shippers who have very high rates, the overall mean rate can go up. Third, the discount rate may be overestimated because important variables were omitted. Shippers, for instance, may prefer trucking because of its highly flexible service. But without the proper incorporation of the effects of flexibility, trucking may appear preferred because of a high discount rate.

6.4. Extended Logistics Choice Model

The basic logistics choice model has major weaknesses. It omits sources of heterogeneity, suggesting the use of a *random coefficient* formulation. Shipper surveys often contain multiple observations from the same respondent. It suggests that an *agent effect* should be incorporated in the model formulation. New developments in discrete choice modeling now permit the inclusion of latent variables in the choice model. The omission of such variables produces biased and inconsistent estimates of the model parameters. Latent variables are unobserved concepts that can be elicited with the help of psychometric indicators. Examples of latent variables in the present context may be *awareness*, *sensitivity to quality* and *image*.

6.4.1. Extension 1: Heterogeneity

Suppose that we observe a population whose modal share is split 50-50 between truck and rail. At one extreme, this might be interpreted as an implication that each shipper in the homogeneous population has a 50% chance of using truck and a 50% chance of using rail. At the other extreme, this might imply that the population is composed of two groups: one group that always uses truck, and another group that always uses rail. A discrete choice model with fixed coefficients accepts the former explanation, assuming that the utility of an individual varies randomly from one choice to the next. Yet, one may find it difficult to accept the concept of identical shippers.

Two types of heterogeneity may occur: inter-shipper heterogeneity and intra-shipper heterogeneity. Intershipper heterogeneity occurs when shipper-specific characteristics, such as attitudinal variations, unobserved perceptions of service quality, choice set generation process and decision protocol, influence modal choices. Intrashipper (or intershipment) heterogeneity occurs when shipment-specific situational constraints, such as shipment type (regular vs. emergency) and customer type (large vs. small), influence modal choices.

6.4.1.1. Treatment of heterogeneity: Random coefficients Ignoring parameter heterogeneity can lead to inconsistent estimates. For instance, suppose that we ignore the heterogeneity in alternative-specific constants even though the population is composed of several market segments which have different preferences for trucking. The estimated coefficients of the common taste model would be incorrect. Two major approaches have been proposed to cope with heterogeneous parameters. One approach is to estimate fixed parameters for each shipper, for example, alternative-specific constants for each shipper or shipper-specific discount rates. This approach is referred to as the fixed effect model. The value in this approach is that the conditional density of a shipper-specific effect (given shipper characteristics) is different from its unconditional density (Hausman, Hall, & Griliches, 1984). Therefore, it can provide segment-specific predictions. However, this approach involves the incidental parameter problem since the number of shipper-specific parameters grows linearly with the number of shippers. Moreover, when there are only a few observations for each shipper, classical estimators will lead to inconsistent estimates of the fixed terms, and of the effects of other variables (Hsiao, 1986).

Assuming that the parameters vary across the population according to some probability distribution, estimating the distribution of parameters would be a more tractable approach than estimating a parameter for each shipper. This approach is called the random coefficients model (Gonul & Srinivasan, 1990; McFadden, 1989). For our study, we adopted the random coefficients model since it ensures parsimony, while providing information (e.g., a distributional form) about the random effect. The discount rate is assumed to be random and, therefore, underlying parameters of its distribution can be estimated.[1]

6.4.1.2. Model specification From Eq. (6.5), we have:

$$\begin{aligned} U_{in} &= \mu(C_{in} + W_{in}\theta + r_n(T_{in} + Z_{in}\gamma)) + \varepsilon_{in}, \\ &= \mu(c_{in} + r_n z_{in}) + \varepsilon_{in} \end{aligned} \tag{6.6}$$

where, $z_{in} = T_{in} + Z_{in}\gamma$ represents a generalized value of discount rate-related variables. Variables in Z_{in} (e.g., in-transit stock holding costs, safety stock costs and loss and damage costs) share the common coefficient (i.e., discount rate) that is

1. One problem with the random coefficient model is that the model assumes that the conditional density of the discount rate given shipper characteristics and its unconditional density are identical. This implies that the model can produce inconsistent estimates if self-selection bias occurs (Chamberlain, 1980). For example, shippers who set quick response as a company policy for unobserved reasons (e.g., perishable product) may make huge investments on establishing efficient trucking connections (e.g., EDI link with trucking companies, dedicated truck transit pass, or 24-hour docking station). We assume that such self-selection bias is negligible in our data, since shippers are rather homogeneous. Even if selection bias did occur, we model the mean and variance of discount rate to vary systematically according to shipper characteristics.

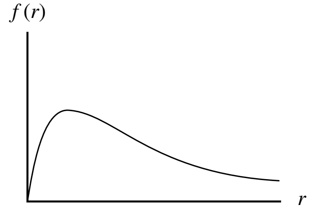

Figure 6.2: Distribution of discount rate.

assumed to be shipper specific. The variable $c_{in} = C_{in} + W_{in}\theta$ represents a generalized value of cost variables.

We represent the variation of the discount rate among shippers with a univariate distribution. We assume that the discount rate is log-normally distributed for the following reasons: (1) the log-normal distribution is defined only in the positive range, (2) the log-normal distribution has a left-skewed unimodal, and (3) a choice model with log-normally distributed discount rates nests the fixed rate model since the two models are the same when the variance of discount rate is zero.[2] If the discount rate is log-normally distributed across shippers, its natural logarithm is normally distributed. That is, if:

$$\ln r_n \sim N(\bar{r}, \sigma_r^2) \tag{6.7}$$

where $\bar{r}$ denotes the expected value of the natural log of the discount rate and σ_r^2 is its variance. With this assumption, the discount rate moments can be computed as:

$$\begin{aligned} \text{Mean} &= \exp(\bar{r} + \sigma_r^2/2) \\ \text{Mode} &= \exp(\bar{r} - \sigma_r^2) \\ \text{Median} &= \exp(\bar{r}) \\ \text{Variance} &= \exp(2\bar{r} + \sigma_r^2)(\exp(\sigma_r^2) - 1) \end{aligned}$$

This is an asymmetric distribution skewed to the left with a minimum value of 0 and tail to the right, as shown in Figure 6.2.

Under our distributional assumptions, the choice probabilities conditional on the discount rate can be computed as:

$$F(i|c_n, z_n, r_n) = \int_{-\infty}^{\infty} \left[\prod_{j \neq i} \Phi(\mu(c_{in} + r_n z_{in}) - \mu(c_{jn} + r_n z_{jn}) + \varepsilon) \right] \phi(\varepsilon) d\varepsilon \tag{6.8}$$

2. Another distribution we experimented with is the Gamma distribution. It does not nest the fixed rate model since when the variance of a gamma variate goes to 0, its mean also goes to 0.

where $\boldsymbol{c}_n$ is the $(J \times 1)$ vector containing all the c_{in}'s, while $\boldsymbol{z}_n$ is the $(J \times 1)$ vector containing all the z_{in}'s.

Unconditional probabilities are obtained by integrating over the distribution of the discount rate, thus leading to:

$$F(i|c_n, z_n) = \int_0^{\infty} [F(i|c_n, z_n, r)] f(r; r, \sigma_r^2) dr \tag{6.9}$$

Numerical integration is performed using a quadrature method.

The MDI approach of Eq. (6.4) was applied for this model, but unfortunately the data do not seem to carry enough information to capture any type of heterogeneity. Many attempts were made, but they were all unsuccessful.

6.4.2. Extension 2: Panel Effect

In the previous section, we did not differentiate inter-shipper and intrashipper heterogeneity. But, since we have multiple observations from the same shipper, the observations may be correlated with each other due to shipper-specific discount rates. The event in which a shipper-specific discount rate persistently influences all observations of the shipper is a specific type of panel effect. The MDI estimation that ignores the panel effect produces consistent estimates. But because it does not account for the fact that multiple observations are from the same shipper, the estimates are relatively inefficient.

In the RP data, we observed the modal shares of two corridors from each shipper, that is, $s_n = (s_{n1}, s_{n2})$. In order to model such multiple choices from a shipper, a subscript t is added to our notation. For example, consider the following utility function:

$$U_{int} = \mu(c_{int} + r_n z_{int}) + \varepsilon_{int} \tag{6.10}$$

where we assume that the discount rate is the same for all observations t from a given shipper, $r_{nt} = r_n \quad \forall t$. We also assume that the shipper specific discount rates are log-normally distributed with parameters $\bar{r}$ and σ_r^2 as follows:

$$\ln r_n \sim N(\bar{r}, \sigma_r^2) \tag{6.11}$$

In order to utilize the information of joint shares, we used the expected MDI estimation criterion which maximizes the expected information value of the joint shares. Given a discount rate r, the conditional information value of joint shares for a given shipper n is:

$$I_n(c_n, z_n, r_n) = \sum_{t=1}^{2} \sum_{i=1}^{J} s_{int} \ln(F(i|c_{nt}, z_{nt}, r_n)) \tag{6.12}$$

where $\boldsymbol{c}_n$ is the $(2J \times 1)$ vector containing all the c_{int}'s, while $\boldsymbol{z}_n$ is the $(2J \times 1)$ vector containing all the z_{int}'s and where:

$$F(i|c_{nt}, z_{nt}, r_n) = \int_{-\infty}^{\infty} \left[\prod_{j \neq i} \Phi(\mu(c_{int} + r_n z_{int}) - \mu(c_{jnt} + r_n z_{jnt}) + \varepsilon) \right] \phi(\varepsilon) d\varepsilon \qquad (6.13)$$

The unconditional expected information value, evaluated over all shippers, is written as:

$$I = \sum_{n=1}^{N} \int_{0}^{\infty} I_n(c_n, z_n, r) f(r; \bar{r}, \sigma_r^2) dr \qquad (6.14)$$

The integral is, again, evaluated numerically using a quadrature method.

6.4.2.1. Estimation results Table 6.3 shows the estimation results of the basic model and the models allowing for extensions 1 and 2. We focus on the changes attributable to the two extensions. The random coefficient modeling improved the basic model only by a negligible amount. Many parameters are fairly similar to the estimates obtained for the basic model.

The results suggest that the natural logarithm of the discount rate is normally distributed with a mean of -0.783 and a variance of 0.01. In terms of the discount rate, the results suggest that the mean rate is 45.9% and the standard deviation is

Table 6.3: Estimation results of the basic, random coefficient and panel effect models

Attributes	**Basic model Estimates**	**Random coefficient Estimates**	**Panel effect Estimates**
Truck-specific constant	–0.138	–0.081	–0.563
Intermodal-specific constant	1.635	1.437	1.264
Value/ton (truck specific)	–0.098	–0.368	–0.404
Distance (truck specific)	0.372	0.438	0.361
Delivery time reliability	–0.811	–0.848	–1.118
Equipment usability	–4.346	–4.723	–4.874
Discount rate-mean	0.456	–0.783	–0.715
Discount rate-standard deviation		0.103	0.481
In-transit stock holding cost	1	1	1
Safety stock costs	0.169	0.033	0.038
Loss and damage	1.645	1.537	0.574
Scale (transportation cost)	–1.036	–1.707	–1.318
Objective function at convergence	–142.77	–142.62	–122.42
ρ^2	0.552	0.552	0.616
Adjusted ρ^2	0.520	0.518	0.584

4.7%. The small variance suggests that the perceived discount rate is high and does not differ much among shippers. The small variance is rather disappointing since high heterogeneity was expected. The major change in parameter estimates occurred with the stock costs variable. It decreased, while the parameter estimate associated with value/ton increased from -0.98 to -0.368.

Allowing for panel effect improves the overall fit. The panel effect modeling gave estimated parameters that have expected signs and comparable magnitude to those of the random coefficient model. The estimated parameter associated with the delivery time reliability variable slightly decreased. In terms of discount rate, the results suggest that the mean rate is 54.9% (a value much higher than in the non panel case) and the standard deviation is 28%. Those values are still high. The inclusion of a latent variable in the model, which we consider next, will greatly improve the results. In the following, we will treat the discount rate as a fixed parameter.

6.4.3. Extension 3: Latent Variables

Shippers' choices are not solely based upon rational assessments of total logistics costs. We assume that shippers' perceptions of service quality influence their choice of mode and extend the choice model to include both total logistics costs and service quality perceptions. The difficulty of this extension is in the measurement of these perceptions. We adopt an approach based on structural equations with latent variables and their corresponding measurement equations.

Quality of service measurements can be extracted from a variety of quantitative attributes, such as transit time. For example, a short transit time not only reduces capital costs committed to in-transit stocks in terms of total logistics costs but also allows shippers to provide customer-responsive service. On-time delivery not only reduces inventory holding costs necessary for meeting safety stock requirements, but also eliminates a need for shippers to prepare for emergency shipments and deal with liability issues. Irregular shipments may cause congestion, confusion, or poor sequencing at shippers' receiving docks. While the exact cost of such irregular shipments cannot be calculated at every incidence, these incidences somehow will be recorded and reflected when shippers evaluate the performance of carriers, and eventually select a mode. In addition to reliability, many other perceptions may also influence modal selection, such as:

Flexibility: Traffic managers must routinely face contingencies. Demand may suddenly peak at unexpected places due to unexpected publicity or promotional campaigns. Production schedules are suddenly delayed due to material supply problems, labor disputes, or bad weather, and so forth. In such cases, shippers need to expedite or delay shipments in order to fulfill their sales commitments or to prevent production disruptions. Naturally, they prefer carriers that can readily provide shipment status, and are easy to work with in changing shipment sizes, contents, schedules, or routing. In addition, attributes such as carriers' responsiveness to inquiries, ease of shipment rescheduling or route change, ready provision of shipment tracing information, or good training of sales representatives will improve the perception of flexibility of a carrier.

Riskiness: Logistics managers strive to minimize the risk of logistics system failure. The perceived risk depends on the uncertainty about the outcome of a carrier selection decision and the magnitude of the consequences of a wrong choice. For instance, unreliable delivery not only increases inventory holding costs for safety stock requirements, but also increases shippers' liability for delayed shipments. While the exact cost of liability can not be calculated, shippers' perceptions about the liability will influence modal selection. Such perception can also be formed from the accuracy of billing statements, the history of loss and damage, the availability of usable equipment, or the reliability of on-time delivery.

Familiarity: When delivery is not made as scheduled, the size or reputation of a carrier can help ameliorate criticism. Logistics managers may be unable to attribute blame to the selection of a specific carrier or mode if the selected carrier or mode has a good reputation. With a lesser-known mode, managers may blame transportation-related problems on the selection of a mode. For instance, shippers may not use intermodal services because they are afraid of using untried services. If railroads develop programs to familiarize shippers with intermodal service, intermodal revenue may grow rapidly.

6.4.3.1. Modeling latent variables The modeling framework is extended to situations where an explanatory variable cannot be observed directly while the survey contains indicators of that variable. This variable, called a latent variable, is added to the choice model specification. The choice model is augmented with measurement equations that relate the latent variable to the indicators. A general presentation of choice models with latent explanatory variables may be found in Ben-Akiva et al. (2002) or Walker and Ben-Akiva (2002).

Although the degree of flexibility to which shippers perceive each transportation mode cannot be directly measured, the survey asks shippers how satisfied they are with the responsiveness of a carrier to their inquiries, how easy it is to work with a carrier, and how flexible a carrier is to schedule changes or shipment size changes. Responses to such questions indicate a level of flexibility to which shippers perceive certain alternatives. Adding latent variables to the basic logistics model framework leads to Figure 6.3.

Latent variables are depicted with oval shapes. In Figure 6.3, latent variables are assumed to be explained by shippers' characteristics and shipment, as well as commodity and mode attributes. Arrows on the right end side indicate a measurement process where respondents reveal their perceptions through indicators. Errors terms affect each component in the figure.

The framework relating the latent variable *flexibility* to explanatory variables and indicators that we consider for estimation is displayed in Figure 6.4. In our framework, flexibility across modes is allowed to be correlated. The model estimation can be performed using a sequential approach or a full information approach. First, the sequential approach involves the estimation of a factor analysis model that relates flexibility to indicators. Then, after extracting the factor, that is, obtaining

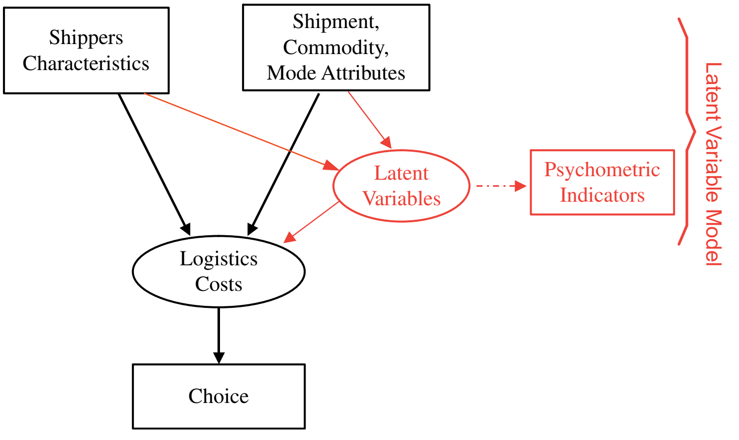

Figure 6.3: Logistics choice model with latent variables.

a fitted flexibility measure, we replace the latent flexibility variable in the choice model with the extracted factor. This approach can be interpreted as an instrumental variable (IV) estimation where the extracted factor works as an instrumental variable in the discrete choice model. The sequential approach leads to consistent but inefficient estimators of the choice model. The variance covariance matrix of the estimators needs to be adjusted to account for the substitution made in the choice model. The full information solution was not implemented because it was found to be too computationally demanding.

The application of the sequential approach to our data results in the estimates summarized in Table 6.4. The estimated model ignores the panel effect and assumes a fixed discount rate. The results show that adding flexibility to the basic RP model formulation only slightly improved the model fit. However, several coefficients show different estimates from those of the basic RP model without flexibility. In particular, the discount rate goes down to 35.6%, which is lower than 45.6%, the previous estimate that did not account for flexibility. The coefficient associated with the value/ton variable is similar to what the random coefficient and the panel effect model produced.

Below, we consider another extension where RP and SP data are combined, and the results show the flexibility variable to be significant.

6.4.4. Extension 4: Combining RP and SP Data

SP data are collected by asking shippers to make choices in hypothetical situations. SP data are different from stated importance data in that it does not ask for the importance of attributes directly. Rather, it tries to infer the importance of attributes by relating preferences to prespecified values of attributes.

Since SP data are usually composed of multiple observations from each respondent, the persistency (panel effect) of respondent-specific heterogeneity may occur more significantly in SP data than in RP data. SP data may contain response biases.

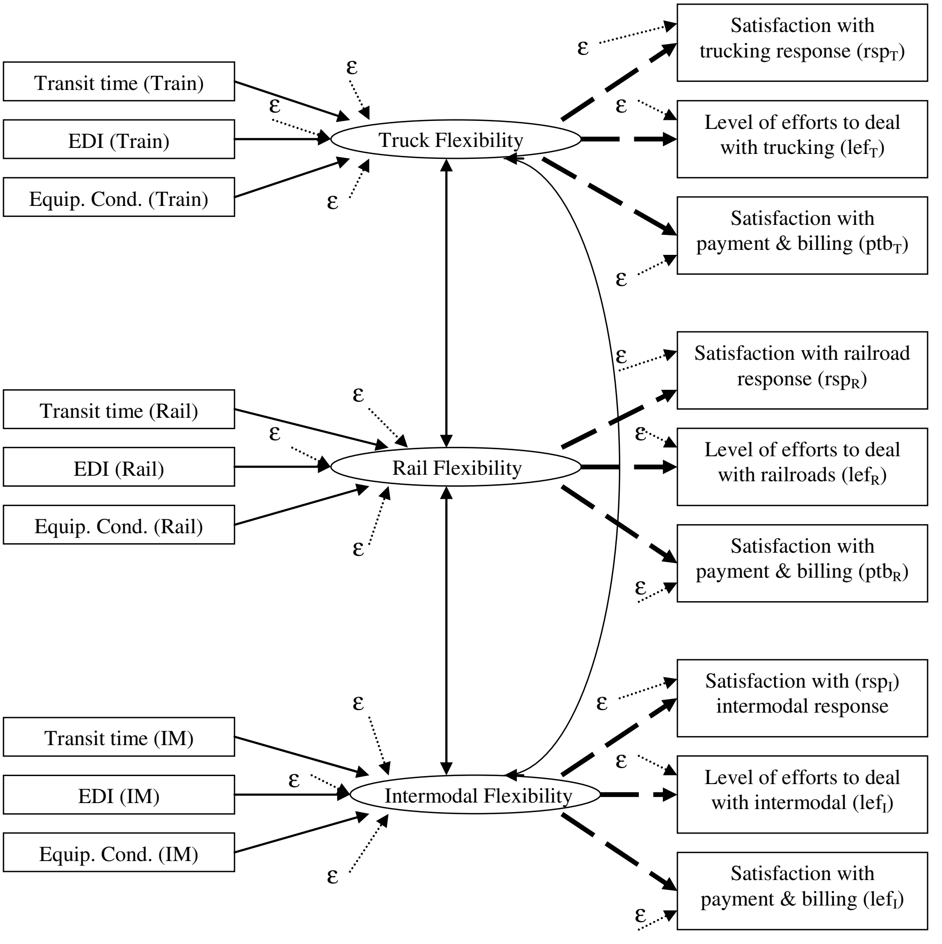

Figure 6.4: Specification of the latent variable model of modal flexibility.

A methodology to combine SP and RP data and correct potential biases within the model specification, with scale adjustments was developed by Ben-Akiva and Morikawa (1990).

Since both data are collected from the same shippers, they will reflect the same preference function. Therefore, we estimate parameters by maximizing the likelihood of observing RP and SP data jointly. By combining both data, we are able to correct response variability inherent in the SP data, while ensuring external validity with the RP data. In this respect, we will improve external validity and increase statistical efficiency. We combine both RP and SP data sets without allowing for heterogeneity of the discount rate. The discount rate was effectively found to be nondistributed in the separate RP and SP estimations.

Several measures were taken to reduce the well known problem of response bias associated with using SP data sets. First, earlier results (Ben-Akiva & Morikawa, 1990) show that alternative-specific constants estimated from RP data are different

Table 6.4: Estimation results of a model with latent variables

Attributes	Basic model		Latent variable model	
	Estimate	***t*-statistics**	**Estimate**	***t*-statistics**
Truck-specific constant	−0.138	−0.478	−0.266	−0.844
Intermodal-specific constant	1.635	2.924	1.648	2.944
Value/ton (truck specific)	−0.098	−2.131	−0.466	−1.768
Distance (truck specific)	0.372	1.945	0.627	1.882
Delivery time reliability	−0.811	−0.918	−0.905	−1.002
Equipment usability	−4.346	−4.032	−4.665	−3.285
Flexibility			−0.769	−1.154
Discount rate	0.456	0.782	0.356	1.921
In-transit stock holding cost	1	N.A.	1	N.A.
Safety stock costs	0.169	1.948	0.097	2.854
Loss and damage	1.645	0.306	2.277	0.849
Scale (transportation cost)	−1.036	−4.831	−1.018	−2.679
Objective function at convergence	−142.77		−137.58	
ρ^2	0.552		0.568	
Adjusted ρ^2	0.52		0.534	

from those estimated from SP data. With this in mind, we estimated mode-specific constants separately for RP and SP data. Secondly, SP responses often have a bias which is that frequently-used modes in actual shipments often receive high preference over infrequently-used modes. In order to capture the degree of this response bias, we included modal shares in RP data as an explanatory variable of SP data, so that SP utility increases smoothly with modal shares.

The estimator used for the combined RP/SP model is a convex combination of the MDI and maximum likelihood estimation (MLE) criteria. Since SP observations correspond to choices, we use MLE. For the RP data, we adopted the MDI criterion because we observed aggregate shares. Both estimation methods produce consistent and asymptotically normal estimators. A convex combination of consistent parameters is also consistent. Finally, a convex combination of asymptotically normal parameters is also asymptotically normal.

6.4.5. *Model Specification*

The utility of shipper n from alternative i at the tth observation is written as follows:

$$\begin{aligned} U_{int}^{RP} &= \mu(c_{int}^{RP} + rz_{int}^{RP} + \pi_{int}^{RP}\alpha^{RP}) + \varepsilon_{int}^{RP}, \\ U_{int}^{SP} &= \mu(c_{int}^{SP} + rz_{int}^{SP} + \pi_{int}^{SP}\alpha^{SP}) + \varepsilon_{int}^{SP} \end{aligned} \tag{6.15}$$

Table 6.5: Estimation results of combined RP and SP model

Attributes	**Latent variable model**		**Combined RP/SP**	
	Estimate	***t*-statistics**	**Estimate**	***t*-statistics**
Truck-specific constant (RP)	– 0.266	– 0.844	– 0.256	– 0.776
Intermodal-specific constant (RP)	1.648	2.944	1.669	4.588
Truck-specific constant (SP)			– 0.201	– 2.557
Intermodal-specific constant (SP)			0.088	1.38
Value/ton (truck specific)	– 0.466	– 1.768	– 0.466	– 1.865
Distance (truck specific)	0.627	1.882	0.602	2.464
Delivery time reliability	– 0.905	– 1.002	– 3.615	– 7.803
Equipment usability	– 4.665	– 3.285	– 1.398	– 7.795
Flexibility	– 0.769	– 1.154	– 0.82	– 7.43
RP modal share in SP (inertia bias)			– 0.118	– 1.735
Discount rate	0.356	1.921	0.202	6.601
In-transit stock holding cost	1	N.A.	1	N.A.
Safety stock costs	0.097	2.854	0.029	0.116
Loss and damage	2.277	0.849	5.049	6.431
Scale (transportation cost)	– 1.018	– 2.679	– 1.029	– 4.592
Objective function at convergence	– 137.58		– 141.32 (RP)	
ρ^2	0.568		0.556	
Adjusted ρ^2	0.534		n.a.	

In this model formulation, the c_{int} and z_{int} generalized variables are assumed to be present in both the RP and the SP model specification. They are also assumed to share common vector of coefficients. This is to ensure that both data reflect the same preferences. Then, π_{int}^{RP} contains variables that are only present in the RP database while π_{int}^{SP} corresponds to variables only present in the SP database. Finally, we assume that the error terms of RP and SP data are independent of each other. Estimation results using this approach are provided in Table 6.5.

The joint RP and SP model estimation produced a significant coefficient of the flexibility latent variable. Table 6.5 shows that using the RP database alone, flexibility was not significant. Combining RP and SP have clearly helped to improve the model. The test of stability of coefficients (not shown here) ensures that the results obtained can be trusted. The value of the discount rate in the RP and SP setting is reasonable and strongly significant, as opposed to the situation obtained in the RP alone.

6.5. Conclusion

Freight choice models were estimated showing the influence of both total logistics costs and service quality on modal selection. Quality perceptions are difficult to analyze since they are unobservable. The incorporation of perceptions into the choice model improved the overall fit and produced more reasonable parameter estimates.

Finally, a model combining RP and SP datasets to increase the estimation efficiency and the external validity of research results was estimated. Our results showed that the utilization of multiple data sources can be of great help in getting accurate parameter estimates.

The major finding is that the joint RP and SP estimation allowed for a significant service quality latent variable in the choice model. Flexibility was insignificant when using the RP database alone. The gain from combining RP and SP was demonstrated. A test of stability of coefficients ensured that the results obtained were reliable and unbiased. The value of the discount rate in the RP and SP model is reasonable and significant, as opposed to the results obtained with RP data alone.

References

Ben-Akiva, M., & Lerman, S. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA: The MIT Press.

Ben-Akiva, M., & Morikawa, T. (1990). Estimation of travel demand models from multiple data sources. In M. Koshi (Ed.), *Transportation and traffic theory*. New York, NY: Elsevier.

Ben-Akiva, M., Walker, J., Bernardino, A., Gopinath, D., Morikawa, T., & Polydoropoulou, A. (2002). Integration of choice and latent variable models. In H. Mahmassani (Ed.), *Perpetual motion: Travel behaviour research opportunities and application challenges* (pp. 431–470). Oxford, UK: Elsevier.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies*, *47*, 225–238.

Chiang, Y. (1979). *Policy-sensitive model of freight demand*, Ph.D. thesis, Civil Engineering, Center for Transportation Studies, MIT, Cambridge, MA.

Friedlaender, A. (1969). *The dilemma of freight transport regulation*. Washington, DC: The Brookings Institution Press.

Gensch, D. H., & Soofi, E. S. (1992). A minimum discrimination information estimation of multiattribute market share models. *Marketing Science*, *11*(1), 54–63.

Gonul, F., & Srinivasan, K. (1990). Modeling multiple sources of heterogeneity in multinomial logit models: Methodological and managerial issues. *Marketing Science*, *12*(3), 213–229.

Hausman, J., Hall, B. H., & Griliches, Z. (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, *52*(4), 909–937.

Hsiao, C. (1986). *Analysis of panel data*. Cambridge, UK: Cambridge University Press.

McFadden, D. (1989). A method of simulated moments for estimation for discrete response models without numerical integration. *Econometrica*, *57*(5), 995–1026.

Meyer, J., Peck, M. J., Stenason, J., & Zwick, C. (1959). *Economics of competition in the transportation industry*. Cambridge, MA: Harvard University Press.

Park, J. K. (1995). *Railroad marketing support system based on the freight choice model*. Ph.D. thesis, Department of Civil and Environmental Engineering, MIT, Cambridge, MA.

Roberts, P. O., Brigham, T. B., & Miller, C. A. (1977). *An equilibrium analysis of selected intercity freight markets: Trucks with double trailers vs. TOFC shuttle trains as energy conservation alternatives*, CTS Report, No. 77–25, MIT.

Roberts, P. O., Terziev, M. N., Kneafsey, J. T., Wilson, L. B., Samuelson, R. D., Chiang, Y. S., & Deephouse, C. V. (1976). *Analysis of the incremental cost and trade-offs between energy efficiency and physical distribution effectiveness in intercity freight markets*. CTS Report, No. 76–14, MIT, Cambridge, MA.

Vieira, L. (1992). *The value of service in freight industry*. Ph.D. thesis, Department of Civil Engineering, MIT, Cambridge, MA.

Walker, J., & Ben-Akiva, M. (2002). Generalized random utility model. *Mathematical Social Sciences*, *43*(3), 303–343.

Chapter 7

A Comparison of Conjoint, Multi-Criteria, Conditional Logit and Neural Network Analyses for Rank-Ordered Preference Data

Michel Beuthe, Christophe Bouffioux, Cathérine Krier and Michel Mouchart

Abstract

The chapter presents a systematic comparison of four different methodologies for separately analysing individuals' stated preferences relative to the choice of alternative solutions of freight transport. These are defined by five 'qualitative' attributes: frequency of service, transport time, reliability, carrier's flexibility and transport losses, plus the monetary cost. As the data consist of alternatives' rankings, the models applied here are somewhat unusual in the field, at least in transportation analysis: conjoint analysis, UTA-type multi-criteria analysis, rank-ordered conditional logit, and neural network analysis. In an earlier paper Beuthe and Bouffioux (2008) applied the ordered logit model to a preference survey of 103 firms at an aggregate level. With linear utility functions, this model's results indicated a wide heterogeneity in decision making. In order to go deeper into the process of decision making, the present chapter applies the above four models, some of them with non-linear utility functions, to the individual rankings of nine firms' transport managers selected from the survey in diverse industrial sectors. The alternatives submitted to their judgement are designed according to an orthogonal fractional factorial design. Each estimation methodology is adjusted and specified to suit the specific data. Over this small set of individual firms, each applied method shows that the cost factor is the most important one in the individual choice making of seven out of nine transport managers. Reliability is often more important than transport

Freight Transport Modelling

ISBN: 978-1-78190-285-1

time, even though its relative importance varies from case to case. The other factors may play a significant role in some cases. These outcomes are mostly coherent with the earlier results obtained at an aggregate level. The two better performing methods are the multi-criteria and the neural network analyses, which both involve non-linear partial utility functions.

Keywords: Freight transport; qualitative attributes; rank ordered preferences; logit methodology; UTA multi-criteria; neural network

7.1. Introduction

Most of the econometric literature on freight transport's 'qualitative attributes', like transport time and reliability, is based on surveys' data that are analysed at an aggregate level, and the models generally assume a linear utility function. Such approaches can be justified by the need to obtain average results, which can be used within larger models of transport analysis, and by difficulties presented by heterogeneous logistic situations. Beuthe and Bouffioux (2008) followed that aggregate approach and applied a rank-ordered logit model on a survey's stated preferences data of 103 individual managers' rankings of transport alternatives. These were defined in terms of their cost and five qualitative attributes: Time, Reliability, Flexibility, Frequency and Loss. Beyond a simple analysis at an aggregate level, that paper was also able to identify to some extent the specific behaviours of transport managers in some subgroups of firms. The latter results indicated a wide heterogeneity in decision making with different weights given to attributes. In contrast, this chapter separately analyses the individual preferences of nine of the transport managers in diverse industries and transport circumstances with different types of models, some of them quite unusual in this line of research and assuming a non-linear utility function. It is devised as a testing of some quantitative models of decision making, also as a means of obtaining additional insights into the managers' decision process and examining whether the often used hypothesis of a linear utility function is appropriate.

Here, four different modelling approaches are compared: conjoint analysis sometimes used in marketing contexts for deriving utilities associated with goods or services' attributes, multi-criteria analysis that is usually applied for assessing projects, a rank-ordered logit model in line with the classic discrete utility models, and a more innovative neural network modelling of preferences. Each model allows the computation of a manager's individual additive utility function, as well as the partial utilities associated with each attribute. The presentation of these models is made here as simple as possible; references for more technical details are given within the chapter.

The 'utility function' terminology is maintained throughout the chapter for convenience rather than for substance. Actually, a more neutral terminology like 'decision function' or 'value function' could be more appropriate. Indeed, in the present case of freight transport, the managers are more likely to minimize some measure of the total transport logistic cost, the generalized cost, which integrates many internal and external logistic factors. These factors are functions of the transport attributes, and may naturally include some subjective judgement as to risk taking, since it bears upon the management of safety stocks.

Additive preference functions imply an assumption of preferential independence, meaning that, if two projects are characterized by the same values for some criteria, the preferences between them depend only on the values taken by the other criteria. Whether this hypothesis is acceptable in practical applications may vary from case to case. It is used as a good approximation in this case where some of the partial utility functions are non-linear and implicitly account for a certain degree of interdependence among attributes.

Beyond this introduction, Section 7.2 gives the main features of the stated preference experiment. Section 7.3 briefly reviews the conjoint analysis models. Section 7.4 explains the UTA-type multi-criteria models. Section 7.5 sets up the rank-ordered logit methodology. Section 7.6 presents the neural network approach, and Section 7.7 presents and compares the results obtained from the various estimated models. Some final comments conclude the chapter in Section 7.8.

7.2. The Stated Preference Experiment

For this experimental analysis, nine Belgian firms from different industries were picked from the larger survey focused on qualitative attributes in transport. They were selected on the basis of the contrasting characteristics of their typical transports as shown in Table 7.1. Obviously, it is not a representative sample of any group of firms. The attention is here centred on individual firms and the diversity of their decision making within different industrial circumstances. The information provided by this somewhat arbitrary small sample should suffice, in a first step at least, to raise some useful questions about better ways to formulate models at an aggregate level.

Table 7.1: Characteristics of firms

	Sector	Mode	Distance (km)	Travel time	Cost (€/Tonne-km)	Shipment Size (Tonne)
A	Steel	multim.	991	240 h	0.038	350
B	Steel	waterw.	404	55 h	0.017	900
C	Textile	multim.	2104	120 h	0.11	15
D	Electron.	road	800	48 h	0.12	23
E	Chemical	rail	1200	48 h	0.002	23
F	Cement	road	123	3 h	0.25	31.5
G	Packing	road	500	10 h	0.16	12
H	Pharma	road	240	24 h	0.96	0.1
I	Materials	waterw.	155	48 h	0.025	1000

In this survey, the face-to-face interviews with each transport manager started with general questions about the firm and its transport organization. Then, the manager had to describe a typical shipment by the firm in terms of origin, destination, mode, distance and transport time, out of pocket monetary cost, the flow's annual tonnage, the shipment size, frequency, the type of consignee and so on. This typical shipment was taken as reference in the preference experiment, which confronted the manager to a set of alternative transport solutions among which he/she had to express preferences. These alternatives were defined by six attributes: the monetary transport Cost, plus five 'qualitative' attributes which are the Frequency of services per week, door-to-door transport Time including loading and unloading, Reliability as percentage of deliveries at the scheduled time, Flexibility of the carrier measured in percentage response to non-programmed shipments, Loss as percentage of commercial value lost from damages, stealing and accidents. Some of these attributes are defined in percentage of occurrence in order to include the idea of probability or risk.

The managers were asked to simply rank the transport alternatives, including the reference shipment, according to their order of preference. Each alternative was presented on a separate card and defined by the six attributes' percentage variation from the typical shipment values. These percentages could take five values, the level '0' corresponding to the current transport solution, and the four other levels being variations with respect to the current situation: $+10\%$, $+20\%$, -10% and -20%. Twenty-five alternatives were submitted to the interviewee, and the corresponding (25×6) matrix was set-up according to a fixed fractional factorial design proposed by Addelman (1962) which has proportional frequencies of the factors levels. This design guarantees that the estimated main effects of attributes are uncorrelated, but obviously ignores the possible interactions between attributes. During the interviews this shorter design turned out rather convenient as it did not impose a burdensome choice task to the interviewees who could freely rank-order and reorder the cards of alternatives in the process of setting up their preferences.

It is worth noting that none of the alternatives is explicitly characterized by a specific mode. That feature was chosen because transport managers often do not have precise information about transport alternatives that they do not use. Providing that information in each case of a large survey would have demanded a difficult research effort. Secondly, the attributes' values themselves are functions of the mode that is used; and, thirdly, this 'abstract mode' approach usefully provides the values that are needed for computing full generalized transport cost, the main criterion for choosing a transport solution that is used in many freight traffic models. In any case, that problem is somewhat circumvented by our knowledge of the mode used by each firm.

7.3. Conjoint Analysis

Conjoint analysis (CA) is based on the premise that individuals value a product or service by combining the values of their attributes or characteristics that are the assumed sources of a consumer's utility. It is an analysis-of-variance model of the attributes' main effects on utility, which can be shown to be equivalent to a regression analysis of an index of preference on the measured levels of attributes. This approach is sometimes used in marketing surveys to analyse observed or stated consumers' preferences towards goods or services with different characteristics (Green, Carroll, & Goldberg, 1981). The dependent variable is usually a preference measure on an ordinal scale, whereas the independent variables are often nominal and sometimes interval-scaled variables. Then, CA decomposes the preferences into components with respect to the attributes of the alternative goods or services.

There are two types of CA: metric and non-metric. Metric CA is applied when the respondents' ratings, expressing strengths of preference for a service, are directly used as dependent variable. A non-metric CA allows for a monotonic transformation of the ratings that best fits the data. The latter approach may be of interest when only rankings are available, in place of ratings. Indeed, a ranking does not normally correspond to a scaled utility since the difference of utility between ranks, say, 25 and 24 can be different from that between 24 and 23. The only information given by a ranking is that alternative 24 is preferred to 25, as 1 is preferred to 2 and so on. Non-metric CA may then be useful for transforming rankings into ratings expressing the possible different strengths of preference between services.

For the reader's convenience, the presentation of the successive models will be given in terms of the survey's data. As indicated in the introduction, six attributes are considered; they take five different discrete values defined in percentage variations from the status-quo situation with zero value. Hence, each of the 25 transport alternatives ($i = 1 \ldots 25$) is characterized by a vector of attributes $\boldsymbol{z_i} = (z_{i1}, z_{i2}, \ldots, z_{ij}, \ldots, z_{i6})$, and the 25 vectors make up all together a (25×6) matrix $\boldsymbol{Z}$. The stated ranking of the 25 alternatives, from 1, 2, 3 … to 25, is represented by a vector $\boldsymbol{R} = (R_1, \ldots R_{25})$, each element of which is associated with a specific transport alternative.

Conjoint analysis is presented here in its linear regression version. It is based on a postulated ranking relationship $r(\boldsymbol{Z}) = (r_1(\boldsymbol{Z}), \ldots, r_{25}(\boldsymbol{Z}))$, which is a linear function of the matrix $\boldsymbol{Z}$. It can be taken as resulting from the alternatives' assessment by an additive utility function:

$$U_i(z_i) = \sum_{j=1}^{6} u_j(z_{ij}) + \varepsilon_i \qquad (7.1)$$

It provides an aggregation of the criteria in a common index for comparing and valuing the alternatives under consideration, and ranks the projects in a complete weak order.

Eq. (7.2) shows how the data are handled in a metric conjoint analysis framework. A particular R_i is represented by R_{klmnpq}, its six indices successively indicating the level taken by the six attributes. Each specific β coefficient is associated with one of the five values that can be taken by an attribute. Taking the relationship (7.2) as a representation of the utility function, each of the products $\beta_{ij} \cdot z_{ij}$ can then be interpreted as the partial utility value, or *part-worth* utility, contributed by an attribute's level. The error terms ε_{klmnpq} are normally and independently distributed, as usual. The partial utility levels of each attribute are normalized to a zero sum.

$$R_{klmnpq} = \mu + \beta_{1k}z_{1k} + \beta_{2l}z_{2l} + \beta_{3m}z_{3m} + \beta_{4n}z_{4n} + \beta_{5p}z_{5p} + \beta_{6q}z_{6q} + \varepsilon_{klmnpq} \qquad (7.2)$$

This decomposition of utility naturally leads to a large number of independent variables, one for each level of each attribute. Given the parameters normalization, there are (24+1) parameters to estimate with only 25 observations. It does not leave any degree of freedom, nor does it provide a clear functional relationship between the ranking and the independent variables.

'Non-metric' CA applies a monotonic transformation on the dependent variable. This approach would be worth considering if it did not require additional parameters. A variety of non-linear transformations can also be applied on the independent variables, but that would only aggravate the problem. It is possible though to reduce its complexity by taking the independent variables as if they were continuous and applying a monotone transformation on the dependent variable. This drastically reduces the number of coefficients to 7, but adds 24 parameters for transforming the dependent variable through an iterative process. The results of such a simplified model are presented in the empirical Section 7.6. CA's for this chapter were computed with the SAS/STAT procedure PROC TRANSREG (Kuhfeld, 1990).

7.4. Multi-Criteria Analysis

Like many multi-criteria approaches, the UTA model, initially proposed by Jacquet-Lagrèze and Siskos (1982), aims at estimating a utility function and allows the computing of each attribute's weight for an individual decision maker. It is specifically designed to handle preference rankings, and is able to estimate non-linear partial utility functions as continuous linear spline functions made of linked linear segments. Thus, it is particularly appropriate to analyse the data of the questionnaire's experimental design. Moreover, it constitutes a real model of decision making, in contrast with the above CA model, which is more like a descriptive statistical tool.

The model assumes the existence of an additive utility function which satisfies the classic axioms of decision theory, namely the axioms of comparability, reflexivity, transitivity of choices, continuity and strict dominance. Using the same notation as above, the expected utility function for the *i*th alternative, which is defined by a set z_i of six attributes values z_{ij}, is

$$U_i(z_i) = \sum_{j=1}^{6} u_j(z_{ij}) \qquad (7.3)$$

with

$$u_j(z_{ij}) \geq 0, \frac{du_j}{dz_{ij}} > 0 \tag{7.4}$$

The two conditions in (7.4) guarantee that the positive partial utilities are monotone increasing with respect to an increasing level of a favourable attribute. It can be seen as a minimal requirement of judgement rationality. Hence, each attribute must be defined as a favourable factor. In our case, this means that the cost and time attributes must be sign-inverted to fit in the model. It is worth noting that no condition similar to (7.4) is imposed on the estimated conjoint models.

The UTA method estimates the utility on a set of ranked alternatives by the method of linear goal programming proposed by Charnes and Cooper (1977), which provides an approximation by linear intervals of a non-linear function, actually a continuous linear spline function.

In order to apply that method, each criterion's variation interval $[z_j^{\min}, z_j^{\max}]$, defined by its least favourable and best values, is divided in equal intervals (four in this case). The variables to be estimated by the linear programme are the partial utilities $u_j(z_{ij})$ at the five bounds. The utility at intermediate values are then given by linear interpolation. The ranking stated by the transport manager imposes a set of 24 linear constraints on the partial utilities to be estimated. Applying the so-called UTASTAR specification proposed by Despotis, Yannacopoulos, and Zopounidis (1990), we have either a non-negative constraint like (7.5) if alternative *i* is preferred to alternative *k*, noted as *i* ***P*** *k,* or an equality constraint like (7.5) if they are indifferent, noted as *i* ***I*** *k*:

$$\sum_{j=1}^{6}\{u_j(z_{ij}) - u_j(z_{kj})\} + \sigma^+(i) - \sigma^-(i) - \sigma^+(k) + \sigma^-(k) \geq \delta \Leftrightarrow i\, \boldsymbol{P}\, k \tag{7.5}$$

$$\sum_{i=1}^{N}\{u_j(z_{ij}) - u_j(z_{kj})\} + \sigma^+(i) - \sigma^-(i) - \sigma^+(k) + \sigma^-(k) = 0 \Leftrightarrow i\, \boldsymbol{I}\, k \tag{7.6}$$

with all σ^+ and $\sigma^- \geq 0$.

$\sigma^+(i)$ is an alternative's specific positive error, whereas $\sigma^-(i)$ indicates a similar negative error. These terms correspond to the possible remaining errors of the estimation process. The parameter δ on the right side of the inequality (7.5) must be set as strictly positive. The objective function F to be minimized is the sum of these errors:

$$F = \sum_{a \in A'}[\sigma^+(a) + \sigma^-(a)] \tag{7.7}$$

Two additional conditions are imposed on the utilities:

$$\sum_{j=1}^{6} u_j(z_j^{\max}) = 1, \text{ and } u_j(z_{j*}) = 0 \quad \forall\, i,j \tag{7.8}$$

The first condition normalizes the partial utility functions. It also indicates that the $u_j(z_j^{\max})$ values correspond to the criteria's relative (average) weights in the utility function. The second condition sets the lower bound partial utility at the zero level.

All these constraints (7.4) to (7.8) together with the minimization of the sum of errors constitute the basic linear goal programming model UTA-UTASTAR that is applied here. Additional details and a few other specifications can be found in Beuthe and Scannella (2001) as well as in Beuthe et al. (2005).

Since it is a linear programming model the problem of degrees of freedom does not arise from a technical point of view. However, as it is often the case, the solution may not be unique. This problem must then be solved by a post-optimality analysis. Jacquet-Lagrèze and Siskos (1982) have simply proposed to use a function that is the average of the extreme optimal functions obtained from a sensitivity analysis applied on the last bounds of each criterion. This procedure was shown to provide a practical and efficient method of estimation.

Another version of this model, named Quasi-UTA, imposes that the utility function be either concave or convex (or linear). Given the equidistance between the successive bounds, this is implemented by imposing a recursive exponential form on the partial utility functions (Scannella, 2001; Scannella & Beuthe, 2001):

$$\begin{aligned} u_j(z_{ij}(p,n)) &= \frac{\gamma - \gamma^p}{\gamma - \gamma^n} \cdot u_j^{\max}, \quad \text{for } \forall \gamma \neq 1, \text{ and } 1 \leq p \leq 5, \\ u_j(z_{ij}(p,n)) &= \frac{k-1}{n-1} \cdot u_j^{\max}, \quad \text{for } \gamma = 1, \text{ and } 1 \leq p \leq 5 \end{aligned} \tag{7.9}$$

where $n = 5$ is the number of bounds of the linear spline functions, p denotes one of the interval bounds numbered from 1 to 5, and $u_j^{\max}$ is the maximum value taken at the last bound, that is $z_{ij}^{\max}$, by the jth attribute's partial utility function. When $\gamma = 1$, the function is linear, $\gamma > 1$ defines a convex function, whereas $\gamma < 1$ leads to concavity. It is seen that this form involves only one curvature parameter plus one relative weight factor for each partial utility function, that is only 11 parameters, whereas the basic UTA involves 23 utility parameters. This specification transforms the model into a non-linear programme.

There are two basic differences between the CAs and the above multi-criteria models. Firstly, CA minimizes the sum of squared errors, whereas UTA models minimize the sum of errors in absolute values. These features imply that CA's results are more influenced by larger residuals. Secondly, the multi-criteria models impose that the partial utility functions be monotonically increasing, whereas no such constraint is imposed in CAs.

The programme MUSTARD by Scannella (2001) is used for the computations.

7.5. The Rank-Ordered Logit Models

Another approach to analyse rank-ordered preferences follows the line of discrete choice modelling, in particular logit-type models. These purport to estimate the probability of choosing an alternative on the basis of the decider's utility, which is assumed to be random with a double exponentially distributed error term. The expected utility, the estimation object, is a function of the alternatives' attributes and/or the decider's characteristics. In the context of our problem, analysing individual data, only the alternatives' attributes are considered. Hence, the model(s) can be placed in the family of conditional logit models.

Rank-ordered data require a particular specification of the stochastic utility model, which is based on the 'Ranking Choice Theorem' (Luce & Suppes, 1965). This theorem states that the probability of a joint ranking of objects can be decomposed in a product of conditional probabilities such that the ranking probability of the *n*th ranked object depends only on the set of objects which have not be ranked above it. Hence, the ranking probability of the top ranked object depends on the full set of objects to be ranked, whereas the probability of the second ranked object depends on the set of objects reduced by the subtraction of the top ranked one, and so on. The probability of the last ranked is equal to one, since there is no other object left to compare with.

This result is obtained from the assumption of independence of irrelevant alternatives, which here determines successive ranking probabilities that are not affected by the ranking of alternatives that are already better ranked. This assumption of independence generally characterizes logit models, where it implies that the ratios of the choice probabilities of two alternatives are not affected by the other alternatives in the choice set. Actually, in the framework of the stochastic utility model, the same result is obtained by assuming that the random error terms ε_j are independent and identically distributed according to a type I extreme-value distribution (Gumbel distribution).

Then, assuming that the utility function of an alternative is stochastic and linear function of its attributes

$$U_i(z_i, \varepsilon_i) = V_i + \varepsilon_i = \sum_j \beta_j z_{ij} + \varepsilon_i \tag{7.10}$$

It follows that the probability of a joint ranking $P(1, 2, \ldots, 25)$ is simply the product of logit choice probabilities with respect to successively reduced choice sets $P(i|i, i+1, \ldots, 25)$ (Beggs, Cardell, & Hausman, 1981):

$$P(1, 2, \ldots, 25) = \prod_{i=1}^{24} \frac{\exp(V_i)}{\sum_{m=i}^{25} \exp(V_m)} \tag{7.11}$$

This decomposition of the probability of a ranking into a product of probabilities of independent choices is referred to as an 'exploded logit', which exploits all the binary comparisons implied by the ranking and allows an efficient use of data by creating statistically independent choice observations (Chapman & Staelin, 1982).

This model must be distinguished from the so-called 'ordered response models' (Maddala, 1983) which would estimate the probability that an alternative falls into one among an ordered set of categories defined by additional parameters. In our case, these categories could have been the mode choice, but it would not be possible to apply that model using the stated preferences of one individual firm with respect to a set of 'abstract' alternatives. Note also that, in our context with 25 choice alternatives, a multinomial unordered choice mechanism with many more parameters was not a feasible option (Bhat & Pulugurta, 1998).

The rank-ordered model assumes that the error terms are independent and identically distributed. However, it could be the case that an interviewee gives less attention to the ranking of less attractive alternatives. Then, one might fear that these observations' errors would have larger variances. On the other hand, our ranking data in principle result from a two-steps process: in order to facilitate the interviewees' task, it was suggested that they first partition the set of alternatives into three groups according to the more or less interest they had for them, then to rank the alternatives inside each group, always with the possibility of switching any ranking to the end.

These two reasons lead us to consider a two-stages 'nested logit model' (McFadden, 1981), which possibly could tackle both aspects while maintaining the rank-ordered specification. In that case, the probability of choosing an alternative that is an element of group l would be equal to the conditional probability of choosing this alternative within group l multiplied by the probability of choosing that particular group. Nevertheless, an estimation of such a model did not provide statistically significant results. This seems to indicate that the suggested initial step of partitioning the alternatives into three groups was not rigorously applied. Both models were estimated with the MDC procedure of SAS/STAT (SAS Institute, 1989) that maximizes the likelihood function using the Newton-Raphson method.

7.6. Neural Modelling

Again, we use the notation defined in Section 7.2 and the (25×6) matrix $\boldsymbol{Z}$ composed of 25 transport alternatives, or scenarios as they are called in neural modelling, each of which is defined by the vector of values taken by the six attributes $\boldsymbol{z}_i = (z_{i1}, z_{i2}, \ldots, z_{ij}, \ldots, z_{i6})$.

The general framework is as follows: to each alternative z_i a utility index $U_i(z_i, \theta)$ is associated that may be interpreted as a utility expected by the decision maker. These utilities provide a ranking $r_i(\boldsymbol{Z}, \theta)$ of the alternatives,

$$r_i(\boldsymbol{Z}, \theta) = 1 + \sum_{k \neq i} \boldsymbol{I}\left\{U_k(z_k, \theta) < U_i(z_i, \theta)\right\} \tag{7.12}$$

where $\boldsymbol{I}(\cdot)$ represents an indicator function with value 1 or 0 according to whether the inequality in parenthesis is satisfied or not. Thus, the vector $\boldsymbol{r}(\boldsymbol{Z}, \theta) = (\boldsymbol{r}_1(\boldsymbol{Z}, \theta), \ldots, \boldsymbol{r}_{25}(\boldsymbol{Z}, \theta))$ gives a scenario's theoretical rank to be compared with the observed ranking R_i.

In this application, the expected utility function $U_i(z_i, \theta)$ is assumed to be a weighted sum of the attributes' partial utility functions subject to a logistic specification:

$$U_i(z_i, \theta) = \sum_{j=1}^{6} \omega_j v_j(z_{ij}, \alpha_j, \beta_j) = \sum_{j=1}^{6} \omega_j [1 + \exp(-\alpha_j - \beta_j z_{ij})]^{-1} \tag{7.13}$$

The logistic functions v_j's take values in the interval (0, 1] and are therefore of the same order of magnitude. The parameters ω_j, called 'neural weights' are non-negative and constrained to a sum equal to 1; they determine the relative weight of each partial utility in the global expected utility U. Reminding that $z_{ij} = 0$ means a status quo (i.e. no change) with respect to a given reference scenario, $\alpha_j > 0$ (resp. $\alpha_j < 0$) means that the partial utility function v_j is lower (resp. higher) than 0.5 at the status quo level. Furthermore $\beta_j > 0$ (resp. $\beta_j < 0$) refers to a desirable (resp. not desirable) attribute. These assumptions imply that the global utility also is scaled in the same interval (0, 1]. Altogether, there are 17 parameters to estimate: six α_j's, six β_j's plus five ω_j's.

Considering the partial utility functions as latent variables, one can take advantage of a neural approach and develop the computations in the framework of a model of intellectual processing and analysis, a so-called perceptron (for detail, see e.g. Bishop (1995) or Haykin, (1999)). This leads to an iterative procedure whereby estimation is realized by the minimization of a discrete loss function L_D, which is the quadratic error between the stated ranking and the estimated ranking produced by the model:

$$L_D(\theta) = \sum_{i=1}^{25} (r_i(Z, \theta) - R_i)^2 \tag{7.14}$$

The structure of the resulting multi-layer perceptron is illustrated on Figure 7.1 for the qth iteration. The logistic functions v_j's and the utilities being unobservable, they constitute two hidden layers of the perceptron. The first layer includes as many

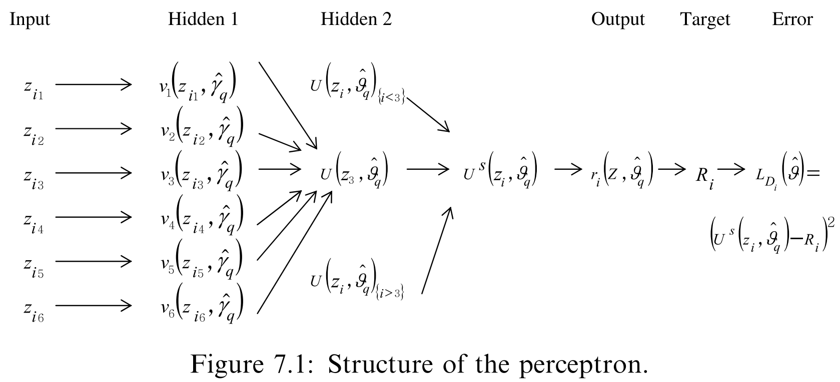

Figure 7.1: Structure of the perceptron.

neurons as attributes (6), and the second hidden layer has as many neurons as scenarios (25). In this schema, $\gamma = (\alpha, \beta)$.

The ranking function is not continuously differentiable, neither is the discrete loss function. Hence, its minimization cannot be carried out by classic algorithms such as gradient methods. In the present case, recourse is made to a so-called 'Pocket' algorithm (Gallant, 1996) which is a heuristic method allowing the minimization of a non-continuous criterion. One advantage of this method is that the detection of convergence is much easier with a smaller number of iterations.

The iterative process goes along the following steps:

1. introduce the 25 alternatives to compute the expected utility $U(z_i, \theta)$ with initial θ_0 values given to all the parameters,
2. standardize the $U(z_i, \theta)$ in such a way that its range of variation coincides with that of the observed ranking, namely [1, 25],
3. compute the distance between the re-scaled $U^s(z_i, \theta)$ and the stated rank $R_{i,}$ and compute the loss function $L_D(\theta)$,
4. update one of the parameters θ_i by a fixed step $\Delta\theta_i$, by choosing the parameter which determines the highest decrease in the loss function,
5. iterate that process by computing again the $U(z_i, \theta)$ with the new parameter $\theta_i + \Delta\theta_i$,
6. iterate steps 2–5 until convergence for obtaining the final estimated ranking.

A more detailed presentation of this procedure, along with some numerical results, may be found in Krier, Mouchart, and Oulhaj (2012).

7.7. Comparative Analysis of Results

As mentioned earlier, the analysed data are rankings of 25 transport alternatives done by nine transport managers during individual face-to-face interviews. The contrasting characteristics of their typical shipments were shown in Table 7.1.

This section presents the relevant results of the four methodologies: simplified non-metric conjoint analysis (CNM), multi-criteria analysis according to the Quasi-UTA specifications (QUTA), rank-ordered (non-nested) logit (RL), and the neural network analysis (NDL).

At this point, it is worth reminding that conjoint analysis minimizes the sum of the squared errors, whereas the multi-criteria approaches minimize the sum of errors in absolute value. Both conditional logit models are estimated through a classical maximum likelihood procedure. The neural network analysis minimizes the sum of squared discrepancies between stated and estimated ranks.

The models also differ in their specification of the partial utility functions. In the simple CNM the transformed ranking is a simple linear function of the six attributes plus one constant. The QUTA model imposes monotonically increasing partial utility functions and allows only for monotonically increasing (weakly) concave or convex utility functions. The RL model, beyond the logistic specification of the probabilities, assumes linear partial utility functions. NDL assumes that the partial utility functions are logistic.

Furthermore, the number of parameters also varies from one model to another. CNM, where the attributes are treated as continuous variables, drastically reduces the number of coefficients to 7, but adds 24 parameters for transforming the dependent variables though an iterative process. QUTA reduces the number of parameters to 11. RL relies on only six coefficients for the six attributes. NDL involves 17 parameters. From a simple descriptive point of view, the number of parameters may not matter very much; but the resulting number of degrees of freedom becomes relevant when assessing the decision making's structure of the models with a measure of its fitting quality.

The appropriate basic criterion for judging the fitting quality of the models is the Kendall's coefficient τ_K of rank correlation between the estimated and stated rankings, which varies between –1 and 1 like all correlation indexes (Kendall & Stuart, 1968). Note that only the two econometric models CNM and RL provide tests of significance for the models' coefficients.

For all models, the relative importance of an attribute for a firm, that is its weight in its decision making, is computed as an average over their domain of variation. These W_j weights directly incorporate the effect of the partial utility functions, and they are somewhat different from the weights ω_j estimated in the neural analysis. As average estimates, the W_j's do not entirely translate the information on marginal effects that is provided by non-linear models. They are computed in the following way: given the utility function $U_i(z_i) = \sum_{j=1}^{6} u_j(z_{ij})$, the weight W_j of each criteria j for a given firm is

$$W_j = \frac{\left| u_j(z_j^{\max}) - u_j(z_j^{\min}) \right|}{\sum_{j=1}^{6} \left| u_j(z_j^{\max}) - u_j(z_j^{\min}) \right|} \tag{7.15}$$

the differences being taken between the partial utilities at the attributes' highest $z_j^{\max}$ and lowest $z_j^{\min}$ values. Their ratios give an average measure of a firm's willingness to substitute a higher quality of an attribute for a lesser quality of another, in particular their basic willingness to pay for a better quality level. In this experimental set-up where the level of each attribute is measured in percentage variation with respect to the firm's current solution, the absolute monetary value of any willingness to pay could be computed with reference to the current levels of transport cost.

Table 7.2 presents the results of the non-metric CNM model. Note that, in this case, a simple metric conjoint analysis would obtain a R^2 as well as a τ_K equal to 1 since it would not have any degree of freedom. Hence, it only gives a description of the data, without providing much insight into the decision making process. Here, CNM still obtains relatively good results.

As summarized by the mean results, Cost appears as the main decision factor, followed by Reliability. The other factors affect much less the decision. However, the factors' role varies from one firm to the other: Time is important in the two cases of

Table 7.2: Weights of the simplified non-metric conjoint analysis (CNM)

	τ_K	R^2	Frequency (%)	Time (%)	Reliability (%)	Flexibility (%)	Loss (%)	Cost (%)
A	0.773	0.961	1.19	1.11	20.02	9.76	10.92	56.99
B	0.713	0.938	5.65	1.24	12.47	6.35	23.36	50.93
C	0.640	0.915	11.37	26.11	1.16	5.81	9.62	45.94
D	0.653	0.909	12.10	35.97	18.44	0.73	1.46	31.30
E	0.740	0.869	8.87	1.76	21.93	4.11	11.62	51.72
F	0.760	0.934	1.44	1.53	31.11	9.65	7.66	48.61
G	0.793	0.942	0.43	4.86	25.35	7.82	3.01	58.54
H	0.500	0.756	17.31	5.07	30.32	3.31	6.04	37.94
I	1.000	1.000	0.00	0.00	16.67	0.00	0.00	83.33
		Mean	6.48	8.63	19.72	5.28	8.19	51.70

Table 7.3: Weights of Quasi-UTA

	τ_K	Frequency (%)	Time (%)	Reliability (%)	Flexibility (%)	Loss (%)	Cost (%)
A	0.931	0.20	0.30	3.00	1.30	1.60	93.60
B	0.922	1.80	4.20	0.10	1.70	35.90	56.20
C	0.930	3.50	31.10	6.30	0.90	20.70	37.60
D	0.866	20.30	34.50	17.80	0.30	0.10	27.00
E	0.848	0.20	0.20	0.90	0.20	0.10	98.40
F	0.893	2.10	14.70	35.20	5.50	1.30	41.30
G	0.966	0.20	0.10	3.20	0.30	0.10	96.10
H	0.895	14.00	1.90	35.20	10.30	20.20	18.30
I	1.000	1.00	0.70	11.70	0.80	0.70	85.10
	Mean	4.81	9.74	12.60	2.37	8.97	61.51

the textile (C) and electronic (D) firms; the Loss factor is important to firm B (steel), and to a lesser extent to firms A (steel) and C; Frequency plays a relatively important role for firms C, D and H (pharmaceutical). Note that many signs, which are not reported here, are incorrect, and that many of the coefficients are not significant at the 5% level. In contrast, all the Cost coefficients are significant with correct signs, and so are the Reliability coefficients but with one exception (C). Note also the special case of firm I (materials), with a R^2 and τ_K equal to 1, and positive weights given only to Cost and Reliability. This indicates that this corporation focuses mainly on the Cost and Reliability factors and neglects the other ones.

Table 7.3 gives the results of the QUTA model. As to the simpler UTA, suffice it to say that it obtains very high rank correlation coefficients since it can fit very well the data with 23 parameters. Nevertheless, QUTA also reaches high levels of τ_K despite its stricter specification and its reduced 11 parameters.

The pattern of weights is similar to the one observed in the CNM results: Cost comes first with a heavier weight and is followed by Reliability. However, Time, Loss and Frequency sometimes play a more important role. The multi-criteria analysis does not provide any measure equivalent to a *t*-value for assessing the significance of the coefficients. In the special case of company I, the τ_K is still equal to one and large weights again are given to Cost and Reliability.

The results of the RL model are given in Tables 7.4 and 7.5. All coefficients of Cost have the correct negative sign and most of them are significant. It is also the case, but to a lesser extent, for Reliability. A few coefficients of Loss and Time are significant and have a correct sign in the cases where the previous analyses indicated an important weight in decision making. The Frequency and Flexibility variables are not significant. We noted in Section 7.5 that the nested specification could not be retained.

Table 7.5 gives the Kendall's coefficients of correlation and the weights obtained from the RL model. The τ_K's are lower than in QUTA. This lesser performance could result from the specification of linear partial utility functions.

The levels of weights are in concordance with the variables' levels of significance. Like in other models, Cost and Reliability have generally heavy weights in decision

Table 7.4: Estimates from the rank-ordered logit (RL)

		Frequency (%)	Time (%)	Reliability (%)	Flexibility (%)	Loss (%)	Cost (%)
A	Estimate	−0.061	0.454	8.968	1.272	−2.647	−21.667
	t Value	*−0.036*	*0.251*	*3.429*	*0.726*	*−1.368*	*−4.582*
B	Estimate	0.108	−4.218	−1.640	0.301	−6.713	−14.384
	t Value	*0.053*	*−2.052*	*−0.891*	*0.174*	*−3.212*	*−4.140*
C	Estimate	1.765	−13.941	1.157	3.002	−8.366	−15.846
	t Value	*1.108*	*−4.057*	*0.660*	*1.435*	*−3.646*	*−4.158*
D	Estimate	6.430	−11.921	5.790	1.312	0.919	−12.390
	t Value	*2.853*	*−3.999*	*2.877*	*0.721*	*0.484*	*−3.837*
E	Estimate	2.226	−0.323	6.912	2.487	2.416	−15.059
	t Value	*1.282*	*−0.177*	*3.139*	*1.411*	*1.359*	*−4.349*
F	Estimate	0.313	−1.183	10.020	2.382	3.053	−17.207
	t Value	*0.201*	*−0.675*	*3.651*	*1.308*	*1.668*	*−4.447*
G	Estimate	−0.625	−1.422	9.699	1.747	2.193	−20.903
	t Value	*−0.352*	*−0.735*	*3.759*	*0.928*	*1.093*	*−4.839*
H	Estimate	2.696	1.122	7.335	0.124	−3.633	−3.645
	t Value	*1.626*	*0.631*	*3.409*	*0.075*	*−2.008*	*−1.830*
I	Estimate	−0.022	−0.275	112.753	0.133	−0.081	−561.886
	t Value	*0.000*	*−0.001*	*0.179*	*0.000*	*0.000*	*−0.199*

Table 7.5: Weights from rank-ordered logit (RL)

	τ_K	Frequency (%)	Time (%)	Reliability (%)	Flexibility (%)	Loss (%)	Cost (%)
A	0.793	0.17	1.30	25.57	3.63	7.55	61.78
B	0.760	0.40	15.41	5.99	1.10	24.53	52.56
C	0.720	4.00	31.63	2.62	6.81	18.98	35.95
D	0.700	16.59	30.75	14.94	3.38	2.37	31.96
E	0.773	7.57	1.10	23.49	8.45	8.21	51.18
F	0.773	0.92	3.46	29.34	6.97	8.94	50.38
G	0.820	1.71	3.89	26.51	4.77	5.99	57.13
H	0.647	14.53	6.05	39.53	0.67	19.58	19.64
I	1.000	0.00	0.04	16.70	0.02	0.01	83.22
	Mean	5.10	10.40	20.52	3.98	10.69	49.31

making. Time and Loss obtain a sizable weight in a few cases. In two cases Frequency appears to play a role (D and H); the Flexibility factor looks like negligible for these nine firms.

The results of the neural network approach are presented in the two following tables. Table 7.6 gives the coefficient estimates of α and β of the logistic partial utility functions, as well as the specific neural weights ω_j. Table 7.7 provides, for each company, the Kendall τ_K and the weights W_j, which are defined in the same way as the weights computed for the previous models.

Most α coefficients are positive, which means that the partial utilities' values are lower than 0.5 at the status-quo level. The β coefficients of Frequency and Cost are positive and negative respectively, as expected. Likewise, the signs of the Reliability coefficients are correct, with one exception (B). We should note that the signs of the β coefficients corresponding to very small ω_j are meaningless, since the parameter ω_j's weigh the partial utility functions. In line with the previous results, the heaviest neural weights are given to Cost followed by Reliability, but there is a wider dispersion than in the previous models. Here, Frequency and Flexibility appear to play a more important role than in previous analyses. Table 7.7 gives the W_j weights that can be compared with those of the previous models.

Here, the Cost factor is given the highest weight among all reviewed methods. Reliability comes next with a weight as low as the one obtained through the conjoint analysis. The other factors are much less important, even though, as before they obtain higher weights for some firms. The Kendall's coefficients are between those of RL and QUTA.

7.8. Concluding Comments

Among the four analysed models, the Quasi-UTA model provides in all cases the best results in terms of the Kendall's coefficients of rank correlation, whereas

Table 7.6: Neural network NDL estimation of the coefficients

ω	**Frequency (%)**	**Time (%)**	**Reliability (%)**	**Flexibility (%)**	**Loss (%)**	**Cost (%)**
A	0.76	3.53	10.74	4.26	6.25	74.46
B	14.26	14.26	14.04	13.97	13.88	29.66
C	9.97	14.74	33.97	14.25	0.00	27.07
D	43.19	36.78	4.86	11.29	0.92	2.96
E	0.28	0.00	0.07	0.49	0.07	99.08
F	0.01	2.01	0.01	16.50	0.01	81.46
G	4.93	4.89	19.53	19.40	4.76	46.49
H	33.21	8.15	50.58	2.62	5.20	0.23
I	0.21	0.19	16.98	0.12	0.17	82.33
α	**Frequency (%)**	**Time (%)**	**Reliability (%)**	**Flexibility (%)**	**Loss (%)**	**Cost (%)**
A	0.00	– 1.20	0.00	1.20	1.80	4.40
B	3.00	2.00	– 4.00	8.50	3.50	2.00
C	3.40	1.00	2.60	8.20	1.00	– 1.40
D	2.60	3.00	2.20	4.20	0.20	– 0.60
E	4.20	0.20	2.60	3.40	7.40	2.60
F	1.00	1.00	0.00	1.00	1.00	2.00
G	5.00	4.20	1.80	2.60	8.20	2.60
H	8.00	– 5.00	7.00	– 10.00	6.00	5.00
I	1.00	0.00	1.00	1.00	1.00	1.00
β	**Frequency (%)**	**Time (%)**	**Reliability (%)**	**Flexibility (%)**	**Loss (%)**	**Cost (%)**
A	3.40	0.40	1.40	0.60	– 1.40	– 8.80
B	0.50	– 1.50	– 4.50	– 22.50	– 32.00	– 17.00
C	12.20	– 31.00	5.00	– 23.00	49.40	– 13.40
D	1.40	– 3.40	7.00	1.00	– 1.80	– 13.40
E	2.60	– 0.60	14.60	3.40	– 31.00	– 95.00
F	0.00	0.00	2.00	0.00	0.00	– 2.00
G	4.20	– 0.60	9.80	1.80	– 35.78	– 38.20
H	10.00	– 2.00	10.00	– 10.00	– 14.00	– 39.00
I	1.00	1.00	0.40	1.00	1.00	– 1.30

the neural network approach often comes close to the former results. The rank-ordered logit estimates come next at a lower level, just before those of the non-metric conjoint analysis. Considering the number of parameters involved in the various models, one may think that the Quasi-UTA model with its 11 parameters is the more reliable model of decision making. The slightly weaker performing neural network

Table 7.7: Weights of NDL

	τ_K	Frequency (%)	Time (%)	Reliability (%)	Flexibility (%)	Loss (%)	Cost (%)
A	0.907	0.67	0.27	4.02	0.69	1.67	92.68
B	0.887	0.33	2.34	1.30	0.65	34.08	61.29
C	0.720	6.30	33.78	11.08	0.87	0.00	47.97
D	0.820	18.53	28.18	16.27	6.50	0.71	29.81
E	0.820	0.00	0.00	0.04	0.02	0.02	99.92
F	0.700	0.00	0.00	0.03	0.00	0.00	99.97
G	0.940	0.11	0.03	17.26	1.56	2.12	78.93
H	0.760	9.16	5.05	37.76	0.32	23.06	24.64
I	1.000	0.00	0.00	17.05	0.00	0.00	82.95
	Mean	3.90	7.74	11.65	1.18	6.85	68.68

requires 17 parameters. Note that neither models provide measures of coefficients' significance like the *t*-values.

These two better performing models are characterized by their non-linear partial utility functions: the Quasi-UTA form can be either convex or concave, whereas the neural network functions follow a logistic form. Thus, it appears that the preferences of the nine interviewed managers, each working in very different circumstances, do not follow a linear relationship.

All methods identify the same factors that are important to the decision makers. From that point of view there is not much difference between the methods, even though there may be differences in the levels of the weights allocated to each attribute. In this small sample, Cost is by far the main factor, as usual, followed by Reliability. Time, Frequency, Flexibility and Loss are important for only some of the firms. The weight of Cost in all models strongly varies between firms, but it is generally heavy. In average over the sample of nine, it varies between 49% and 69% according to the model. This indicates that, despite the heterogeneity of behaviour, the quality attributes taken all together weigh heavily in the decision making.

This relative importance of the quality attributes and the diversity in weighting attributes also appeared in the large sample analysis (Beuthe & Scannella, 2008) where the rank-ordered logit model was applied. There, in analyses with specific cost coefficients attached to each set of subgroups of firms, all coefficients were significant and all likelihood ratio tests were conclusive. As could be expected in an analysis at an aggregate level, the Kendall rank correlations were much lower, most often below 0.40. However, Time came up as the second main factor followed by Reliability and the other factors. This outcome is rather surprising since Reliability is given more importance in most cases analysed in the present chapter. This issue should deserve additional attention.

To conclude, the present research has identified two non-linear models which performed better on this small sample of nine very diverse firms. It would be worthwhile to apply them to all firms in the survey, and see whether their aggregated results would provide more useful information. It would also be interesting to go back to the rank-ordered logit formulation to devise a model with a linear spline or logistic function, or try out a formulation where the coefficients would vary with the current transport mode of reference, a model somewhat closer to the un-ordered multinomial logit.

Nevertheless, in the field of freight transport there is such a variety of situations and logistic strategies in diverse industries that it may not be possible to explain the modal choices, even inside a specific industrial group, with just a few variables. The localization of the firm on the networks, its supply and distribution chains, its marketing strategy, the goods transported etc., are all factors that bear on the transport mode and means choices. This is a serious problem for the builders of large scale freight transport models in which transport cost and time factors are often the only modal choice variables beside, in the best cases, additional modal constants that aggregate all missing factors.

Acknowledgements

This chapter is one output of research led by a Belgian consortium directed by M. Beuthe (UCLouvain – Mons), H. Meersman and E. Van de Voorde (U. Antwerp), M. Mouchart (UCLouvain), and F. Witlox (U. Ghent). We thank the Belgian Federal Office for Scientific, Technical and Cultural Affairs (OSTC) for the financial support it granted to this project. This is an updated version of an earlier chapter (Beuthe, Bouffioux, Krier, & Mouchart, 2008).

References

Addelman, S. (1962). Orthogonal main-effects plans for symmetrical factorial experiments. *Technometrics*, *4*, 21–46.

Beggs, S., Cardell, S., & Hausman, J. (1981). Assessing the potential demand for electric cars. *Journal of Econometrics*, *17*, 1–19.

Beuthe, M., & Bouffioux, Ch. (2008). Analysing qualitative attributes of freight transport from stated orders of preference experiment. *Journal of Transport Economics and Policy*, *42*(Part 1), 105–128.

Beuthe, M., Bouffioux, Ch., De Maeyer, J., Santamaria, G., Vandresse, M., Vandaele, F., & Witlox, F. (2005). A multi-criteria methodology for stated preferences among freight transport alternatives. In A. Reggiani & L. Schintler (Eds.), *Methods and models in transport and communications: Cross Atlantic perspectives*. Berlin: Springer Verlag.

Beuthe, M., Bouffioux, Ch., Krier, C., & Mouchart, M. (2008). A comparison of conjoint, multi-criteria, conditional logit and neural network analyses for rank-ordered preference data. In M. Ben-Akiva, H. Meersman & E. Van de Voorde (Eds.), *Recent developments in transport modelling: Lessons for the freight sector* (pp. 157–178). Bingley, UK: Emerald Publishing Group Ltd. Chapter 9.

Beuthe, M., & Scannella, G. (2001). Comparative analysis of UTA multi-criteria methods. *European Journal of operational research*, *130*(2), 246–262.

Bhat, Ch. R., & Pulugurta, V. (1998). A comparison of two alternative behavioral choice mechanisms for households auto ownership decisions. *Transportation Research-B*, *32*(1), 61–75.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford: Oxford University Press.

Chapman, R. G., & Staelin, R. (1982). Exploiting rank ordered choice set data within the stochastic utility model. *Journal of Marketing Research*, *30*, 288–301.

Charnes, A., & Cooper, W. W. (1977). Goal programming and multiple objective optimizations part 1. *European Journal of Operational Research*, *1*, 39–54.

Despotis, D. K., Yannacopoulos, D., & Zopounidis, C. (1990). A review of the UTA multicriteria method and some improvements. *Foundations of Computing and decisions sciences*, *15*, 63–76.

Gallant, S. I. (1996). Perceptron based learning algorithms. *IEEE Transactions on Neural Networks*, *1*, 179–191.

Green, P. E., Carroll, J., & Goldberg, S. (1981). A general approach to product design optimization via conjoint analysis. *Journal of Marketing*, *93*, 17–35.

Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.

Jacquet-Lagrèze, E., & Siskos, J. (1982). Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. *EJOR*, *10*, 151–164.

Kendall, M. G., & Stuart, A. (1968). *The advanced theory of statistics* (Vol. 3). New York, NY: Hafner Publishing Co.

Krier, C., Mouchart, M., & Oulhaj, A. (2012). Neural modelling of ranking data with an application to stated preference data. *Statistica*, *LXXII*(3), 255–269.

Kuhfeld, W. F. (1990). SAS Technical Report R-108, Algorithms for the PRINQUAL and TRANSREG Procedures; Cary, NC: SAS Institute Inc., 21pp.

Luce, R. D., & Suppes, P. (1965). *Handbook of mathematical psychology* (Vol. III). New York, NY: Wiley.

Maddala, G. S. (1983). Limited-dependent and qualitative variables in econometrics'. Econometric Society Monographs No. 3. Cambridge: Cambridge University Press.

McFadden, D. (1981). *Econometric models of probabilistic choice*. Cambridge, MA: MIT Press.

SAS Institute Inc. (1989). *SAS/STAT® User's Guide* (Version 6, 4th ed., Vol. 2, Cary, pp. 1511–1611).

Scannella, G. (2001). *Multicriteria assessment of uncertain projects with UTA and Quasi-UTA approaches*. Ph.D. thesis, FUCAM.

Scannella, G., & Beuthe, M. (2001). Assessing risky public investments with MUSTARD. *Journal of Multicriteria Decision Analysis*, *10*, 287–302.

Chapter 8

Modelling the Emergence of Spatiotemporal Structures in Commodity Transport

Gernot Liedtke, Stefan Schröder and Li Zhang

Abstract

Logistics and transport systems transform commodity flows between single manufacturing and trading firms into spatiotemporal movement patterns of vehicles on the infrastructures. In order to analyze the impacts of policy measures aiming at influencing the behavior of various logistics actors (such as road pricing instruments, regulation, and market interventions), it is necessary to model decision making in logistics explicitly including the emergence of spatiotemporal logistics structures. This would be similar to activity-based models for passenger transport that deduce the spatiotemporal movement patterns of individuals from their planning problems. Such types of models consider the complex reactions of individuals coherently and allow for a deduction of individual welfare changes.

To model individual logistics decisions, optimization methods could be employed (network planning, or vehicle routing). The scope and constraints of the individual optimization problems, however, often result from coordination mechanisms between the different actors involved in the logistics systems. Multiagent systems allow for modelling the local interactions of these actors. The present overview describes some models dealing with the emergence of spatiotemporal structures in freight transport by integrating optimization methods and coordination mechanisms.

Three modelling efforts are presented: The first model is on the formation of regional logistics agglomerations. The second model is on the formation of tours at a national level. The third one is on the formation of transport chains

Freight Transport Modelling

ISBN: 978-1-78190-285-1

and vehicle utilization. The presentation of these models gives an outlook on the prospects of integrating optimization methods and coordination mechanisms in a multiagent system.

Keywords: Spatiotemporal structures; Optimization; Coordination mechanisms; Multiagent system; freight modelling

8.1. Introduction

Modern logistics concepts, that lead to an increasing number of small consignments and the fast growing trans-national (road) freight transports, give rise to a couple of challenges for transport policy. In particular, congestion, accelerated wear and tear of the infrastructures, particle matter emissions, and CO_2 emissions can be mentioned. Because of the strained situation of public budgets, extensive infrastructure investments cannot be expected thus, there is a need for using infrastructures and transport capacities more efficiently.

To encourage an efficient usage of existing infrastructures, transport policy has developed a set of instruments that have their origins in traffic planning, economic policy, and competition policy. With respect to commodity transport, following instruments are discussed and applied: Investment grants, regulation (vehicle technology, traffic, working conditions), fostering of collaboration, complex road pricing, vehicle and fuel taxation, subsidies for promoting rail and inland waterway transportations.

Transport models are applied to quantify the impacts of policy instruments on the transport sector. The reference standard for modelling transport — the ‘‘four-step model’’ — has initially been developed for passenger transport. Further developments integrate discrete choice models and activity-based micro-simulation which derive transport patterns from the activities of agents in space and time. The latter developments explicitly allow for studying reactions of households’ activities and mobility patterns on transport supply and incentive-oriented policy instruments.

An equivalent and coherent micro-foundation of freight demand models has not yet been achieved — there is a so-called micro–macro gap between the activities of single firms and the macrophenomenon of vehicle flows. This gap could be explained by the existence of *logistics meso-structures* (Liedtke & Friedrich, 2012; Sjöstedt, 2004). Logistics meso-structures comprise logistics networks (e.g., distribution systems or tours) of individual firms as well as collaborations and associations of the logistics networks of different firms resulting in entangled networks. Logistics meso-structures result from optimization of individual firms and from business relationships. The emergence of single logistics networks is driven by economies of scale (Bryan & O’Kelly, 1998); collaborations and associations of networks result in cost saving from synergies and an increased range of logistics service products that can be offered to the customers (‘‘demand-side economies of scale’’, cf. Rochet & Tirole, 2003). Logistics meso-structures are set up by private companies and show a high degree of variability and flexibility in adapting to demand changes. In responding to the needs of the customers/consignees, they transform the flows of goods into shipments, and transform shipments into spatiotemporal usage patterns of the infrastructures.

> *Definition:* Spatiotemporal structures in commodity transport comprise the nodes of private transport networks (warehouses, production sites, transshipment points, etc.) and movements of different types of vehicles at the right time and in a right frequency between them.

Logistic meso-structures impact the daily logistics operations of firms and thus, the vehicle flows over the infrastructure networks. They open the door to the understanding of the spatiotemporal structures in commodity transport.

The research summarized in the present overview chapter deals with methods that model the emergence of logistics meso-structures. At first it seems obvious to use optimization methods. There is a wide range of optimization tools to study the formation of distribution systems, private transport networks or tours. However, an important precondition must be met before using an optimization model — the problem needs to be well defined. In passenger travel modelling, the movement patterns of persons and vehicles can be modelled as the result of an activity scheduling problem subject to time and resource-availability constraints. In commodity transport, however, the scopes of optimization problems of individuals result from upstream coordination processes. On the one hand, demand and supply for logistics services is coordinated via markets. On the other hand, service providers and their customers coordinate their logistics processes in form of bilateral agreements.

Various commodity transport models try to cope with the phenomenon of logistics meso-structures and their impact on traffic flows. Micro-simulation models (in particular: multiagent models) go above and beyond flow-based logistics network models by mapping the local interaction between individual logistics agents pursuing their own optimization programs. A major challenge is a coherent modelling of both (i) the individual logistics optimization and (ii) of the coordination and interaction of different agents.

The present chapter shows some recent developments in the modelling of spatiotemporal structures in interregional as well as regional freight transport incorporating market-based coordination mechanisms. It is structured as follows: After these introductory remarks, methods dealing with the emergence of spatiotemporal structures in commodity transport are presented and shortly commented in Section 8.2. Section 8.3 carries out a simplified commodity transport systems analysis providing insights into the particularities of the space-time structures in commodity transport and the interdependence between market interaction and formation of spatiotemporal meso-structures. Section 8.4 presents three market-based approaches to the formation of logistics meso-structures. In Section 8.5, conclusions are drawn and a short outlook is provided.

8.2. Literature Overview

State-of-the art commodity transport models can be classified into tour-based models, flow based models and hybrid models. Tour based models for commercial transport model focus on the formation of tours generated by single delivery vehicle fleets and by sales-, trade- and craftsmen. Flow-based models deal with the routing of commodity flows through a sequence of intermodal connection nodes and warehouses (transport network models and logistics network models). Hybrid models deal with flows of goods and with tours. The following literature overview shows some typical models dealing with the emergence of spatiotemporal patterns in commodity transport and analyses briefly the modelling methods and assumptions.

Most tour-based models try to generate a large number of tours based on a random sample of observed tours. This approach can be compared to the activity-based passenger models with given and fixed household activity schedules. Schneider (2011) duplicates observed tours by putting them into a new spatial context. The duplicated tours are homomorphous to the original tours with respect to angles between tour segments, and the distances between the stops and the origin. Hunt and Stefan (2007) simulate the execution of tours based on a statistical model. Lohse, Teichert, Dugge, and Bachner (1997) and Wang and Holguín-Veras (2009) map the formation of tour patterns on a more aggregate level using statistical values, multilinear balancing/entropy maximization. All sketched models deal with the formation of spatiotemporal tour patterns, however, they are neither based on optimization and nor do they allow mapping interactions with consumers and competitors. An exception is the study of Holguín-Veras (2000), which develops an integrated approach to the modelling of vehicle routing and of oligopolistic interactions between carriers.

The model described by De Jong and Ben-Akiva (2007) is an example of an advanced flow-based model. In a multinomial logit model, (i) shipment sizes and (ii) flows through an intermodal transport network (that has been predetermined by the planner) are modelled. In this approach, flows are rerouted in the virtual and physical networks until an equilibrium state is reached. During these iterations, the network attributes are updated continuously. By doing so, the effects of consolidation (a positive interaction effect) can be mapped explicitly. In certain cases it can be imagined that links and nodes can "run dry" with the consequence that the average cost increases enormously. In the real world this means that those links and nodes would be closed. Despite of the mentioned traits of dynamic supply-demand interactions, the emergence of new structures is not part of the described modelling approach.

Wisetjindawat, Sano, Matsumoto, and Raothanachonkun (2007) present a hybrid model, which transforms commodity flows between firms into shipments, and shipments into tours. The decisions of relevant decision makers are simulated by drawing random numbers on estimated distributions. Choice models determine whether in-house transport is used or a forwarder. Forwarders are selected using a choice model, too. Liedtke (2009) simulates the coordination between shippers and carriers in interregional transport using the means of simulated calls for tenders. The number of carriers and vehicles was given and fixed. The formation of triangular and quadrangular tours could be observed, but the tours were significantly less efficient than it was observed in reality. Several studies compiled by Holmgren (2010) close the gap between optimization and simulation: A multiagent architecture is proposed to coordinate and to optimize logistics processes across company borders.

8.3. Commodity Transport Systems Analysis

8.3.1. Spatiotemporal Structures in Passenger and Commodity Transport

Activity-based modelling in passenger transport derives transport from the need of humans to pursue activities distributed in space (Bhat & Koppelman, 2003). Activity-based models reply to the challenge of mapping the effects of incentive-oriented transport policy instruments and traffic management on the spatiotemporal structure of transport demand. The approach addresses changes in mobility behavior once the households are confronted with congestion, scheduled public transit systems and incentive-oriented policy instruments aiming at increasing the efficiency of transport systems (for instance, complex road pricing). The activity-based approach requires time-use survey data. Some models consider the generation of activity episodes as exogenous inputs. Other models deal with variable activity plans using rescheduling and plan generation algorithms.

The last-mentioned activity-based simulation models with activity (re)scheduling map the emergence of spatiotemporal structures (movement patterns of people in time and space) as a result of an optimization program of households. They have the following characteristics and features:

- Consideration of constraints and conflicts at a micro level (individuals or households).
- Consideration of heterogeneity (space and time dimensions, household characteristics).
- Mapping of evasion strategies in a consistent fashion.
- Time-sensitive demand for dynamic assignment/scheduled transport modes.

A drawback of activity-based simulation with activity (re)scheduling should be mentioned: Often, creative aspects in model calibration dominate over microeconometric estimation procedures.

A formulation of a similar modelling framework for commodity transport faces the challenge of being confronted with a completely different space-time problem (see Table 8.1).

The discrepancy between the spatiotemporal structures in commodity and passenger transport has its origin in different underlying transport needs: A person needs to be consecutively present at different locations in order to execute different activities. Thus, a traveling salesman problem with a generally limited number of stops and preferred stopping durations is defined. A company, in contrast, wants to source or to deliver commodities from and to different locations dispersed in space.

Table 8.1: Characteristics of the space-time problems of households and firms

Characteristic	Passenger transport	Freight transport
Involved actors	Households and persons transport companies	Consignees, shippers, carriers and forwarders
Time-space structure	Activity scheduling with limited time budget	Time-window structure, vehicle dispatching
Microscopic flow of the moved objects	cycles	"trees"
Microscopic flow of the vessels/vehicles	cycles	cycles
Communication and coordination	Negligible interhousehold communication; intrahousehold communication	Strong interactor relations and interactions
Activity	Time consuming process at a specific location	Actions of forwarding freight, decisions concerning transport and logistics operations
Cause of transport	Time-space shifting between activities, which satisfy human needs	Provisioning and distributing for assuring the companies' functioning
Nature of constraints	(a) Physical limits of human beings (b) Limited budget (c) Limited resources (car availability)	(a) Properties of the goods (b) Legislation (such as limited behind-wheel time) (c) Coordination institutions (such as time-windows or time-tables) (d) Limited resources
Routines	Occurrence of patterns (e.g., of a regular working day)	Relevance of contracts, optimized regular working processes, for example, a regular tour
Archetypes	Some rather meaningful homogeneous family or actor types with similar behavior	Large diversity (economic activity, logistics activity, size)

Table 8.1: (*Continued*)

Characteristic	Passenger transport	Freight transport
Optimization criterion	Utility, including supposedly irrational criteria (such as prestige)	Costs and customer satisfaction. (it might be that companies have other optimization criteria and principal-agent phenomena occur)

Source: Author's own representation based on Liedtke (2006).

In a first instance, a vehicle routing problem with a variable and heterogeneous vehicle fleet is defined. In a second instance, however, further synergies and cost savings could be realized by (i) establishing hub-and-spoke structures and by (ii) combining shipments with shipments of other shippers. This leads to the formation of another type of a logistics actor — logistics service providers. In addition — because of the difference between shipments and vehicle tours — the group of carriers emerges. Consignees, shippers, logistics service providers and carriers are coordinated by market-based processes (contract and spot markets, auctions, calls for tender, brokers). Coordination on a day-to-day basis is assured through time-tables, time-windows and other formal informal agreements.

Because of the peculiarities of the space-time structure of commodity transport demand and its supply structures, a widened definition of ''activity-based freight transport modelling'' should be provided (see Liedtke & Schepperle, 2004):

> *Definition:* The activity-based modelling approach of freight transport explains, how individual operational decisions concerning logistics and transports are undertaken, in order to give indications to a traffic planner, how the whole transport system reacts on trans-national and federal transport policy measures.

The daily logistics operations of firms constitute the counterpart to activity-schedules and round-trips of travelers. The traffic-side phenomenon of the logistics operations is the spatiotemporal utilization of the infrastructures. These operations are impacted by logistics meso-structures. Amore detailed analysis of the processes giving rise to the emergence of logistics structures and logistics operations is provided in the next subsection.

8.3.2. Actors, Decisions, and Markets

Commodity transport can be explained as the result of sequence of logistics decisions. In particular, location choices, supplier choices, transport and logistics

Table 8.2: Decision problems and decision-makers in freight transport and logistics

Denomination	Description, subdivision	Responsible actor		
Production	*Location problem:* To build a factory *Production problem:* To assign production factors to production processes	Strategic level of a manufacturer	**DEMAND**	
Attraction	Supplier choice; Distribution area choice	Purchasing department (sourcing); Marketing and sales (distribution)		
Logistics network design	*Location of warehouses with commissioning and storage function. Volumes and replenishment strategies for stocks (resulting in regular shipment sizes).*	Producer or recipient or retailing corporation		**NETWORKS**
Transport logistics service provider choice	*The transportation service market can be subdivided into a set of transportation submarkets. This choice is sometimes equivalent to mode choice.*	*Producer*: Logistics department		
Transport logistics network design	Location of transportation nodes (transshipment nodes = depots and intermodal connections) and connections between them.	Transport logistics service provider ("*forwarder*"), but often the result of a self-organization process (associations)	**SUPPLY**	
Carrier choice	*To select carriers to carry shipment cases between nodes of the transportation system.*			
Dispatching and tour planning	Vehicle routing (En-route pickup and delivery and vehicle routing problems), scheduling and empty trips	Dispatcher of the *carrier or forwarder*		**TOURS**
Route choice	Detailed routing	Driver		

Source: Authors' own representation based on Liedtke, Tavasszy, and Wisetjindawat (2009).

chain choices, and vehicle routing/tour construction can be mentioned (Friedrich (2010) and Liedtke, Tavasszy, and Wisetjindawat (2009) based on Manheim (1979)). A more differentiated presentation of the decisions and systems giving rise to vehicle flows is provided in Table 8.2. It should be noted that this table presents a rather general case. It might happen in reality that there are additional decision makers; certain decisions could be skipped or undertaken as bundle decisions; complex decisions might degenerate into rather easy problems to be solved.

From Table 8.2 can be seen that several actors make decisions that influence transport demand — manufacturing and trading firms, transport logistics service providers, carriers and drivers. Each agent is in charge of a certain optimization problem: Manufacturing firms make decisions about their logistics network or just about stock replenishing strategies; transport logistics service providers deal with the construction of transport networks and transport chains (some logistics service providers in contract logistics additionally offer storage and commissioning services); fleet dispatchers deal with tour construction; drivers choose optimal routes.

The logistics processes of different actors need to be coordinated. At first, a suitable service provider has to be selected. These are the choices of (i) suppliers, (ii) transport service providers, and (iii) carriers. A strong interdependency between the individual optimization problems and the coordination processes can be identified. Transport companies extend their set of regular customers preferably by transport relations that fit well into their existing operational structures. In addition, the service providers and their clients have to coordinate their processes: The constraints in a service provider's planning problem, for instance, might be defined by the planners in charge of the clients' logistics systems in the case of a strong power asymmetry or agreed upon in negotiation processes.

Thus, when modelling commodity transport in a spatial context and in an activity-based fashion, it is necessary to understand the formation of spatiotemporal structures in commodity transport as a result of both optimization and market interactions. In the mid and long run, the number of firms and structures should be variable.

8.4. Modelling Approaches

8.4.1. Overview

In the present section, three modelling approaches to the emergence of spatiotemporal structures in freight transport are presented. They are based on following principles:

- An optimisation/behavior model of the clients (shippers).
- Economies of scale concerning the emergent spatiotemporal structures.
- Competition between different spatiotemporal supply-structures.
- Flexible transport capacities.
- Free market entries and exits.

In the three approaches, transport tariffs are determined on the basis of full cost. Oligopolistic interactions are not considered. Instead, the structures/service providers/resources are modelled endogenously. Thus, the coordination processes are mapped similarly to a monopolistic competition model (Chamberlin, 1934). Figure 8.1 gives an overview on the three approaches.

The first case (at the top of the figure) relates to the modelling of spatial location patterns of inland terminals. The star-symbol (★) shows the emergence of a terminal. There is spatial competition between inland terminals. The model maps market entries and exits of terminals.

The second case relates to the emergence of interregional truck tours performed by independent truckers. Shippers try to beat down transport tariffs. There might be transaction cost, lack of transparency, and heterogeneity. Carriers enter into the market and start a new tour only, if they can hope to recover full cost.

The third case is on the competition between different carriers operating with different truck types and forwarders employing transhipment centres.

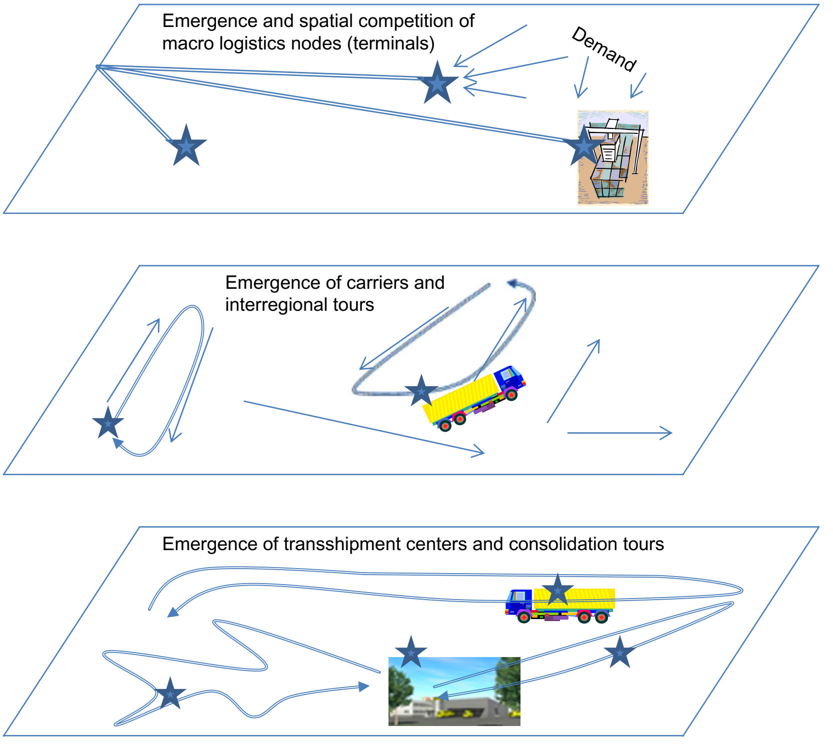

Figure 8.1: Modelling the emergence of (a) nodes (terminals), (b) tours, and (c) transport chains.
Source: Author's own representation.

8.4.2. Terminal Choice and Terminal Formation

8.4.2.1. Case description Macrologistics nodes — such as port, airports, or inland terminals — have a significant impact on the routing of transcontinental and national freight flows over infrastructure networks. Even if the node has been financed or constructed by the public sector, it can only be successful, if it attracts a sufficient critical mass of turnover. Macrologistics nodes do not only fulfill an intermodal transshipment function. In many cases, the physical transshipment facility also serves as a nucleus attracting additional logistics activities. All of these services increase the attractiveness of the node and transform an intermodal facility into a macrologistics node, which then could become a dominant factor for the regional economy. The emergence of a macrologistics node is an example of a cluster formation.

Macrologistics nodes are often subject to price and access regulation. They are in competition with nodes in other locations and with substitutive monomodal transport chains. Since macrologistics nodes are open to each transport company willing to use the node, there is strong competition between transport companies on each side of the node. This competition prevents the active transport companies from achieving high profit margins — a so-called negative network externality, (cf. Rochet & Tirole, 2003). Thus, transport companies cannot internalize agglomeration benefits.The latter ones do not only express cost savings due to collaboration between logistics firms offering complementary services nearby a logistics node. They also comprise demand side economies of scale, which describe the increase of the attractiveness due to an increasing number of attractive connections and the location of additional value added services nearby the terminal.

Due to (i) intramodal competition, (ii) regulation, (iii) the involvement of the public sector, and (iv) strong spatial competition, it could be expected that all agglomeration benefits of a macrologistics node are distributed to the shippers and forwarders using that node. A cost-related price setting and absence of oligopolistic interactions can be expected.

8.4.2.2. Strategies The strategies of three types of agents — shippers, investors, and terminal operators — are briefly described as follows.

Strategies of the shippers:

- Each shipper has a given and fixed transport demand (number of containers destined to a seaport and using intermodal transport).
- Shippers can choose between different terminals. Transport cost from and to terminals can be neglected; there are several terminals in proximity to the shipper).
- Road haulage is excluded.
- Shippers are temporally tied to a terminal.
- Law of motion: Shippers switch from terminal i to another terminal j in a time period dt according to following infinitesimal probability: $dp = \mu \cdot \exp((\alpha/2)C_i - (\alpha/2)C_j) \cdot dt$ (where c designates the total user cost of a terminal taking into account the demand side economies of scale).

Strategies of investors:

A new terminal is constructed if it is expected to survive economically in the long run. It should be noted that the construction of new transhipment infrastructure is heavily subsidized in Germany. In return they are subject to economic and access regulation.

Strategies of terminals operators:

- Terminals set their prices on the basis of full cost (including the quasi fixed cost for keeping the terminal running). Agglomeration benefits and cost savings are transmitted to the final customers (i.e., shippers): The average costs for the shippers (c) are composed of transhipment costs minus user benefits from agglomeration.
- A terminal is not competitive in the long run, if it suffers a net loss of customers below some critical mass. At this point, it becomes increasingly unattractive and finally it closes down.
- A new terminal is set up if it can be expected that it will attract enough regular customers accepting a price set on the basis of average cost.

8.4.2.3. Equilibrium The strategies describe a monopolistic competition equilibrium. Carrillo Murillo and Liedtke (2013) describe a model to the formation of macrologistics nodes based on the master function (cf. Weidlich, 2000). Another approach is rooted in a monopolistic competition model similarly to the approaches described by Anderson and De Palma (1992). These approaches are congruent with the presentation of the three agent types' strategies.

It could be shown that the average number of shippers choosing a terminal at long-term equilibrium is determined by (see Carrillo Murillo & Liedtke, 2013):

$$n^* = -\frac{1}{\alpha \cdot AC'(n^*)}$$

where
n^* = equilibrium number of cluster members
α = parameter for the sensitivity to cost differences
AC' = first derivative of the average cost function

8.4.2.4. Empirical evidence of the model The model predicts a constant terminal size (in terms of turnover). In his thesis, Carrillo Murillo (2011) analyzed the accuracy of the presented model to describe the emergence of terminals in space. For this purpose, the demand for intermodal transport was determined for the agglomeration regions in Germany. This was done by breaking down the intercontinental trade flows from and to the German regions using employment statistics. Figure 8.2 presents the number of active terminals by region against the modelled regional demand for intermodal transports.

There is a clearly visible linear component in the relationship between the number of terminals and demand. The constant (see the intercept with the vertical axis) can be attributed to the flows from and to adjacent regions using the terminals located in

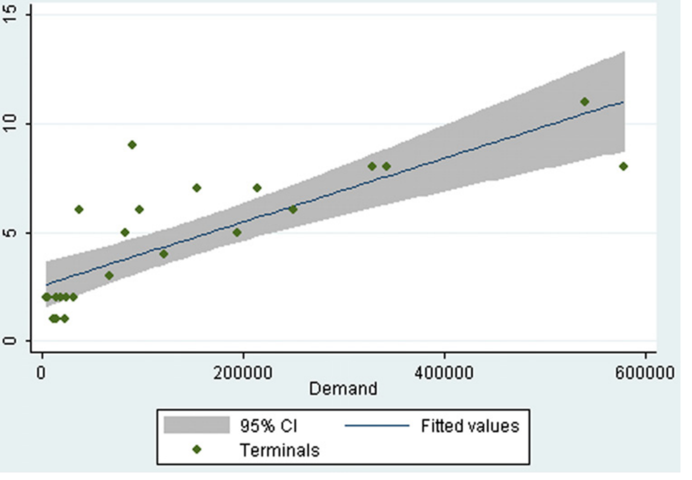

Figure 8.2: Number of terminals in function of regional demand for intermodal transports.
Source: Carrillo Murillo and Liedtke (2013).

the agglomeration regions. It can be questioned whether there is a non-linear component in the relationship between demand and terminals. This possible tendency toward larger terminals could be explained by the exhaustion of the space of characteristics of the terminals (Carrillo Murillo & Liedtke, 2013).

8.4.3. Carrier Choice and Tour Formation

8.4.3.1. Case description Interregional transports of full and partial truckloads are often operated by independent carriers. There are no significant barriers for market entries and exits because of the existing leasing market for trucks. The independent carriers try to acquire return cargo in order to construct efficient tours. For this purpose, they accept orders from different shippers (including transport cases from forwarders). In cases of regular shipments, regular tours result. A certain regularity can also be expected for consolidated mixed cargo (on the main runs) and for transport relations with high transport volume.

Because of the high (spatial) heterogeneity in this market, dynamic market entries and exits and the significant fixed cost of tours, the coordination between shipment cases and tours can be mapped by a monopolistic competition model. In our model, the dimension of time is not yet considered.

8.4.3.2. Strategies

Strategies of the shippers:

- Each shipper has a given and fixed transport demand (fixed shipment sizes, fixed frequencies) on one up to several transport relations.

- For each regular shipment, shippers conclude a transport contract with a carrier.
- If a transport contract expires, the tariff is beaten down by 10%.
- If the carrier accepts the new tariff, the contract is prolonged. Otherwise, the willingness to pay is incremented step by step until a transport company accepts the contract.

Strategies of the carriers:

- In each region, there is a potential carrier.
- Using random drawing, the carrier constructs tours (shuttle tours, triangle tours, simple quadrangular tours) bundling several shipment cases that are currently on the market (all transport contracts that will expire soon). The carriers ask the responsible shippers for their willingness-to-pay. A tour is constructed, if the tariffs for all transport cases on that tour allow for recovering the full tour cost.
- Cost coverage of the existing tours is continuously monitored. If the revenues do not allow for the full-cost coverage, the tour is closed down (special termination of the transport contracts).

8.4.3.3. Equilibrium These strategies describe a monopolistic competition equilibrium. Heterogeneity results from the spatial configuration of transport relations and tours. A certain lack of transparency results from the restricted number of market interactions (the construction of tours as well as the requests for the willingness to pay of shippers consume resources). Fixed cost relates to single vehicles. In simulation experiments relating to a single transport relation with uneven transport flows, the transport tariff drops down to marginal cost on the direction with less demand.

8.4.3.4. Simulation results We have constructed a scenario with 3 locations (zone A, zone B, and zone C) to model the impact of the transportation demand structure (Table 8.3) on the resulting market prices and on the structure of the tours (Figure 8.3). The travel distance between any two zones is 100km. It is assumed that the transport cost depends only on the travel distance. The cost of a loaded trip is 1.1€/km, whereas an empty trip causes a cost of 1.0 €/km (thus, the marginal cost of cargo on board is 0.1 €/km). Shuttle or triangle tours with maximally one empty trip are permitted.

Table 8.3: Aggregated shipper demand (in truck loads) in the transportation market

	Zone A	**Zone B**	**Zone C**
Zone A	0	20	5
Zone B	5	0	5
Zone C	50	30	0

Source: Authors' own representation.

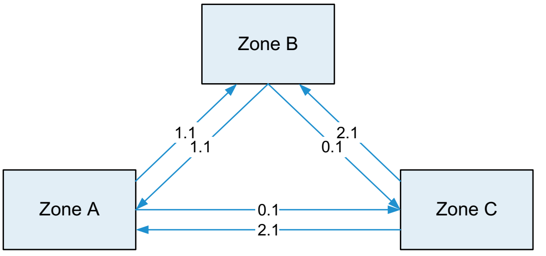

Figure 8.3: Simulated market price (in €/km) for each transport relation. *Source*: Authors' own representation.

In each simulation period, the market demand is fully met and the market price is updated. Through an iterative improvement of the tours, a market price matrix is obtained (Figure 8.3). The outbound price and inbound price of zone C differ significantly because of the considerable imbalance of transportation flows out of and into zone C. However, there is no price difference for the two transport relations between zone A and zone B because of the capability of the carriers to build triangle tours.

In the future, we are able extend this model to the German long-haul trucking market. Current developments deal with a reformulation of the zone-to-zone model into a firm-to-firm model. There are first tests dealing with the incorporation of a vehicle routing model, which finally could open the door to simulate tours in time and space — a real spatiotemporal model. The resulting price matrix will indicate the average transport price between any two NUTS2 regions and serves as a reference price on the German long-haul trucking market. In such a large-scale scenario, it seems adequate to assume that the carriers are not able to evaluate each possible route connecting the shipments that are currently on the market. In reality, this is due to transaction and information procurement cost; in the computer simulation this is due to the limited computational performance. In such situations we expect a much smoother transition of the transport tariffs between full- and marginal costs.

8.4.4. Logistics Service Provider Choice and Transport Path Formation

8.4.4.1. Case description Interregional transport systems for less than truckloads are often organized as hub & spoke systems. These systems involve the utilization of a transshipment facility that causes capital cost and operational logistics expenses. Alternatively, there is the possibility to set up en-route pickup and delivery tours (sometimes called "consolidation tours," cf. Liedtke & Schepperle, 2004).

The structure of both alternative systems and the resulting movement patterns of trucks differ significantly. There are economies of scale and density for both systems. Complex instruments of transport policy can impact the competition between these different types of logistics systems: Policies aiming at reducing the flows of heavy goods vehicles, for instance, are expected to facilitate the emergence of hub & spoke structures. Efficient delivery tours with small vehicles could lead to a reduction of the transport tariffs for small shipments and to a reduction of warehouse costs. But on the other hand, increased flow of light vehicles could negatively impact the traffic situation in the cities. To evaluate these different effects and trade-offs, a simulation model has been constructed.

8.4.4.2. Strategies

Strategies of the shippers:

- Agents choose frequency and transport service providers such that total logistics cost is minimized.
- They ask the Transport Service Provider Agents (TSP) for transport tariffs for each possible discrete shipment size.
- Having the tariff information, the shippers select a TSP agent according to a logit model (in the current model version there are only two TSPs).

Strategies of the transport logistics service providers:

- The behavior of the TSP Agent consists of building transport chains, commissioning carriers and setting up a tariff table.
- We chose to model two TSP Agents. The first TSP Agent does not have any logistics facilities at its disposal, that is, it can only offer direct transport services. The second TSP Agent operates a logistics network containing a transshipment centre.
- Agents know all modelled Carrier Agents and request services from all of them.
- During the course of the simulation, the TSP Agent collects information about all available Carrier Agents and in the process creates its own tariff table.

Strategies of the carriers:

- Carriers construct tours that start and end at their home base. Vehicle routing is modelled with an optimization approach based on the Ruin and Recreate principle (see Schrimpf, Schneider, Stamm-Wilbrandt, & Dueck, 2000).
- The price setting strategy is cost oriented. If the carrier already has some experience with similar shipments, it takes the price from its internal tariff table. If not, the associated cost is calculated based on average marginal costs, that is, the average marginal contribution to total tour costs.
- The carrier updates its tariff table in the course of a simulation.

8.4.4.3. Equilibrium In contrast with the two models presented above, this equilibrium is not a pure monopolistic competition equilibrium. Total cost recovery in the long run depends on the calculation of transport tariffs by the carriers. The prices are set on the basis of the average marginal cost (similarly to the principle of the Shapley value) and thus, some kind of activity-based costing schema results. However, the fixed cost for the warehouse and for the vehicles (time dependent component of depreciation and capital cost) are not yet considered. In simulation experiments (see the following subsection) the simulation system tends toward a dynamic equilibrium.

8.4.4.4. Simulation results The scenario is based on two time periods and a simple 8×8 checkerboard with a spike (see Schröder, Zilske, Liedtke, & Nagel, 2012). The checkerboard represents a simplified urban area. It is an undirected graph where all nodes and links have equal characteristics. Each link has a length of 1 kilometer, a capacity of 1000 vehicles per hour and a free-flow speed of 50 km/h. The spike represents the connection from our city to a distant industrial location. It is 80 kilometers long and has a free-flow speed of 100 km/h.

Four Shipper Agents are modelled, each producing commodities in the industrial zone at the right-hand side of Figure 8.4. The transport contracts are listed in Table 8.4. Each contract defines a commodity flow originating from the right hand side of the spike, and ending at the corresponding consumers' location at the left hand side of the urban area. For instance, Shipper Agent 1 has committed to consumer 5 to send 10 units with a value of 3500€ per unit. Each shipper-agent is provided with a warehouse at the consumer's location. Concerning the industry location (right-hand-side), warehousing operations and production processes (such as batch manufacturing) are not considered.

There are four carriers in our experiment. Two of them are located in the middle of the northern edge of our checkerboard (Carrier 1 and 4), whereas the other two are located in the middle of the southern edge (Carrier 2 and 3). Each carrier is equipped with exactly one vehicle. We model three types of vehicles: heavy (30 units), medium (20 units) and light vehicles (10 units). Carrier 1 and Carrier 2 are equipped

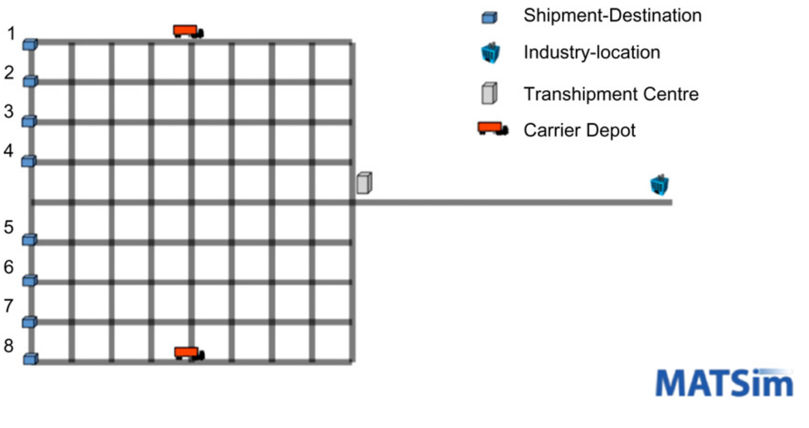

Figure 8.4: Scenario for emerging carriers and transport chains.
Source: Schröder et al. (2012).

Table 8.4: Shipper Agents and its commodity flows

Shipper	From	To	Size	Value
Shipper 1 (high value)	Industry	1	10	3500
		3	10	3500
		5	10	3500
		7	10	3500
Shipper 2 (low value)	Industry	1	10	500
		3	10	500
		5	10	500
		7	10	500
Shipper 3 (mixed value)	Industry	5	10	1200
		6	10	1000
		7	10	2030
		8	10	1500
Shipper 4 (mixed value)	Industry	1	10	1170
		2	10	2510
		3	10	1600
		4	10	2220

Source: Schröder, Zilske, Liedtke, and Nagel (2012).

with light vehicles. Carrier 3 can operate with medium sized vehicles. Carrier 4 deploys heavy vehicles.

Following scenarios have been simulated:

- *Case 1*: Reference.
- *Case 2*: Introduction of a new vehicle type. The 'heavy' carrier can now load 60 units.
- *Case 3*: Heavy vehicles in cities are prohibited.
- *Case 4*: Introduction of a city toll that amounts to 100€/day for medium vehicles, plus a toll for long distance transport amounting to 0.2 €/km.

All cases are built upon its preceding cases. For instance, when we introduce the toll, heavy vehicles are still prohibited, and the capacity of the heavy vehicle is still 60 units.

For each case, we conducted 10 model runs consisting of 50 iterations. We summarize the results of these model runs yielding to the average distance travelled, the average volumes assigned to the transport service providers as well as the average logistics costs of the shippers (see Table 8.5). It should be noted that total logistics costs relate to both inventory and transport costs.

In Case 1 — the reference scenario — all carriers travelled in total 1283 kilometers. The highest share of total kilometers exhibits Carrier 4, which has a capacity of 30 units. Carrier 1 and 2, which are with the small vehicles, travel in

Table 8.5: Simulation results

Average transport distance (in meters)					
Case	**Carrier 1**	**Carrier 2**	**Carrier 3**	**Carrier 4**	**Total**
Case 1	36,840	37,440	131,740	1,077,140	1,283,160
Case 2	18,420	0	94,300	752,580	865,300
Case 3	151,820	113,800	125,420	648,360	1,039,400
Case 4	194,820	235,660	34,800	664,580	1,129,860

Average volumes assigned to Transport Service Providers (in units)			
Case	**TSP (with TSC)**	**TSP (without TSC)**	**Total**
Case 1	0	160	160
Case 2	0	160	160
Case 3	160	0	160
Case 4	160	0	160

Average logistics costs (in €)					
Case	**Shipper 1**	**Shipper 2**	**Shipper 3**	**Shipper 4**	**Total**
Case 1	689	340	511	516	2056
Case 2	556	229	351	381	1518
Case 3	687	358	481	534	2059
Case 4	741	393	511	573	2217

Source: Schröder et al. (2012).

average less than 40 kilometers. Both of them cannot compete with the carriers employing bigger vehicles (in terms of costs). When it comes to the contract assignment to TSP Agents, operating a single leg transport chain is the most favorable solution here. The total logistics costs of all shippers amount to 2056 €, where shipper 1 — the one with the high value commodities — exhibits the highest total logistics costs. Consequently, he decides to send the quantity of its commodity flows as two shipments in almost all model runs. In contrast, Shipper 2 sends the whole quantity at once, since transport costs of a second shipment would be higher than the savings in inventory cost.

In Case 2, a new vehicle with a capacity of 60 units is introduced; it is allocated to Carrier 4. Carrier 4 can now use its large vehicle to organize a round tour from the industry area to the customers. He has an enormous consolidation advantage and takes over the transport cases of carrier 2 who drops out of the market. Total mileage traveled falls down by about 33 percent. The TSP Agent offering only direct transport chains can offer the lowest price, and thus, he is exclusively chosen by the Shipper Agents. In other words: The TSP Agent offering multileg transport chains does not reach a critical mass of turnover to divide the transport process into two separate legs. Compared to case 1, total logistics cost is reduced by 25%.

In Case 3, a ban for heavy vehicles in cities is implemented. This implies that Carrier 4 cannot operate in the urban city and thus, the efficient round-tour in Case 2 is not feasible anymore. For simplicity, we still allow Carrier 4 to use urban roads to enter and exit the city area. The solution found here is to operate a logistics network with the logistics centre at the entry point to the city. The TSP agent allocates the main leg (from the industry to logistics centre) to Carrier 4. Carrier 4 can then use its consolidation advantages in the long distance. Right from the logistics centre Carrier 1 to 3 take over the shipments. They then organize round tours from the logistics centre to final consumers. Total logistics costs rise up as the result of the changes in the transport system and now amounts to 2059 €. It should be noted that it is the existence of the extra large vehicle type enables the emergence of multileg transport chains. Without such a vehicle, the simulation system would construct direct transports (case 1).

In Case 4, we introduce a city toll for medium vehicles. The toll amounts to 100 € per day and vehicle and fall when the vehicle enters the urban area. Additionally, we introduce a toll being charged for vehicles using the interurban road. This toll amounts to 0.2 € per kilometer. As can be seen, total vehicle kilometers increase by 10 percent and contracts are shifted from Carrier 3 to the carriers with the light vehicles. Total logistics costs increase by 10 percent either, which means transportation costs are the main driver here and cannot be compensated by changing transport frequencies/shipment sizes.

8.5. Conclusions and Outlook

The state-of-the-art models for interregional commodity transport have always been behind their passenger travel counterparts. A major challenge lies in an adequate consideration of private logistics networks. Modelling spatiotemporal structures and activities in commodity transport means modelling supply and demand. When modelling the demand for infrastructure use, it is necessary to consider the networks and operations of logistics service providers.

To model the emergence of spatiotemporal structures as a result of entrepreneurial decisions of logistics firms, the problem of coordination between different actors cannot be circumvented. Besides several possibilities — such as auctions, list prices, and grandfather rights — coordination can be achieved by the market. Applying a market model as a component of a commodity transport model requires coherence between the market model, the shape of the cost functions, and the way how agents interact with each other. The market model involving economies of scale and highlighting dynamic market entries and exists is referred to as the monopolistic completion market model. The present chapter evaluates possibilities to extend this basic approach in order to model complex spatiotemporal structures:

- The first one determines the structures in equilibrium endogenously. It can be applied to spatial distributions of nodes and connections. Furthermore, econometric parameter estimations are possible.

- The second one is based on simulation. It can be applied to model the emergence of tour structures. Calibration of the model needs applying pattern matching. The approach also provides price matrices that can be used for calibration purposes, too.
- The third one is also based on simulation. It can be applied to tour structures and transshipment and consolidation operations in logistics nodes. For the moment, the model is based on crude behavior assumptions. There is a need to analyze which parameters impact the equilibrium solutions and which ones only have technical relevance.

The three approaches demonstrate that it is possible to integrate market based coordination mechanisms leading to the formation of spatiotemporal structures into a freight model. Our experiences with the demonstrated models indicate possible directions to combine these approaches:

- The equilibrium in Section 8.4.2. relates to the formation of clusters of firms (or of firms that are considered as clusters). Cluster formation happens, if there is preference for variety. Preference for variety could be expressed in form of a choice model. The spatial distribution of customers and the explicit consideration of the dimension of time might lead to additional heterogeneity that is not yet considered explicitly by this model.
- The approach highlights that new firms could only enter into the market successfully, if they have a sufficiently large mass of customers (critical company size). A simulation system of commodity markets should incorporate mechanisms modelling market entries in such an environment, for instance by mapping entrepreneurial decision making.
- If market entries and exists are enabled, long-term cost functions are relevant. They include fixed cost recovery and express economies of scale. By adding heterogeneity, market equilibriums with several up to many suppliers can be achieved.
- Since the equilibrium in Section 8.4.2. can also be achieved through agents that update their decisions from time to time, it is perfectly compatible with multiagent modelling systems.
- In Section 8.4.4 it was demonstrated that it is possible to deal with heterogeneous shipments (in terms of their value density and the firm-to-firm flows of goods). In fact, heterogeneous shipments are just one additional dimension in the space of the product characteristics of transport services. For the moment it is necessary to calculate the transport tariffs based on a cost allocation model.
- The model described in Section 8.4.3 provides a template for adding dynamic elements (in particular: market entries and exists) into the model of Section 8.4.4. Furthermore, it gives an idea how to replace cost allocation schemes completely.
- Alternatively, the model of Section 8.4.4. could map the long-hauls in an inter-regional transport model dealing with flows, shipments, and multistep transport chains.

References

Anderson, S. P., & De Palma, A. (1992). Multiproduct firms: A nested logit approach. *Journal of Industrial Economics*, *40*(3), 261–276.

Bhat, C. R., & Koppelman, F. S. (2003). Activity-based modelling of travel demand. In R. Hall (Ed.), *Handbook of transportation science* (Vol. 56, 2nd ed., 752pp.). International Series in Operations Research & Management Science. Netherlands: Kluwer.

Bryan, D., & O'Kelly, M. E. (1998). Hub and spoke networks in air transportation: An analytical review. *Journal of Regional Science*, *39*(2), 275–295.

Carrillo Murillo, D. (2011). *Demand and supply interactions in transport models: The case of hinterland transportation*. Baden-Baden: Nomos Verlag.

Carrillo Murillo, D., & Liedtke, G. (2013). A model of the formation of colloidal structures in freight transport — The case of German hinterland terminals. *Transportation Research Part E*, *49*, 55–70.

Chamberlin, E. (1934). *The theory of monopolistic competition*. Cambridge, MA: Harvard University Press.

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

Friedrich, H. (2010). *Simulation of logistics in food retailing for freight transportation analysis*. Dissertation, KIT, Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW). Retrieved from http://digbib.ubka.uni-karlsruhe.de/volltexte/1000020602

Holguín-Veras, J. (2000). A framework for an integrative freight market simulation. 2000 IEEE Intelligent Transportation Systems Conference Proceedings, October 1–3, Dearborn, MI.

Holmgren, J. (2010). *On the integration of optimisation and agent technology for transportation and production planning*. Doctoral dissertation series no. 2010:09. ISSN 1653-2090, Blekinge Institute of Technology, Blekinge.

Hunt, J. D., & Stefan, K. J. (2007). Tour-based microsimulation of urban commercial movements. *Transportation Research B*, *41*(9), 981–1013.

Liedtke, G. (2006). *An actor-based approach to commodity transport modelling*. Baden–Baden: Nomos-Verlag.

Liedtke, G. (2009). Principles of micro-behaviour commodity transport modelling. *Transportation Research Part E*, *45*, 795–809.

Liedtke, G., & Friedrich, H. (2012). Generation of logistics networks in transport models. *Transportation*. doi:10.1007/s11116-012-9386-9, available online: http://www.springerlink.com/content/0049-4488

Liedtke, G., & Schepperle, H. (2004). Segmentation of the transport market with regard to activity-based freight transport modelling. *International Journal of Logistics*, *7*(3), 199–218.

Liedtke, G., Tavasszy, L., & Wisetjindawat, W. (2009). *A comparative analysis of behavior-oriented commodity transport models*. 88th Meeting des Transport Research Board (TRB), Washington, DC.

Lohse, D., Teichert, H., Dugge, B., & Bachner, G. (1997). Ermittlung von Verkehrsströmen mit n-linearen Gleichungssystemen. In Schriftenreihe des Instituts für Verkehrsplanung und Straßenverkehr der TU Dresden, Nr. 5.

Manheim, M. (1979). *Fundamentals of transportation systems analysis*. Cambridge, MA: The MIT Press.

Rochet, J.-Ch., & Tirole, J. (2003, June). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990–1029.

Schneider, S. (2011). *A methodology for the extrapolation of trip chain data*. Dissertation. DLR-Forschungsbericht. DLR-FB 2011-18, 151 S.

Schrimpf, G., Schneider, J., Stamm-Wilbrandt, H., & Dueck, G. (2000). Record breaking optimization results using the ruin and recreate principle. *Journal of Computational Physics*, *150*, 139–171.

Schröder, S., Zilske, M., Liedtke, G., & Nagel, K. (2012). *A computational framework for a multi-agent simulation of freight transport activities*. 91th Meeting des Transport Research Board (TRB), Washington, DC.

Sjöstedt, L. (2004). A conceptual framework for analyzing policy-maker's and industry roles. In Piet R. Roger (Ed.), *Barriers to sustainable transport*. London: Tylor and Francis Group.

Wang, Q., & Holguín-Veras, J. (2009). *Tour-based entropy maximization formulations of urban freight demand*. 88th Meeting des Transport Research Board (TRB), Washington, DC.

Weidlich, W. (2000). Sociodynamics: *A Systematic Approach to Mathematical Modelling in the Social Science*, Harwood Academic Publishers, ISBN 90-5823-049-X.

Wisetjindawat, W., Sano, K., Matsumoto, S., & Raothanachonkun, P. (2007). Micro-simulation model for modelling freight agents interactions in urban freight movement. Paper presented at the 86th annual meeting of Transportation Research Board, Washington, DC.

Chapter 9

Accounting for WTP/WTA Discrepancy in Discrete Choice Models: Discussion of Policy Implications Based on Two Freight Transport Stated Choice Experiments

Lorenzo Masiero and Rico Maggi

Abstract

A key input in cost-benefit analysis (CBA) is represented by the marginal rate of substitution which expresses the willingness to pay (WTP), or its counterpart willingness to accept (WTA), for both market and non-market goods. The consistent discrepancy between these two measures observed in the literature suggests the need to estimate reference dependent models able to capture loss aversion by distinguishing the value attached to a gain from the value attached to a loss according to reference dependent theory. This chapter proposes a comparison of WTP and WTA measures estimated from models with both symmetric and reference dependent utility specifications within two different freight transport stated choice experiments. The results show that the reference dependent (or asymmetric) specification outperforms the symmetric specification and they prove the robustness of reference dependent specification over datasets designed according to different ranges of attribute levels ranges. Moreover the study demonstrates the policy relevance of asymmetric specifications illustrating the strong implications for CBA in two case studies.

Keywords: WTP/WTA discrepancy; freight choice; policy evaluation

Freight Transport Modelling

ISBN: 978-1-78190-285-1

9.1. Introduction

A key input in the economic evaluation of transport measures is represented by the marginal rate of substitution which expresses the willingness to pay (WTP), or its counterpart willingness to accept (WTA), for both market and non-market goods. Indeed, in the analysis of travel demand a lot of effort has been put into modelling individual preferences in order to obtain the trade-off between time and cost, commonly known as value of travel time saving (VTTS). In this context, Hensher (2001) reports that in the quantification of user benefits for transport project appraisal the VTTS accounts for 60%. Mackie, Jara-Diaz, and Fowkes (2001) indicate that around the 80% of the monetized benefits within cost-benefit analysis (CBA) is attributable to VTTS.

Revealed preferences (RP) and stated preferences (SP) are the main methods for collecting data suitable for the estimation of WTP and WTA measures within the discrete choice class of models (Ben-Akiva, Bolduc, & Bradley, 1993; McFadden, 1974; Train, McFadden, & Ben-Akiva, 1987). In particular, stated choice experiments have become a consolidate instrument that allows for the analysis of individual preferences by letting the respondent choose among a set of hypothetical choice situations.

Increasing attention has been paid to generating experiment designs by pivoting the hypothetical situations around individual specific reference alternatives. However, the data collected are typically modelled by specifying a symmetric utility function (that is, without distinguishing between positive and negative deviations from the reference alternative levels). Within a symmetric utility specification the WTA and WTP values are identical and this is consistent with the Hicksian surplus theory in a context where WTP and WTA are relatively small compared to the income (see Randall & Stoll, 1980 for a proof). However, the significant discrepancy between WTP and WTA measures observed in the literature[1] suggests the need to estimate asymmetric models able to capture loss aversion by distinguishing the value attached to a gain from the value attached to a loss as proposed by the reference dependent theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1991, 1992). In this regards, recent studies have analysed reference dependent utility specifications in a stated choice framework supporting the hypothesis that classic symmetric models tend to over-estimate WTP and under-estimate WTA (see e.g. De Borger & Fosgerau, 2008; Hess, Rose, & Hensher, 2008; Masiero & Hensher, 2010). Indeed, WTA/WTP discrepancy, within reference dependent theory, has been confirmed in a laboratory experiment (Bateman, Munro, Rhodes, Starmer, & Sugden, 1997) and in a stated choice experiment (De Borger & Fosgerau, 2008).

Although well recognized and discussed in several papers (e.g. Brown & Gregory, 1999; Graves, 2009a, 2009b; Hanemann, 1991) the divergence between WTP and

1. A review by Horowitz and McConnell (2002) based on 45 studies sets the median of the ratio WTA/WTP to 2.6.

WTA is not taken into account in the majority of the discrete choice model specifications resulting in potential upward biased estimates of WTP measures for policy makers. On the other hand, the estimation of reference dependent discrete choice models re-opens the debate on which measure between WTP and WTA is most desirable in the economic evaluation of transport projects.

In this study we propose a comparison of WTA and WTP measures estimated from models with both symmetric and reference dependent utility specifications within two different freight transport stated choice experiments conducted among Swiss logistics managers in 2003 and 2008, respectively. The freight transport sector occupies a minor part in the research literature involving the transport sector in general. However, the impact of the value of freight transport time saving (VFTTS) in the evaluation of the profitability of investments in transport infrastructures must not be neglected since it can represent up to 50% of the potential VTTS (Zamparini & Reggiani, 2007). In particular, we focus the analysis on testing the robustness of the loss aversion property (and WTA/WTP divergence) within two pivoted freight transport stated choice experiments defined under different experimental design assumptions. The results are based on the estimation of random parameters logit models on both the single dataset collected in 2003 and the pooled dataset containing the two stated choice experiments.

In the derivation of WTP and WTA measures, the selection of the density function for the random parameters has a great impact. Indeed, if all parameters are set to be random then the estimation of the marginal rate of substitution involves the ratio of two random distributions leading to substantial evaluation problems. Train and Weeks (2005) propose the estimation of discrete choice models in WTP space overcoming the problem of ratio distributions as the WTP distribution is directly introduced in the model estimation. However, the estimation of models in WTP space requires the normalization of the model for the cost attribute. This is a restriction for reference dependent models since they involve two cost attributes: associated with gains and losses, respectively (see Rose & Masiero, 2010). Masiero and Hensher (2010) specify a reference dependent model where the random parameters are assumed to be triangular distributed and constraining the standard deviation of the coefficient to be equal to the mean. Although the use of constrained triangular distribution leads to desirable estimates of the parameters, the heterogeneity across the sample is only assumed and not estimated. In order to analyse the spread of the random parameters distribution associated with gains and losses, we decide to fix the cost parameter and let the parameters associated with the other attributes to be Normal distributed. This method has good properties in terms of model identification, WTP estimation and rational assumption about the cost coefficient (see Revelt & Train, 2000).

A specific purpose of this study is to discuss the policy implications that arise from the WTA/WTP discrepancy. In this context, we propose to reconsider the concept of WTP and WTA in transport investment appraisal focusing the discussion on the rationale of using asymmetric WTP and WTA instead of symmetric WTP. We illustrate the discussion with two hypothetical infrastructure investments, involving either an improvement or a maintenance of the current situation.

The chapter is organized as follows. In Section 9.2 we describe the two stated choice experiments used in the analysis. The methodological background is presented in Section 9.3 whereas the model estimates are shown in Section 9.4 along with comments on the results. In Section 9.5 we outline the potential policy implications associated with WTA/WTP discrepancy. Conclusions and final remarks are provided in Section 9.6.

9.2. Data

The data refers to two freight transport stated choice experiments conducted among Swiss logistics managers in 2003 and 2008, S-2003 and S-2008, respectively. The first experiment involves the evaluation of relevant service characteristics in freight transport (see Maggi & Rudel, 2008 for details) whereas the second experiment is part of a project[2] aimed to analyse the infrastructure vulnerability of the Gotthard corridor, one of the most important European transport corridors (see Masiero & Maggi, 2012 for details).

The freight transport services considered in the two stated choice experiments refer to conventional origin-destination services and are characterized by cost (CHF per transport service), time (hours per transport service) and punctuality (percentage of transport services arriving on time per year). An additional attribute is considered in the first experiment expressing the yearly percentage of transport services which register damages to the goods transported.

The cost and time attributes for the hypothetical alternatives included in the design of the two stated choice experiments are constructed in terms of positive and negative percentage deviations from the values of a typical freight transport service (i.e. reference alternative) previously described by the logistics managers. Although logistics managers also reported reference values for punctuality and damages, these attributes are presented in absolute values for technical convenience. The levels associated with each attribute in the two experiments are shown in Table 9.1 which also highlights the main differences between the two experimental designs.[3]

Both experiments have been created using orthogonal design techniques[4] (performed with Sawtooth software) and the data have been successively collected through face-to-face Computer Assisted Personal Interviews (CAPI), where logistics managers were asked to indicate their preferred alternative in each choice task. The

2. NFP54 ‘Sustainable Development of the Built Environment’, funded by the Swiss National Science Foundation.

3. The attribute levels for dataset S-2003 differ from those reported in Maggi and Rudel (2008) which by mistake are not correctly reported there.

4. We acknowledge the recent move form orthogonal to efficient design techniques (see e.g. Bliemer & Rose, 2009). For additional details on efficient designs involving the reference alternative see Rose, Bliemer, Hensher, & Collins (2008).

Table 9.1: Description of the stated choice experiments

	DATASET S-2003	**DATASET S-2008**
Attributes and Levels		
Transport cost (CHF)	−40%, −20%, Reference, +20%, +40%	−10%, −5%, Reference, +5%, +10%
Transport time (hours)	−40%, −20%, Reference, +20%, +40%	−10%, −5%, Reference, +5%, +10%
Transport punctuality (%)	96%, 98%, 100%	96%, 98%, 100%
Damages (%)	6%, 4%, 2%	
Design		
Experiment	Unlabeled	Labeled
Alternatives	Alternative A and Alternative B	Road, Piggyback and Combined transport
Reference in design	Not included	Road
Number of choice tasks	20	15

sample focused on medium (50–249 employees) and large (more than 249 employees) companies. Regarding S-2003 data, 35 firms of the food and wholesale sector were involved and a subset of the sample participated to the experiment twice, for inbound and outbound transport services, respectively. After having removed the extreme cases in order to obtain similar range of minimum and maximum values (in terms of cost, time and punctuality values revealed by logistics managers) across the two experiments, S-2003 data consists of 42 cases, representing 840 choice observations. The sample associated with S-2008 data is composed of 27 firms operating in the manufacturing sector, representing 405 choice observations. By pooling the two datasets we obtain 69 valid cases, representing 1245 choice observations. The descriptive statistics of the reference transport services described by logistics managers are reported in Table 9.2.

9.3. Methodology

Within the random utility models (RUM) framework, the utility function associated with respondent n for alternative j in choice task s is defined as the combination of a systematic component and an unobserved component, where the systematic part is typically assumed to be linear in parameters such that

$$U_{njs} = \alpha_j + \sum_{k=1}^{K} \beta_{nk} x_{njsk} + \varepsilon_{njs} \tag{9.1}$$

Table 9.2: Descriptive statistics for attributes of the reference transport service

Variable	Mean		Median		SD		Min		Max	
	S-2003	S-2008	S-2003	S-2008	S-2003	S-2008	S-2003	S-2008	S-2003	S-2008
Cost (CHF)	894.4	1300.1	800	1000	533.1	1152.9	120	136	2500	5400
Time (hr)	15.1	33.3	7	24	26.3	27.3	2	2	168	96
Punctuality (%)	98.5	96.5	99	98	1.7	3.0	94	90	100	100
Damages (%)	0.3	-	0	-	0.6	-	0	-	2	-

where α_j represents the alternative specific constant, $\boldsymbol{\beta_{nk}}$, is the vector of k coefficients associated with the set of attributes, and the unobserved part, $\varepsilon_{\mathrm{njs}}$, is independent and identically distributed (IID) extreme value type 1. The subscript n in $\boldsymbol{\beta_{nk}}$ refers to the random parameters logit class of models, where the coefficients (all or a subset) are assumed to be heterogeneous across the respondents according to a specific density function. In this context, the Normal distribution is the most referred in the literature although log-normal and triangular distributions are also used (see Hensher & Greene, 2003).

The derivation of the marginal rate of substitution between the k-th attribute and cost is straightforward and leads to WTP and WTA estimates. For symmetric specification models they are defined as follows:

$$\mathrm{MRS} = \frac{(d/dx_{njsk})\beta_{nk}x_{njsk}}{(d/dx_{njs,\mathrm{cost}})\beta_{n,\mathrm{cost}}x_{njs,\mathrm{cost}}} \Rightarrow \mathrm{WTP} = \mathrm{WTA} = \frac{\beta_{nk}}{\beta_{n,\mathrm{cost}}} \tag{9.2}$$

As shown in Eq. (9.2), symmetric models assume by construction that WTP and WTA are identical.

A departure from the classic symmetric model specification, formulated in Eq. (9.1), is represented by the reference dependence model specification which allows the estimation of different coefficients for both positive and negative deviations from the reference values. The utility function is then defined as follows:

$$U_{njs} = \alpha_j + \sum_{k=1}^{K} \beta_{nk}(\mathrm{dec})x_{njsk}(\mathrm{dec}) + \sum_{k=1}^{K} \beta_{nk}(\mathrm{inc})x_{njsk}(\mathrm{inc}) + \varepsilon_{njs} \tag{9.3}$$

where (dec) and (inc) indicate decreases and increases respectively, and x_{njsk} (dec) = $\max(x_{\mathrm{ref}} - x_j, 0)$ and $x_{njsk}(\mathrm{inc}) = \max(x_j - x_{\mathrm{ref}}, 0)$. The estimation of different parameters for gains and losses with respect to the reference values allows to test for asymmetries in the utility function[5] and eventually for the presence of loss aversion. Moreover, the WTP and WTA measures are not constrained to be

5. Note that the reference dependence specification nests the symmetric specification.

identical anymore since they are separately estimated according to the following relations:

$$\text{For undesirable goods: WTP} = \frac{\beta_{nk(\text{dec})}}{\beta_{n,\text{cost(inc)}}}; \quad \text{WTA} = \frac{\beta_{nk(\text{inc})}}{\beta_{n,\text{cost(dec)}}} \tag{9.4}$$

$$\text{For desirable goods: WTP} = \frac{\beta_{nk(\text{inc})}}{\beta_{n,\text{cost(inc)}}}; \quad \text{WTA} = \frac{\beta_{nk(\text{dec})}}{\beta_{n,\text{cost(dec)}}} \tag{9.5}$$

The relationship between loss aversion and WTA/WTP divergence can then be easily shown from Eqs. (9.4) and (9.5). In fact, loss aversion holds if the absolute value of the coefficient associated with losses is bigger than the absolute value of the coefficient associated with gains. That is, for undesirable goods: $|\beta_{nk(\text{inc})}| > |\beta_{nk(\text{dec})}|$; whereas for desirable goods: $|\beta_{nk(\text{dec})}| > |\beta_{nk(\text{inc})}|$. If loss aversion holds for both coefficient at the numerator and coefficient associated with the cost attribute, then WTA > WTP, independently of the degree of loss aversion.

Given the panel structure of the data and the use of the random parameters logit class of models, the estimation of the utility parameters is derived from the maximization of the following simulated log likelihood:

$$LLn = \sum\nolimits_n \ln \frac{1}{R} \sum\nolimits_r \prod\nolimits_s \frac{\exp(\boldsymbol{\alpha}_j + \boldsymbol{\beta}_n' \mathbf{x}_{njs})}{\sum_j \exp(\boldsymbol{\alpha}_j + \boldsymbol{\beta}_n' \mathbf{x}_{njs})} \tag{9.6}$$

where $s = 1, \ldots, S$ represents the number of choice situations whereas $r = 1, \ldots, R$ refers to the number of draws.[6]

9.4. Model Results

The estimation of symmetric and reference dependent models is performed firstly on the S-2003 data and then on a joint dataset, where S-2003 and S-2008 datasets are pooled (for model estimation based on S-2008 see Masiero & Hensher, 2010[7]). This allows us to test for robustness of reference dependent specification across different datasets. The estimation of the models for the pooled dataset includes the computation of the scale parameters for the three alternatives of dataset S-2008 allowing to control the difference in the scale associated with different datasets. In particular, as suggested by Brownstone, Bunch, and Train (2000), the scale

6. Refer to Train (2003) for details.

7. Note that Masiero and Hensher (2010) use constrained triangular distribution for random parameters whereas here we use unconstrained normal distributions for attribute parameters and a fixed cost coefficient.

parameters can be derived by introducing a set of alternative-specific constants having zero mean and free variance.[8] In doing this, we normalize the scale of S-2003 dataset to one upon the second dataset (see Hensher, 2008 for additional details on this procedure). The estimation of the models is based on 500 Halton draws and performed using Nlogit 4.

The model results are shown in Table 9.3. The first two columns (M1 and M2) refer to symmetric model specification and reference dependent model specification for S-2003 data whereas the last two columns (M3 and M4) refer to the same model specifications but for the pooled dataset. The overall evaluation of model fits is based on the log-likelihood at convergence, the Akaike's Information Criterion (AIC) and the pseudo ρ^2.[9]

Comparing these three measures we register that the reference dependent model specification outperforms the symmetric one in both the datasets used. In particular, the pseudo ρ^2 increases from 0.3652 to 0.4068 for S-2003 data and from 0.3396 to 0.3973 for the pooled dataset. These findings further exclude the hypothesis that the restricted symmetric models are more parsimonious than the unrestricted reference dependent models.

The scale parameters estimated for the alternatives of dataset S-2008 result statistically different from 1 (for the two hypothetical alternatives) providing evidence for a significant difference in the scale of the two datasets used in the analysis. In particular, the scale parameters for piggyback and combined transport alternatives indicate that the associated unobserved effects are characterized by a considerably lower variance compared to dataset S-2003. On the contrary the unobserved effect associated with the reference alternative does not report a different variance if compared with the alternatives in dataset S-2003.

Examining the coefficient estimates for the symmetric models (M1 and M3) we observe that they all are of the expected sign that is, negative for damages, cost and time attributes and positive for punctuality. Both mean and standard deviation (for random parameters) estimates result statistically significant at an alpha level of 0.05 except for the standard deviation of the time parameter in S-2003 data which results statistically significant at an alpha level of 0.10. Looking at the reference dependent model specifications (M2 and M4), where cost, time and punctuality attributes are defined in terms of gains and losses, we observe a similar consistency. That is, parameters associated with gains (cost decrease, time decrease and punctuality increase) are positive in sign whereas the parameters associated with losses (cost increase, time increase and punctuality decrease) are negative in sign. Moreover, we find that loss aversion holds for all the three attributes and in both S-2003 and pooled datasets. In fact, the parameters associated with losses are in absolute value greater than the parameters associated with gains. The standard deviation for the random

8. Indeed, the scale parameter (λ) can be derived from the following relationship: $\lambda = \pi/\sqrt{6}\sigma_{ASC}$, where σ_{ASC} is the standard deviation of the alternative-specific constant.
9. Note that the pseudo ρ^2 is calculated as 1-(L(β)/L(ASC)).

Table 9.3: Model results

	M1		M2		M3		M4	
	Symmetric		Reference Dependent		Symmetric		Reference Dependent	
	Par.	**(t-ratio)**	**Par.**	**(t-ratio)**	**Par.**	**(t-ratio)**	**Par.**	**(t-ratio)**
Means for Random and Non-Random parameters								
ASC Alternative A	0.1223	(0.93)	0.1284	(1.16)	0.1480	(1.52)	0.1599	(1.36)
ASC Piggyback	–	–	–	–	−1.0933	(−1.71)	0.8495	(1.05)
ASC Combined transport	–	–	–	–	−0.8716	(−1.37)	1.0408	(1.28)
Cost	−0.0038	(−12.59)	–	–	−0.0036	(−12.68)	–	–
Time	−0.0691	(−2.91)	–	–	−0.0740	(−3.49)	–	–
Punctuality	0.2890	(6.37)	–	–	0.2880	(9.45)	–	–
Damages	−0.3959	(−10.76)	−0.4870	(−10.39)	−0.4042	(−10.74)	−0.5303	(−10.54)
Cost decrease	–	–	0.0033	(5.69)	–	–	0.0041	(6.71)
Cost increase	–	–	−0.0052	(−7.88)	–	–	−0.0060	(−8.56)
Time decrease	–	–	0.0662	(1.82)	–	–	0.0809	(2.49)
Time increase	–	–	−0.0718	(−2.39)	–	–	−0.1315	(−2.83)
Punctuality decrease	–	–	−0.3454	(−2.94)	–	–	−0.6127	(−4.15)
Punctuality increase	–	–	0.2640	(2.11)	–	–	0.2272	(2.76)
Standard deviations for Random parameters								
Ns Time	0.0586	(1.92)	–	–	0.08504	(3.32)	–	–
Ns Punctuality	0.3395	(5.90)	–	–	–	–	–	–
Ns Time decrease	–	–	0.0772	(2.78)	–	–	0.1017	(2.76)
Ns Time increase	–	–	0.1013	(1.91)	–	–	0.1807	(2.75)
Ns Punctuality decrease	–	–	0.6099	(5.18)	–	–	0.8077	(5.52)
Ns Punctuality increase	–	–	0.3812	(3.15)	–	–	0.3215	(4.10)

Table 9.3: (*Continued*)

	M1		**M2**		**M3**		**M4**	
	Symmetric		**Reference Dependent**		**Symmetric**		**Reference Dependent**	
	Par.	**(t-ratio)**	**Par.**	**(t-ratio)**	**Par.**	**(t-ratio)**	**Par.**	**(t-ratio)**
Scale parameters								
Scale ALT Piggyback	–	–	–	–	19.384	(8.72)[a]	15.952	(5.52)[a]
Scale ALT Combined transport	–	–	–	–	6.854	(1.85)[a]	6.704	(2.26)[a]
Scale ALT Reference	–	–	–	–	0.417	(−0.31)[a]	0.297	(−0.58)[a]
Conditional WTP measures [standard deviation]								
Travel Time	17.69 [7.22]		12.61 [6.16]		20.39 [12.49]		13.23 [6.33]	
Punctuality	62.80 [48.70]		52.44 [36.57]		79.16 [0.00]		36.30 [27.13]	
Conditional WTA measures [standard deviation]								
Travel Time	17.69 [7.22]		21.72 [13.52]		20.39 [12.49]		32.71 [24.23]	
Punctuality	62.80 [48.70]		101.65 [137.57]		79.16 [0.00]		152.22 [146.38]	
Model Fits								
Number of Observations	840		840		1245		1245	
Log-likelihood (ASC)	−582.05		−582.05		−1016.21		−1016.21	
Log-likelihood (β)	−369.48		−345.28		−671.13		−612.47	
Number of Parameters	7		12		11		17	
AIC normalized	0.8964		0.8507		1.0958		1.0112	
Pseudo ρ^2	0.3652		0.4068		0.3396		0.3973	

L(ASC) refers to the log-likelihood function for the model estimated with alternative specific constants only.
L(β) refers to the log-likelihood function at convergence.
[a]The *t*-ratio is calculated on the assumption that the scale parameter is different from one.

parameters results higher for the parameters associated with losses meaning that the preferences of the logistics managers are more heterogeneous when logistics managers are faced with losses.

The conditional estimates for WTP measures from symmetric models are in line with current research literature (see Zamparini & Reggiani, 2007 for a review). In particular, the WTP for time is 17.7 CHF/hour (approximate exchange rate, 1 CHF = 1.05 USD) and 20.4 CHF/hour for symmetric models M1 and M3, respectively. The WTP for punctuality is a key factor, as also recognized by similar studies (e.g., Danielis, Marcucci, & Rotaris, 2005; Fowkes, 2007), where logistics managers show a considerable sensitivity regarding punctuality of the transport service. For symmetric models the WTP for punctuality is 62.8 CHF and 79.16 CHF per percentage point for M1 and M3, respectively.

Looking at the reference dependent model specifications in M2 and M4, we are able to distinguish between WTP and WTA. In particular, referring to the estimates for the pooled dataset (M4) we find that the WTP for time is 13.23 CHF/hour whereas the WTA for time is 32.71 CHF/hour. On the other hand, the WTP for punctuality is 36.30 CHF for an increase in punctuality by one percentage point whereas the WTA is 152.22 CHF for a decrease of punctuality of one percentage point. Punctuality still remains a crucial factor, especially when logistics managers are faced with a reduction of this service attribute. The WTA/WTP discrepancy registered is fairly marked for both the marginal rates of substitution considered. In particular, the ratio WTA/WTP is 2.5 for time and 4.2 for punctuality resulting in line with past studies (see e.g. Horowitz & McConnell, 2002 for a review).

9.5. Policy Implications

In the previous section we found that the estimation of reference dependent choice models can lead to asymmetric estimates of WTP and WTA measures in line with the literature on WTA/WTP discrepancy. This evidence has implications on the economic evaluation of transport investments and on the appropriate use of WTP and WTA measures. Different policy measures or infrastructure investments are designed for different purposes implying the use of either WTP or WTA values. In this section we focus on infrastructure investments and in particular on transport projects, defining three categories according to the expected outcome of consumers WTP and WTA values.

In Table 9.4 we introduce the expected consumers' WTP and WTA values associated with a new infrastructure depending on whether the impact on actual conditions represents a worsening, a conservation or an improvement in terms of consumers' utility. Infrastructures that lie in the worsening category are those which carry considerable environmental consequences such as the construction of a nuclear power station. In this case, the expected WTP for having a new nuclear power station is expected to be negative whereas the WTA is expected to be particularly high. Typically the calculation of the social impact associated with such infrastructures is

Table 9.4: Expected consumers WTP and WTA values due to an infrastructure investment

	Impact on actual conditions		
	Worsening	**Conservation**	**Improvement**
WTP	Negative	Zero	Positive
WTA	Positive	Positive	Positive
Appropriate measure	WTA	**WTA**	**WTP**

based on ad-hoc stated choice experiments designed directly in the WTA space. Since in this study we are interested in the economic appraisal of transport infrastructure investments we do not discuss this category any further.

Within transport projects, many investments deal with the conservation of the current conditions. Indeed, transport infrastructure operation and maintenance are necessary in order to maintain a certain level of quality (e.g. travel time) that would otherwise be impossible to maintain due to the constant increase of traffic flows. These infrastructure investments can often be very expensive, depending on the transport network involved, and the convenience of the investment needs to be evaluated. In this case, the WTP is expected to be zero since we are asking the users to face a situation where the quality of service remains stable at the actual level. Therefore, the user benefit associated with such investments should be calculated using their WTA for a loss in service quality (e.g. an increase in the travel time) which would be the consequence if the investment were not realized.

The typical situation in the economic appraisal of a transport project is however the evaluation of an investment against an improvement of the actual condition. This is the case of a new transport infrastructure, where the WTP is now positive and reflects the maximum (marginal) amount that consumers are willing to pay for the improvement (e.g. a reduction of the travel time). Therefore, the WTP should be used in the computation of user benefits.

9.5.1. Case Studies on Freight Transport

Based on the estimates from models M3 and M4 (reported in Table 9.3) we illustrate the implication of WTA/WTP discrepancy in the case of hypothetical policy measures for freight transport in Switzerland. In particular, we compare two CBAs distinguishing between the two categories highlighted in Table 9.4, conservation and improvement, respectively.

We hypothesize two different large investments along the Gotthard corridor, one of the most important link across the Alpine region. The first investment refers to the construction of a second 'Gotthard road tunnel' increasing the number of lanes

Table 9.5: Case studies assumptions

Setting	
Initial cost	900,000,000
Annual maintenance cost	50,000
Discount rate	4.5%
Population	650,000
Infrastructure lifetime (years)	50
Scenario	
Change in Time attribute	10%
Change in Punctuality attribute	1%

from two to four and representing a significant improvement in terms of travel time and punctuality. The second investment consists of protective galleries and tunnels on the north and south access to the Gotthard road tunnel. This represents a maintenance intervention assuming that climate change leads to a dramatic increase of hazards.

Table 9.5 describes the case studies. We realistically assume for both projects an identical initial cost of 900 millions CHF[10] and we set the annual maintenance cost to 50,000 CHF. The population is set to 650,000 units according to the Swiss transport policy goal regarding the annual number of trucks allowed to cross the road corridor after 2018. The infrastructure lifetime and the discount rate are assumed to be 50 years and 4.5%, respectively.

The hypothetical scenario envisages a reduction of freight travel time of 10% and an increase in the punctuality of the freight transport services of 1%. In the first case these improvements are due to the elimination of queues caused by the current bottleneck. In the second case we assume that the increasing hazards would cause an increase in the travel time and punctuality which could be avoided by the investment. Given this scenario and given the WTP and WTA estimates from models M3 and M4 (for convenience reported in Table 9.6), we calculate the average generalized cost of the transport services described by logistics managers and the average generalized cost of the same transport services but under the scenario assumptions.[11] In particular, we apply asymmetric WTP in the first case and asymmetric WTA in the second case further reporting, for comparison, symmetric WTP for both cases. The benefits for the freight transport sector associated with the scenario are then derived

10. The reference cost for the second 'Gotthard road tunnel' is based on the estimate published in 'Ticino Business', *Camera di commercio, dell'industria, dell'artigianato e dei servizi del Cantone Ticino, Lugano*, November 2008.

11. The generalized cost is calculated as the sum of the cost, time and punctuality where time and punctuality are expressed in monetary values according to the WTP and WTA estimates.

Table 9.6: Case studies results

	Case 1 Improvement Asymmetric WTP	Case 2 Maintenance Asymmetric WTA	Case 1 = Case 2 Symmetric WTP = WTA
Travel time	13.23	32.71	20.39
Punctuality	36.30	152.22	79.16
Net present value	– 57 millions	1988 millions	698 millions

by taking the difference of the generalized costs, under actual and scenario conditions, both calibrated to the population considered.[12]

The results, shown in Table 9.6, suggest the relevance of estimating WTP and WTA separately and applying them appropriately. Using the traditional approach, and hence overestimating WTP for improvement and underestimating WTA for maintenance, both projects would be accepted. Using the asymmetric approach, the improvement would result in a negative net present value indicating that a 900 millions investment would not be justified for a 10% and a 1% percent improvement in travel time and punctuality, respectively. On the other hand, an equally expensive maintenance investment with same impact is largely justified. In a general sense it is therefore demonstrated that the use of symmetric WTP may result in over- or under-estimating the economic appraisal of transport investments.

9.6. Conclusions

This chapter has investigated the policy implications of WTA/WTP discrepancy in a freight transport context. The analysis has focused on the estimation of discrete choice models for two freight transport stated choice experiments. In particular, we estimated a set of random parameters logit models comparing between the classic symmetric specification which does not distinguish between WTP and WTA, and the reference dependent specification which allows for the estimation of different parameters associated with gains and losses. We outlined then the policy implications supporting the discussion with hypothetical examples on the freight transport sector in Switzerland.

12. To be noted that in the computation of the benefits we did not distinguish for intra-country transports and transports that use the corridor as connection between different countries. Indeed, it is reasonable to assume lower WTP values for the latter transport segment. However, we are convinced that our estimates are still conservative since we fixed the population to 650,000, the Swiss policy objective, representing around the 50% of the actual figure.

The results show that the reference dependent specification outperforms the symmetric specification and they support the robustness of the reference dependent specification for datasets designed to accommodate different attribute level ranges. Loss aversion has been registered for all attributes investigated in the analysis leading to a significant WTA/WTP discrepancy. In particular, our results confirm the findings reported in the recent research literature suggesting that symmetric models tend to overestimate WTP values and to underestimate WTA values.

The policy implications associated with WTP and WTA measures estimated from reference dependent choice models are indeed interesting. We defined three main categories of infrastructure projects labelled worsening, conservation and improvement, respectively. For each category, the two measures (WTP and WTA) have been discussed and the most appropriate measure for the evaluation of the investment has been selected. A detailed analysis, based on case studies, has then been proposed on the two categories that typically reflect transport projects, that is transport infrastructures aimed to conserve or to improve the actual quality of the service. We pointed out a major difference between these two categories suggesting that the infrastructures aimed to conserve the actual conditions should be evaluated using the consumers WTA whereas the infrastructures aimed to improve the actual conditions should be evaluated using the consumers WTP. Based on reference dependent model estimates and given our distinction for the type of infrastructure, we conclude that the evaluation of investments aimed to conserve (improve) the actual conditions is underestimated (overestimated) if symmetric models are used.

Finally, we strongly encourage policy oriented analysts to further estimate and test reference dependent choice models appropriately derived from reference pivoted choice experiments. The persistence in only using symmetric discrete choice models as an instrument for deriving marginal substitution effects for policy purposes might most probably lead to biased evaluation in the form of consistent overestimation or underestimation of the economic benefits of transport projects.

References

Bateman, I., Munro, A., Rhodes, B., Starmer, C., & Sugden, R. (1997). A test of the theory of reference-dependent preferences. *Quarterly Journal of Economics*, *112*, 479–505.

Ben-Akiva, M., Bolduc, D., & Bradley, M. (1993). Estimation of travel mode choice models with randomly distributed value of time. *Transportation Research Record*, *1413*, 88–97.

Bliemer, M. C. J., & Rose, J. M. (2009). Designing stated choice experiments: The state of the art. In R. Kitamura, T. Yoshi, & T. Yamamoto (Eds.), *The expanding sphere of travel behaviour research, selected papers from the 11th International Conference on Travel Behaviour Research*, pp. 499–538. Chapter 25.

Brown, T. C., & Gregory, R. (1999). Why the WTA-WTP disparity matters. *Ecological Economics*, *28*, 323–335.

Brownstone, D., Bunch, D. S., & Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B*, *34*, 315–338.

Danielis, R., Marcucci, E., & Rotaris, L. (2005). Logistics managers' stated preferences for freight service attributes. *Transportation Research Part E*, *41*(3), 201–215.

De Borger, B., & Fosgerau, M. (2008). The trade-off between money and travel time: a test of the theory of reference-dependent preferences. *Journal of Urban Economics*, *64*, 101–115.

Fowkes, A. S. (2007). The design and interpretation of freight stated preference experiments seeking to elicit behavioural valuations of journey attributes. *Transportation Research Part B*, *41*(9), 966–980.

Graves, P. E. (2009a). *If the large WTA-WTP gap for public goods is real (and there are good reasons to think so) conventional welfare measures are simply incorrect*. The Selected Works of PHILIP E GRAVES. Retrieved from http://works.bepress.com/philip_graves/35

Graves, P. E. (2009b). *The WTA-WTP gap and welfare measures for public goods*. The Selected Works of PHILIP E GRAVES. Retrieved from http://works.bepress.com/philip_graves/41

Hanemann, M. (1991). Willingness to pay and willingness to accept: How much can they differ? *The American Economic Review*, *81*(3), 635–647.

Hensher, D. A. (2001). Measurement of the valuation of travel time savings. *Journal of Transport Economics and Policy*, *35*(1), 71–98.

Hensher, D. A. (2008). Empirical approaches to combining revealed and stated preference data: Some recent developments with reference to urban mode choice. *Research in transportation Economics*, *23*, 23–29.

Hensher, D. A., & Greene, W. H. (2003). Mixed logit models: state of practice. *Transportation*, *30*(2), 133–176.

Hess, S., Rose, J. M., & Hensher, D. A. (2008). Asymmetric preference formation in willingness to pay estimates in discrete choice models. *Transportation Research Part E*, *44*(5), 847–863.

Horowitz, J., & McConnell, K. E. (2002). A Review of WTA-WTP Studies. *Journal of Environmental Economics and Management*, *44*, 426–447.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: an analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Mackie, P. J., Jara-Diaz, S., & Fowkes, A. S. (2001). The value of travel time savings in evaluation. *Transportation Research Part E*, *37*(2-3), 91–106.

Maggi, R., & Rudel, R. (2008). The value of quality attributes in freight transport: Evidence from an SP-experiment in Switzerland. In M. E. Ben-Akiva, H. Meersman & E. van de Voorde (Eds.), *Recent developments in transport modelling*. Bingley, UK: Emerald Group Publishing Limited.

Masiero, L., & Hensher, D. A. (2010). Analyzing loss aversion and diminishing sensitivity in a freight transport stated choice experiment. *Transportation Research Part A*, *44*(5), 349–358.

Masiero, L., & Maggi, R. (2012). Estimation of indirect cost and evaluation of protective measures for infrastructure vulnerability: A case study on the transalpine transport corridor. *Transport Policy*, *20*, 13–21.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics*. New York, NY: Academic Press.

Randall, A., & Stoll, J. R. (1980). Consumer's surplus in commodity space. *The American Economic Review*, *70*(3), 449–455.

Revelt, D., & Train, K. E. (2000). *Customer-specific taste parameters and mixed logit*. WP E00-274. Department of Economics, University of California, Berkeley, CA.

Rose, J. M., & Masiero, L. (2010). A comparison of the impacts of aspects of prospect theory on WTP/WTA estimated in preference and WTP/WTA space. *European Journal of Transport and Infrastructure Research*, *10*(4), 330–346.

Rose, J. M., Bliemer, M. C. J., Hensher, D. A., & Collins, A. C. (2008). Designing efficient stated choice experiments involving respondent based reference alternatives. *Transportation Research Part B*, *42*(4), 395–406.

Train, K. E., McFadden, D., & Ben-Akiva, M. (1987). The demand for local telephone service: A fully discrete model of residential calling patterns and service choice. *Rand Journal of Economics*, *18*, 109–123.

Train, K. E., & Weeks, M. (2005). Discrete choice models in preference space and willingness-to-pay space. In R. Scarpa & A. Alberini (Eds.), *Application of simulation methods in environmental and resource economics*. Dordrecht: Springer Publisher.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, *106*, 1039–1061.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

Zamparini, L., & Reggiani, A. (2007). Freight transport and the value of travel time savings: A meta-analysis of empirical studies. *Transport Reviews*, *27*(5), 621–636.

Chapter 10

Endogenous Shipment Size in Freight Mode Choice Models: Theory and Empirical Testing

François Combes, Kees Ruijgrok and Lóri Tavasszy

Abstract

One of the research questions in freight demand modelling is how to improve the representation of logistics choices underlying the decision of the mode of transport. On theoretical grounds it is already widely accepted that the choice of shipment size influences the choice of mode. Mostly due to a lack of data, however, there have been few attempts to empirically test the quality of this relationship. In this chapter we propose a model for the choice of mode of transport in which the costs of transport follow the logic of economics of shipment size choice. We specify a discrete choice model where the generalized cost depends on transport distances, mode abstract values of time and continuous shipment sizes. The model is estimated on observations of about 10,000 individual shipments from the French ECHO survey. The results show that mode choice is strongly consistent with the economics of shipment size choice. Possibilities for the further extension and application of the model lie in simultaneous mode and shipment size choice, in network based choice where modes interact and in geographically specified models.

Keywords: Freight transport demand modelling; mode choice; shipment size; discrete choice models; EOQ

10.1. Introduction

Recent advances in the state of the art in freight demand modelling have been focusing more and more on the question of how to introduce principles of logistics

Freight Transport Modelling

ISBN: 978-1-78190-285-1

into these models (Chow, Yang, & Regan, 2010; De Jong, Vierth, Tavasszy, & Ben-Akiva, 2012; Tavasszy, Ruijgrok, & Davydenko, 2012). Through a more accurate representation of the behaviour of shippers and carriers, it is believed that the models would become more broadly applicable and more accurate. Research is evolving in different directions, studying new choice problems that were hitherto not modelled explicitly, linking choices into composite ones and refining the mechanisms of standard choice problems that have already been represented for a long time in freight models. In the latter category, one of the research problems that has been studied since the late 1980s is the choice of mode in freight transport (see a more detailed review of this stream of work further in the chapter). Our chapter deals with the relationship between mode choice and the size of shipments, one of the characteristics of the goods in transit that is the outcome of a broader logistics optimization process. Only recently have attempts been made to develop demand models that include an endogenous representation of shipment size, and that were developed from a rigorous theoretical background. This chapter serves two objectives. Firstly, it introduces the importance of modelling shipment size and mode choice simultaneously, from the broad and practical perspective of a modelling framework that minimizes total logistics costs. Secondly, it gives a formal treatment of the relationship and empirical proof from French shipper survey data. The sections of the chapter follow this structure.

10.2. Shipment Size as Central Variable in Logistics Optimization

We assume that rational supply chain managers try to minimize their logistic costs while maintaining a certain service level that is required for their customers. These service levels are very much correlated with the value density of the products involved (Christopher, 1992; Simchi-Levi, Kaminsky, & Simchi-Levi, 2000) so the supply chain optimization problem can be reduced to a generalized cost minimization problem per product type. This optimization problem involves the choice of production and storage locations, the frequency of replenishment shipments, the choice of mode and the inventory policy used. We propose to use generalized cost functions that take into account the most likely choice for mode of transport.

The definition of generalized logistic costs we propose that takes into account the drivers mentioned above is:

$$C_i = I_i + H_i + T_i, i = 1, \ldots, I \tag{10.1}$$

where, C_i — the generalized cost for product *i* out of the set of all products — is the sum of I_i (Inventory), H_i (Handling) and T_i (Transportation) costs.

Important drivers of a realistic generalized cost function are:

- pipeline costs (including transport and inventory costs for products in the pipeline),

- product value density ($\boldsymbol{vd} = \boldsymbol{V}/\boldsymbol{D}$ in \$/m^3), where $\boldsymbol{V}$ stands for value and $\boldsymbol{D}$ for total demand measured in m^3: for products with a higher value density inventory costs are more important than for other products,
- order size $\boldsymbol{O}$: the larger the order size, the larger the shipment size $\boldsymbol{s}$, the lower transport and handling costs per unit can be. This is due to the utilization factor of vehicles and economies of scale and scope that lead to lower logistic costs per unit for larger vehicles,
- shipping frequency $\boldsymbol{f}$: higher frequencies lead to lower waiting costs, and therefore more reliable lead times, lower safety stocks and also to network synergies: frequency and shipment size are inversely related,
- variance in lead time $\sigma^{\mathbf{supply}}$ and demand volumes $\sigma^{\mathbf{demand}}$, the higher the volatility, the higher the demand that ask for responsive, and therefore more expensive services that can meet demand, or that require carrying larger amounts of stocks: if the demand is stable, goods can be forecasted accurately and can be shipped before the actual demand is realized and using a cheap mode of transport.

Inventories are made up by cycle, safety and pipeline inventories. Cycle inventories are dependent on order size. Safety inventories depend on frequency and variability of demand and supply. Pipeline inventory costs include deprecation of the value of goods during transport.

$$I_i = I_i^{\text{cycle}} + I_i^{\text{safety}} + I_i^{\text{pipeline}} \tag{10.2}$$

$$I_i^{\text{cycle}} = f(\text{order size } \boldsymbol{O}) \tag{10.2a}$$

$$I_i^{\text{safety}} = f(\text{frequency} \boldsymbol{f}, \text{order size } \boldsymbol{O}, \text{demand variance } \boldsymbol{\sigma}^{\text{demand}}, \text{lead time variance } \boldsymbol{\sigma}^{\text{supply}}) \tag{10.2b}$$

$$I_i^{\text{pipeline}} = f(\boldsymbol{t}, \boldsymbol{r}, \boldsymbol{V}) \tag{10.2c}$$

where $\boldsymbol{t}$ = Transport time, $\boldsymbol{r}$ = interest rate including risk for obsolescence of unsold products, $\boldsymbol{V}$ = value of goods transported.

$$H_i = \text{handling costs} = f(\boldsymbol{pd}) \tag{10.3}$$

where $\boldsymbol{pd}$ = packaging density expressed in number of colli per m^3. The number of colli (or packages per m^3 is an indication for the amount of handling activities necessary to handle this amount. It refers to the amount of packages that are handled individually, so if they are loaded on a pallet and the pallet is handled by a forklift, the loaded pallet is the handling unit and the package density is redefined accordingly

The higher $\boldsymbol{pd}$, the more individual activities are needed to handle the product and therefore the higher the need for automation in this process in order to reduce manpower costs.

$$T_i = \text{transport costs} = f(\boldsymbol{d}, \boldsymbol{O}, \boldsymbol{f}, \boldsymbol{vd}, \boldsymbol{m}, \boldsymbol{sp}, \boldsymbol{b}) \qquad (10.4)$$

where $\boldsymbol{d}$ = distance, $\boldsymbol{f}$ = frequency = $\boldsymbol{V}/\boldsymbol{O}$, O = *order size*, $\boldsymbol{vd}$ = value density, $\boldsymbol{m}$ = mode, $\boldsymbol{sp}$ = speed (depending on the mode of transport), $\boldsymbol{b}$ = reliability of mode.

Note that the components of the generalized cost function are not independent; an important linkage is through shipment size $\boldsymbol{s}$. The shipment size is directly related to the order size: given the quantity of goods ordered per period the number of times orders are placed determine the order size $\boldsymbol{O}$. The Economic Order Quantity (EOQ) model produces the optimal order size $\boldsymbol{O^*}$ that minimizes the sum of order (transport and handling) costs and cycle inventory costs. Through this linkage, the shipment size variable becomes instrumental in minimizing total logistic costs. As transport costs also depend on shipment size, the relationship becomes circular and the EOQ becomes an integral part of a mode choice model.

Heavier transport modes are characterized by a higher b_m and a lower c_m, and are thus competitive for larger commodity flows, all other things equal, provided their travel times are not comparatively too large. So transport costs are highly sensitive to the type of mode that is used and this also determines the maximum load that can be transported per shipment. In Figure 10.1 the shipment sizes and transport costs of some modes of transport are visualized. From this picture it becomes clear that huge differences exist between the respective modes of transport, both in shipment size and in average transport charges.

In the remainder of this chapter we discuss the more formal treatment of the relations between EOQ and mode choice and describe an empirical test that confirms the validity of a simultaneous model.

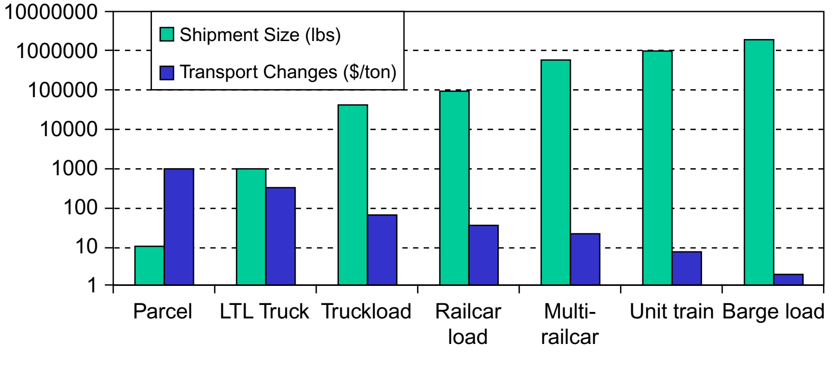

Figure 10.1: Shipment size and tariffs for different modes of transport. *Source*: Rodrigue (2006).

10.3. Shipment Size in Freight Transport Demand Modelling

The formal relationship between shipment size and the decision of transport mode was described mathematically a long time ago with the seminal paper of Baumol and Vinod (1970), building on the classic EOQ model (Harris, 1913; Wilson, 1934). Baumol and Vinod characterize the behaviour of a shipper with a total logistic cost function, and show how shipment size, mode choice and safety stock are determined jointly to minimize this cost function. This function depends on the available transport alternatives, more precisely of their maximum allowable shipment sizes, their costs, travel times and travel time reliabilities. It also depends on the logistic context of the shipper-receiver's supply chain: the commodity's capital opportunity cost and depreciation cost, warehousing costs, the variability of demand and the sensitivity of customers to delays in delivery, or stock outs.

Other theoretical papers discussed these topics: Langley (1980) and Hall (1985) examined the specific role of freight tariffs in the mode choice-shipment size decision. McFadden, Winston, and Boersch-Supan (1985) derived a dynamic inventory-theoretical model with a similar objective. McCann (2004) extended the model to include the decision of ship size. De Jong and Ben-Akiva (2007) describe a still ongoing effort to develop a specialized freight transport demand model where shipment size is explicit and endogenous. This model is particularly interesting, as its aim is to account for most of the developments discussed in this section.

10.3.1. Elements of Inventory Theory: The EOQ Model and its Extension to Mode Choice

The formal principles of the EOQ model, and of its extension to mode choice, are now summarized. Consider a firm sending a regular commodity flow of constant rate Q, from a given location to another by a given transport mode, in shipments of identical size s. Each shipment is dispatched as soon as possible, so that the average origin inventory is $s/2$ (in other words, safety stock is not considered). The average destination inventory is assumed to be the same, so that the total inventory level is s, on average. Now denote a the commodity value of time, that is the willingness of the shipper to pay for a reduction of the time that elapses between the moment a unit of commodity is produced and the moment it is sold, or used. The cost to transport a shipment with a given transport mode m is assumed to consist of a fixed cost b_m, independent of shipment size, and a variable cost $c_m s$, proportional to shipment size, where c_m is expected to increase with the distance between the origin and the destination. The travel time is denoted by t. Given the fact that Q/s shipments are dispatched per period of time, it is now possible to introduce the following total logistic cost function, denoted by g_m:

$$g_m(s) = as + (at + c_m)Q + \frac{Qb_m}{s} \tag{10.5}$$

Minimizing this function yields the well-known EOQ equation:

$$s^* = \sqrt{\frac{Qb_m}{a}} \tag{10.6}$$

Then, if s^* is replaced in Eq. (10.1), one obtains a total logistic cost for each mode (expressed on a per ton basis in the equation below), where shipment size is endogenous:

$$\frac{g_m(s^*)}{Q} = 2\sqrt{\frac{ab_m}{Q}} + at_m + c_m \tag{10.7}$$

The total logistic cost function for mode m is a non linear function of a, Q, and of the cost and quality of service characteristics of the transport mode m. Heavier transport modes are characterized by a higher b_m and a lower c_m, and are thus competitive for larger commodity flows, all other things equal, provided their travel times are not comparatively too large.

Clearly, inventory theory and the EOQ model in particular constitute an appealing theoretical basis to design freight transport demand models. By making shipment size an endogenous variable — either implicitly or explicitly, and more generally by representing the logistic contexts of freight transport operations, they are more relevant than the classic modelling approaches, which only consider costs and travel times.

10.3.2. Empirical Validation of the EOQ Model

The difficulty in introducing such concepts in freight transport demand models is mainly a matter of data availability. Indeed, one of the important explanatory variables in both Eqs. (10.2) and (10.3) is the commodity flow rate Q. This flow rate is usually not observed in shipment surveys. Despite this fact, there have been a number of empirical studies on the relationship between mode choice and shipment size. Most of them had to rely on exploratory specifications only remotely related to inventory theory (Abdelwahab & Sargious, 1992; Holguín-Veras, 2002; McFadden et al., 1985). They all confirm that there is a close relationship between shipment size and mode choice, but not in ways that can be easily used to design freight transport demand models. Other studies are actually based on inventory theory (de Jong & Johnson, 2009; Windisch, de Jong, van Nes, & Hoogendoorn, 2010). They associate a total logistic cost function to each mode and/or shipment size alternative; however, they are limited by the lack of adequate data, especially regarding the commodity flow rate Q.

One database allows to bridge this gap: the French ECHO shipper survey, conducted in 2004 and 2005, has observations on about 10,000 shipments of very different types, origins and destinations (Guilbault, 2008). It notably describes the logistics contexts within which these shipments are sent, with variables such as whether the shipments are part of a scheduled programme or sent on demand, the average inventory level at the shipper's, etc. Among other variables, the total commodity flow between the shipper of a given shipment and its receiver is observed. With such information, it is possible to test whether the equations above make sense empirically or not.

For example, linear estimation techniques are easily applied to check Eq. (10.6). Once the logarithm of both sides is taken, it becomes:

$$\ln s* = \frac{1}{2}\ln b_m + \frac{1}{2}\ln Q - \frac{1}{2}\ln a \tag{10.8}$$

Of course, a cannot be observed directly. In this case, it can be replaced by a_{dens}, the value density of the shipment. If a and a_{dens} are proportional, then the factor between them should be captured in the constant. The shipment size independent component of transport freight tariffs b_m is also not observed, and as a result captured in the constant.

Eq. (10.8) was estimated for each mode available in the ECHO database. The results are summarized in Table 10.1:

These results are strongly consistent with the theory underlying the EOQ model. Even if it does not perform as good for heavy land-based modes (especially rail and inland waterway) as for light modes, and if the coefficients are not always close to their theoretical values, the fit is rather good and the coefficients' signs are as expected. A more precise discussion of the database, methodology, and results of the estimation of the EOQ model is available in Combes (2012). The results presented in Table 10.1 are drawn from Lloret-Batlle (2011).

10.3.3. *Estimation of an EOQ-based Mode Choice Model*

As illustrated in the previous part, it is not difficult to estimate an EOQ-like shipment size model with the ECHO database. As a matter of fact, it is equally straightforward

Table 10.1: Estimation of the EOQ model per mode

Mode	**Intercept**	**ln Q**	**ln a_{dens}**	**adj. R^2**	***N***
Private carrier	1.838***	0.457***	−0.525***	0.657	934
Common carrier	1.777***	0.506***	−0.475***	0.771	3603
Rail	2.276*	0.460***	−0.189*	0.287	115
Combined transport	0.822	0.473***	−0.241*	0.627	74
Inland waterway	5.627**	0.110	−0.166	0.238	19
Sea	2.721***	0.412***	−0.370***	0.637	449
Air	−0.957*	0.484***	−0.217***	0.449	513

p-values are symbolized as follows: ***significant at 0.001 level; **significant at 0.01 level; *significant at 0.05 level; other estimates not significant at a level more than 0.05.

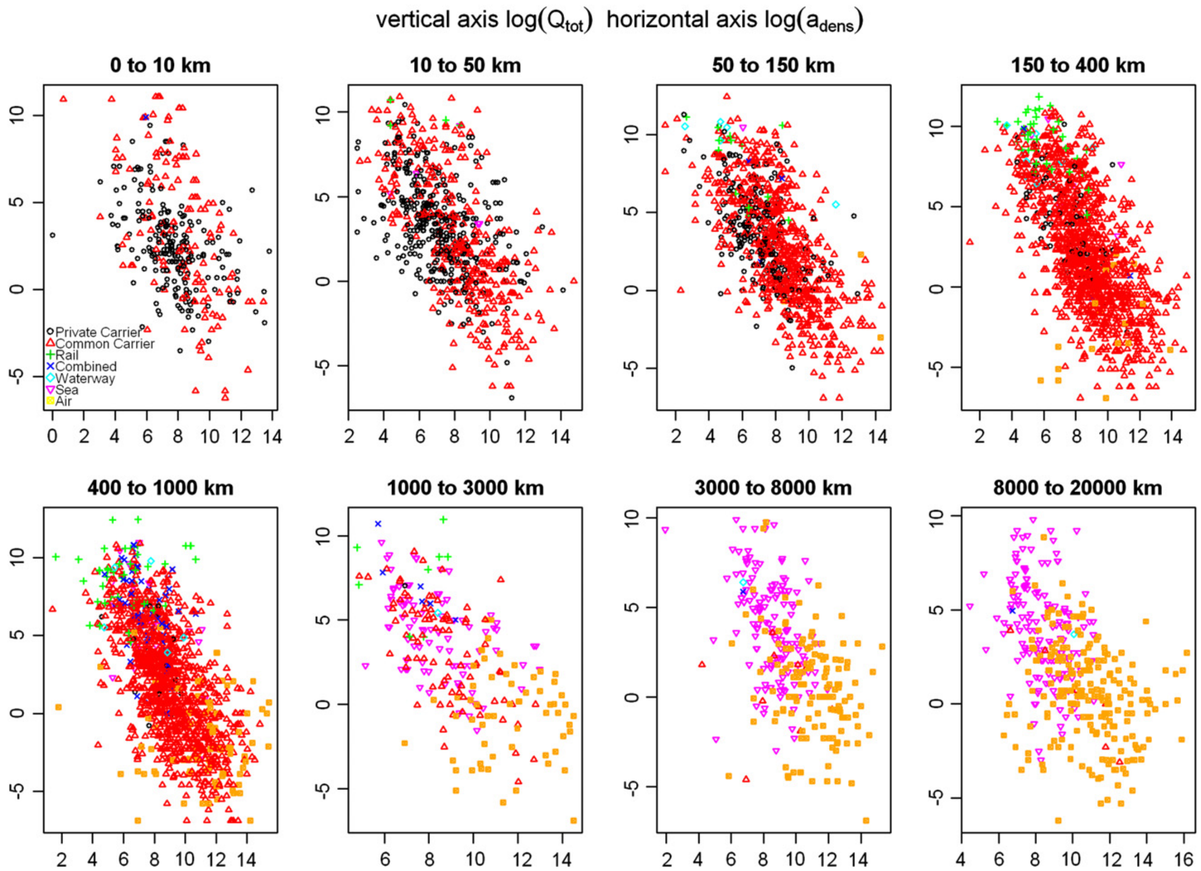

Figure 10.2: Observed mode choices in the ECHO database.

to estimate a mode-choice modelled inspired from Eq. (10.3). What follows is a short summary of Lloret-Batlle and Combes (2013) is presented.

First of all, in order to illustrate qualitatively the relevance of the EOQ model, consider Figure 10.2. Per band of distance, it illustrates how mode choice depends on both the commodity flow, represented along the x-axis with a logarithmic scale, and the value density, represented along the y-axis with a logarithmic scale as well. For example, on medium distances (50–1000 km), heavy modes are restricted to large commodity flows and low value densities; for larger distances (above 1000 km), where sea and air transport are dominant, there is a clear specialization of these modes: large flows and low value densities for sea transport, small flows and high value densities for air transport.

In order to confirm quantitatively this intuition, logit models inspired from Eq. (10.3) were estimated. Each mode was characterized by a utility function with a specification similar to:

$$V_m = ASC_m + \beta_m \sqrt{\frac{a_{\text{dens}}}{Q_{\text{tot}}}} + \gamma_m d + \varepsilon_m a_{\text{dens}} d + \sum_{\text{dummies}} \delta_{im} X_i \tag{10.9}$$

Note that these models are entirely demand oriented: there is no description of the transport alternatives. To overcome the lack of a travel time variable, the d and $a_{\text{dens}}d$ terms are introduced, in order to capture the travel cost and travel time components respectively. The dummies concern commodity type variables (fragile, refrigerated, hazardous materials etc.) as well as the availability of some infrastructures at the origin of the transport operations (active rail terminals or active inland waterway terminals).

A number of model variants were estimated with Biogeme (Bierlaire, 2003). The detailed results will not be reported here. The main conclusions are that first, even without transport supply data, the model is relatively efficient at predicting mode choice. Secondly, all the variables in Eq. (10.9) are significant, and their coefficients have the expected signs and hierarchies. In particular, the $(a_{\text{dens}}/Q_{\text{tot}})^{1/2}$ term improves significantly the model likelihood, while the β_m are consistent with the EOQ model (a heavier mode is expected to have a lower β_m, which is indeed observed).

This approach should be considered more as proof of concept rather than as a fully fledged model. Nevertheless, it confirms the potential of inventory theory for freight modelling and the importance of having adequate data to design, calibrate and estimate such models.

10.4. General Discussion and Further Research Directions

The theory and examples discussed here show the potential of inventory theory in freight transport modelling. The EOQ model and its extensions are theoretically sound and empirically valid, as illustrated by the above estimations.

Efforts should be made to improve these models, theoretically and empirically. Firstly, the methodologies described in this section do not account properly for the capacity constraint of vehicles. To overcome this issue, two possible directions can be considered: (a) to build a model of joint mode choice and shipment size and (b) to generalize Eq. (10.9) so as to make it consistent with capacity constraints. A second possible improvement concerns the distinction between order size and shipment size. Current models depart from the assumption that these can be treated as equivalents. Shipper surveys are now not unambiguous about what is exactly measured. Models and data bases should be refined further and aligned in terms of their definitions.

Models that are sensitive to the change of shipment sizes also offer a lot of opportunity to address the issue of the choice of logistic networks by shippers: should production-consumption flows be direct or go through break-bulk platforms or warehouses? How many such platforms should there be? The underlying problem of the choice of mode and shipment size, and its relationship with the commodity flow rate variable has an important impact for such decisions. Further investigation is needed to develop such integrative models.

Finally, these models should be implemented in the frame of specialized, that is geographically specified, freight transport demand modelling. This involves significant improvements of freight modelling techniques; the four stages of generation, distribution, mode choice and assignment have to be adapted to account for additional variables and cost structures. It also involves changes in the way freight transport systems are observed and measured.

References

Abdelwahab, W. M., & Sargious, M. A. (1992). Modeling the demand for freight transport: New approach. *Journal of Transport Economics and Policy*, *26*(1), 49–72.

Baumol, W. J., & Vinod, H. D. (1970). An inventory theoretic model of freight transport demand. *Management Science*, *16*(7), 413–421.

Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland.

Chow, J. Y. J., Yang, C. H. Y., & Regan, A. C. (2010). State-of-the art of freight forecast modeling: Lessons learned and the road ahead. *Transportation*, *37*(6), 1011–1030.

Christopher, M. (1992). *Logistics and supply chain management*. London: Financial Times Pitman Publishing.

Combes, F. (2012). An empirical evaluation of the EOQ model of choice of shipment size in freight transport. *Transportation Research Board 91th Annual Meeting*, Washington, DC.

De Jong, G., & Ben-Akiva, M. (2007). A micro-simulation model of shipment size and transport chain choice. *Transportation Research Part B*, *41*, 950–965.

De Jong, G., & Johnson, D. (2009). Discrete mode and discrete or continuous shipment size choice in freight transport in Sweden. *Proceedings of the European Transport Conference*, Leeuwenhorst, The Netherlands.

De Jong, G., Vierth, I., Tavasszy, L. A., & Ben-Akiva, M. (2012). Recent developments in national and international freight transport models. Transportation: DOI 10.1007/s11116-012-9422-9

Guilbault, M. (2008). Enquête ECHO envoi - chargeurs - opérateurs de transport. Résultats de référence. Synthèse no 58, INRETS.

Hall, R. W. (1985). Dependence between shipment size and mode in freight transportation. *Transportation Science*, *19*(4), 436–444.

Harris, F. W. (1913). How many parts to make at once. *Factory, the Magazine of Management*, *10*(2), 135–152.

Holguín-Veras, J. (2002). Revealed preference analysis of commercial vehicle choice process. *Journal of Transportation Engineering*, *128*(4), 336–346.

Langley, C. J. (1980). The inclusion of transportation costs in inventory models: Some considerations. *Journal of Business Logistics*, *2*(1), 106–125.

Lloret-Batlle, R. (2011). Modal choice and shipment size models based on the French ECHO database, Universitat Politècnica de Catalunya, and Ecole des Ponts ParisTech, Novembre 2011, MSc thesis.

Lloret-Batlle, R., & Combes, F. (2013). Estimation of an inventory-theoretic model of mode choice in freight transport. In *Transportation Research Board 92th Annual Meeting*, Washington, DC.

McCann, P. (2004). A proof of the relationship between optimal vehicle size, haulage length and the structure of distance-transport costs. *Transportation Research Part A*, *35*, 671–693.

McFadden, D., Winston, C., & Boersch-Supan, A. (1985). Joint estimation of freight transportation decisions under nonrandom sampling. In A. F. Daughety (Ed.), *Analytical studies in transport economics* (pp. 138–157). Cambridge, UK: Cambridge University Press.

Rodrigue, J. P. (2006). Transportation and the geographical and functional integration of global production networks. *Growth and Change*, *37*(4), 510–525.

Simchi-Levi, D., Kaminsky, P., & Simchi-Levi, E. (2000). *Designing and managing the supply chain*. McGrawHill Higher Education.

Tavasszy, L. A., Ruijgrok, C. J., & Davydenko, I. (2012). Incorporating logistics in freight transport demand models: State of the art and research opportunities. *Transport Reviews*, *32*(2), 203–219.

Wilson, R. H. (1934). A scientific routine for cost control. *Harvard Business Review*, *13*, 116–128.

Windisch, E., de Jong, G. C., van Nes, R., & Hoogendoorn, S. P. (2010). A disaggregate freight transport model of transport chain and shipment size choice. *Proceedings of the European Transport Conference*, Glasgow, Scotland, UK.

Chapter 11

Supply-Chain Risk Analysis with Extended Freight Transportation Models

Hanno Friedrich and Andreas Balster

Abstract

The analysis and management of supply-chain risks has become more important in a world with globally integrated production networks in which extreme events seem to occur more frequently. But it seems that transparency, in particular for the overall economic, production and logistics system, is missing. Research is needed that generates knowledge on the interrelationships and impact of risks within the system.

This chapter discusses the possibility of using approaches from freight transportation analysis to analyze consequences of supply-chain risks on overall logistics and freight transportation. A literature review shows that there is a research gap for these models especially for the analysis of short and medium-term impacts of risks (days to weeks). An overview of possible reactions of economic actors to risks and challenges of modeling these risks recommends that such models should be explanatory, map time explicitly, and probably be specific to sectors and countries.

At the end of the chapter a research plan to develop such a model for the analysis of the impact of risks on food distribution in Germany is given. This model uses elements of SYNTRADE, a model originating from the area of freight transportation demand analysis.

Keywords: Risk analysis; risk management; freight transport modeling; supply chain; food supply; logistics network

Freight Transport Modelling

ISBN: 978-1-78190-285-1

11.1. Introduction

Risks, as potential losses with a low probability of occurrence, have increased in supply chains in the last decades. This is the result of expanding global production networks and more and more distributed value creation, in which low probabilities add up to a substantial level. Furthermore, it seems that more extreme events occur, increasing the possibility of a supply-chain disruption. Extreme events include extreme natural phenomena, technical failures, epidemics, along with anthropogenic events such as strikes, terrorism, and war. The Centre for Research on the Epidemiology of Disasters (CRED) found in 2006 that economic damages due to natural disasters almost doubled between 1987 and 2006. Even though this might also be the result of more accurate recording, it is clear that production networks and supply chains today are more complex and thus more vulnerable to extreme events if no precautions are taken.

Media reports relating extreme events and the authors' own interviews with logisticians show that companies have little or no information on actors two or three steps up or down their supply chain. A detailed overview of the whole supply chain does not exist. This causes uncertainties for companies in the case of extreme events such as the ash cloud that followed the volcanic eruptions on Iceland in 2010 or the nuclear accident in Fukushima in 2011. German companies, for example, were not sure to what extent they would be affected by these events. Governments face similar issues. Like companies, they do not know how robust or resilient supply chains are because they lack detailed knowledge on the connections between companies and their globally distributed suppliers. Hence, taking the right measures to support preparation or reaction is difficult.

To fill this information gap, research with the goal of bringing transparency into these complex dependencies is required. In this context, the transportation process is especially interesting, since it is a potential cause of disruptions, for example through damage to transportation infrastructure. At the same time, the transportation process also provides possible ways of reacting to disruptions, for example air transport for emergency deliveries.

The development of analytical, empirically calibrated models can be a helpful method to fill this gap. They can provide answers for questions such as: How vulnerable, robust, and resilient are the existing logistics networks? What are the dependencies between industries in different countries? What are potential consequences of extreme events for logistics and production systems?

A number of existing modeling approaches already address these questions in part. There are models that encompass the overall economy, such as input-output (IO) models or spatial computable general equilibrium (SCGE) models. These models give insight into the connections between sectors and countries on an aggregated level. A second class of models, from logistics research, focuses on individual supply chains or logistics networks. In this area, optimization models and simulation models of logistics processes are dominant, while descriptive models are rare. However, this research provides insight into how actors should behave in different situations. Finally, there are models from the area of transportation system analysis that investigate the impact of infrastructure disruptions due to natural disasters on overall traffic flow. Passenger transportation accounts for the largest share of these models. Only a small amount of research has been done on freight transportation.

Existing models for freight transportation focus on other purposes than supply-chain risk analysis. However, models developed recently attempt to include more and more logistics details. This could also be interesting for the problem at hand. Parts of different types of models can be used to analyze the consequences of supply-chain risks for the overall economy.

This chapter discusses the possibility of using approaches from freight transportation analysis to analyze consequences of supply-chain risks on overall logistics and freight transportation. For this purpose, a literature review will be conducted in Section 11.2, collecting relevant approaches and models. Then, categories of risk management measures that may be taken by companies will be described in Section 11.3. This will give an idea of the necessary scope of a model that is sensitive to behavior. Section 11.4 will discuss additional challenges and simple findings for modeling supply-chain risks. A research plan for the development of a simulation model analyzing shortfalls in food distribution will be given as a first example in Section 11.5. In the final section, conclusions will be drawn.

11.2. Literature Review

In this section, relevant literature from different areas will be presented briefly. First, basic definitions and risk classifications will be discussed. Then, empirical studies on supply-chain risks will be reviewed. Finally, models on the level of supply chains or logistics networks, as well as on the level of overall economic activity or the transportation system will be discussed.

11.2.1. Basic Definitions of Supply-Chain Risk Management

Schmidt and Schaich (2011) define risk as ‘‘an eventuality of damage or an absence of an advantage, which can occur with low or unknown probability.’’ This definition emphasizes two characteristics of risks: their downward potential and their low or unknown probability. Such risks can be classified by their cause, in our case, as internal or external, in relation to the supply chain (Christopher & Peck, 2004). ‘‘Internal to the supply chain’’ can refer to risks caused by the firm itself as well as risks caused by partners within the supply chain. Internal risks can be broken down further into the categories supply, demand, process, and control/behavioral risks, while external risks can be categorized as political, social, economic, and technological risks (Christopher & Peck, 2004; Tang & Tomlin, 2008).

Knemeyer, Zinn, and Eroglu (2009), distinguish risks by the level of probability and the level of impact. According to the chosen definition of risk, we only consider low-probability risks. The definition of vulnerability (Sheffi, 2006) goes one step further, stating that a process or company is vulnerable to a certain event if the possibility of an effect is high and the impact itself is significant. Take, for instance, the risk of an extreme winter storm: the probability of a winter storm taking place is low, but, if a winter storm occurs, the probability of impact on critical infrastructure and the resulting damage to it is high. Studying such extreme events, Sheffi (2009) differentiates risks by their causes: natural disasters, accidents, negligence, and intentional disruptions.

Supply-chain risk management attempts to analyze risks related to the supply chain and minimize the negative impact of these risks. Norman and Jansson (2004) describe a three-step risk management process: risk analysis, risk assessment, and risk management. For the management of risks, measures must be planned, implemented, and controlled (Pfohl, 2008).

Such measures can either be proactive or reactive (Tomlin, 2006). Proactive measures are taken before the event occurs. This can mean that something is avoided or that the potential impact is reduced (Köhler, 2011). This can also be referred to as prevention. Reactive measures address the consequences by transferring the risk to a third party or by handling the consequences. This can either mean simply bearing the consequences (i.e., absorbing them) or limiting them (Köhler, 2011).

In the context of a logistics network or a supply chain, the two concepts of resilience and robustness of a system are also of interest. Christopher and Peck (2004, p. 4) define resilience as the ability of a system to return to its original state or move to a new, more desirable state after being disturbed. Monti (2011) declares a system to be robust when it can continue functioning in the presence of internal or external challenges without fundamental changes to the original system. Both can be transferred to logistics processes.

11.2.2. Empirical Studies on Supply-Chain Risks

Two areas of empirical studies can be differentiated: studies focusing on individual supply chains or networks and studies analyzing risks to the overall economic or transportation system. In the first categories, different risks are often analyzed for a single network or supply chain. There are case studies and surveys. The work of Sheffi (2006), which gives an overview of risks connected to extreme events, is a collection of case studies. Another example is the work of Tuncel and Alpan (2010), which demonstrates the usefulness of Petri nets for the analysis of supply-chain risks using an industrial case study. Petri nets are a graphical and mathematical modeling tool used to describe and study distributed, nondeterministic and stochastic systems. They are directed and weighted bipartite graphs consisting of two sets of nodes, referred to as "transitions" and "places," where marks can shift from a place to a transition or the reverse (Murata, 1989; Tzes, Kim, & McShane, 1996). The work of Wagner and Bode (2006) is an example of a survey. They conducted a survey of 760 executives of firms operating in Germany and showed that their vulnerability depends on the number of customers and suppliers. Another example gives Skelton (2007), whose study is based upon work with a number of different companies. He describes how companies can use their knowledge from audits and experience as well as existing relationships with logistics service providers to design flexible supply chains.

The second category tends to concentrate more on specific events that are either observed or outlined as potential scenarios. For example, McKinnon (2006) analyzes the potential impact of an organized strike of truck drivers in the United Kingdom. The consequences for the economy and consumers are described for each day that the strike continues. The analysis is based on available data, literature, and experience from past events. Another example is the report of NEA on the economic damage of a ship accident on the River Rhine in Germany in 2011 (NEA, 2011).

11.2.3. Models

In modeling, there is already significant existing research in different areas as well. As for empirical research, there are models that focus on analyzing risks to individual supply chains or networks and there are models that analyze the consequences of extreme events on the overall transportation system or on the overall economy. There is an interesting recent trend in general freight-transportation modeling to include logistics details. These models could be used as a basis for the development of models that analyze the risks to overall freight transportation and logistics.

The first category of this research, concentrating on individual supply chains or networks, includes models from the area of operations research, such as models determining the optimal level of safety stock or the optimal design of networks. One example is the analysis of Simchi-Levi (2012) on the optimal level of flexibility of production networks in the context of vulnerability.

Other models do not try to optimize the system, but rather attempt to explain how it functions. Such models are used to analyze the likely impact of events on supply chains. Take, for example, the work of Wilson (2007) or Brock, Zhang, Hayden, Matteis, and Gross (2012). Wilson (2007) developed a system dynamics model to analyze the impact of transportation interruptions on supply-chain performance. Brock et al. (2012) analyze the consequences for a whole logistics network, considering changes in traffic situations and logistics tour planning. Other researchers, such as Holguín-Veras, Jaller, van Wassenhove, Pérez, and Wachtendorf (2012), concentrate on humanitarian logistics. They identify the difference between commercial and humanitarian logistics, along with research gaps that need to be filled to enhance the efficiency of short-term and long-term disaster responses.

There are many examples of models that analyze the impact of extreme events on transportation networks — these form the second category. In most cases, they analyze the direct effects on traffic flow or changes in flow distribution via different routes. The analysis of the consequences for transport demand is rare and, to our knowledge, only exists for passenger transport. For example, the analysis of impacts on traffic flows with microscopic simulation models is done in different studies: Boden and Weger (2008) simulate traffic interruptions to optimize the evacuation of the population during an Elbe flooding in Dresden, Germany, Tamminga, Tu, Daamen, and Hoogendoorn (2011) use microscopic traffic simulation to analyze the impacts of various departure time spans on evacuation time and network performance and Burgholzer, Bauer, Posset, and Jammernegg (2012) focus on the analysis of intermodal transport networks by using a traffic microsimulation. Examples for more aggregated analysis of traffic flows are the works of Murray-Tuite and Mahmassani (2004), who establish a vulnerability index to measure the importance of a specific link to the connectivity of an origin–destination pair, of Jenelius, Petersen, and Mattsson (2006), who also derive several link-importance and site-exposure indices based on the increase in generalized travel cost and of Snelder (2010), who develops a general design method for robust road networks. Examples for the analysis of the impact on transportation demand are given by Schulz (2011) and Erath, Birdsall, Axhausen, and Hajdin (2009), who model the impact on modal choice and destination choice for passenger transport.

The third category consists of models that describe the consequences for the overall economic system. The known methodologies are IO analysis, which describes the relationship between the sectors and was introduced by Leontief (1986), and the general equilibrium models that are based on the assumption of market equilibria (Bröcker, 1998). There are models for both methods that include transportation aspects. One example of an application for the analysis of the impact of risks is the TransNIEMO model of Park et al. (2011). They analyze the impact that terrorist attacks on major bridges would have on the movement of commercial goods and on trade between metropolitan areas. For this purpose, they use a multiregional input-output model (MRIO) which is extended by freight-transportation networks. The National Cooperative Highway Research Program (NCHRP) describes two approaches to analyzing the impact that bottlenecks and interruptions have on freight transportation and the resulting economic impact. They built a high-level methodology which provides the user with the estimated economic costs associated with a particular type of disruption and a more detailed methodology based on the analysis of the supply-chain dynamic (NCHRP, 2012).

In the area of freight transportation modeling, the impact of risks or catastrophes has not yet been in the focus of research. However, several models have been developed that include logistics details (De Jong & Ben-Akiva, 2007; Friedrich, 2010; Holmgren, 2010; Liedtke, 2009; Maurer, 2008; Ramstedt, 2008; Roorda, Cavalcante, McCabe, & Kwan, 2010; Samimi, Kawamura, & Mohammadian, 2011; Tavasszy, Smeenk, & Ruijgrok, 1998; Wisetjindawat, Sano, Matsumoto, & Raothanachonkun, 2007). At this place a detailed review is not helpful since risk aspects are not included yet in these models, we therefore refer to overview papers (like Chow, Yang, & Regan, 2010; De Jong, Vierth, Tavasszy, & Ben-Akiva, 2012; Tavasszy, 2006). Many aspects of these models can be used to analyze the impact of a risk. This will be shown at the end of this chapter, using the example of the SYNTRADE model developed by Friedrich (2010).

11.3. Supply-Chain Risk Management Measures

This section discusses possible categories of supply-chain risk management measures for logistics actors. What can companies do? How do they react? The ideas are inspired by literature and by a case study of a European automobile producer in China analyzing potential measures (Friedrich, Despotov, Zhang, & Kroner, 2012). This research resulted in 10 categories of measures:

- Safety stock and safety capacity: Additional stock and capacity are the most intuitive aspects when considering risk reduction in supply chains. ''Capacity'' in this context can refer to any kind of resource that is needed for operations, such as transportation, storage or production capacities. In logistics research, many studies and models exist that discuss the appropriate level of safety stock. However, disruptions with low probabilities are frequently not covered by these measures, due to cost considerations.
- Alternative transport paths: Transport paths represent the overall transport chain from origin to destination. This includes routes, transportation modes, and transshipment points, excluding warehousing. Shippers usually have business relations with different logistics service providers to ensure access to alternative transport paths in case of problems. From this perspective, this category is closely related to the category of alternative suppliers.
- Alternative supply locations or warehouses: Some products can be delivered from different locations. Either the supplier produces at several locations or there are widespread warehouses that stock the product. In both cases, risks that are spatially limited can be addressed.
- Alternative suppliers: It is possible to have different suppliers for the same product. Multiple, or at least dual sourcing is often established in order to have a better negotiation position. It also limits dependencies on a single supplier, which reduces an important supply-chain risk.
- Flexibility in production: Over the course of the last two decades, a changing business environment has led to the necessity of quickly adapting production to changing market needs. Therefore, concepts such as modularization, standardization of production processes, or flexibilization of production have been developed. These concepts can also be used to handle different types of supply-chain risk.
- Proactive company culture: Authors like Sheffi (2006) recommend that companies establish a proactive (i.e. entrepreneurial and flexible) company culture to improve changeability and increase reactivity in emergency situations.
- Scenario planning: To be prepared, different types of planning may be useful, especially for recurring situations. For instance, in the case of demand fluctuations, where planning, experience and forecasting methodologies can be helpful. Emergency plans exist in many industries to answer specific situations. An example is the supply-chain continuity framework described by Blos, Wee, and Yang (2010). It characterizes six stages of a business continuity planning process together with operational constructs giving a competitive advantage in case of a supply-chain disruption.
- Information/decision support systems: Transparency through information about the situation and possible consequences helps make better decisions. Today, powerful information systems already support processes in daily operations. New technologies such as radio-frequency identification (RFID) or the Global

Positioning System (GPS) make more and more data available. The problem lies in making use of this data and having appropriate decision-support systems for a specific situation.

– Collaboration: Collaboration to reduce supply-chain risk can take place within or outside the supply chain. Within the supply chain, examples of collaboration include the exchange of information or mutual support in critical situations. Outside the supply chain, there might be collaborations between companies in the same sector or region. Some examples of collaboration are the exchange of experience or joint emergency planning. The cooperation between private and public organizations (Public – Private Partnerships (PPP)) is especially useful if prevention measures require funding and would not be put in place by companies on their own.
– Financial hedging or supply-chain insurance (SCI): More and more insurance companies offer supply-chain insurance. While the risks of specific processes, such as the loss of containers at sea or events at specific locations such as fire or flooding, have always been covered, insuring the entire supply chain against disruptions is new. However, calculating such insurance is very complex since the damage is highly dependent on the ability of actors to react to events. In addition to this, indirect effects are hard to quantify. A more direct method is the financial hedging of certain risks. One example would be hedging oil price fluctuations, as described by Elbert, Bogusch, and Özsucu (2012).

This list already shows the wide range of possible measures. However, it is not exhaustive. Other measures could be imagined. Empirically, it is also difficult to clearly identify measures as risk-management measures, since they may have multiple reasons. Some measures are helpful for different risks but also have a positive impact on operational efficiency. Dual sourcing, for example, is also used to maintain a good negotiating position toward suppliers. Flexibility in production processes also improves utilization. Additional stock can also be used to manage price variations. Another reason for the empirical difficulty is the loose definition of the term "risk management." How can risks be distinguished from the ordinary operational problems that have to be handled by management? For example, should disruptions that are caused by poorly educated employees or insufficient process documentation be categorized as failures of risk management? This simple question reveals that, in many companies, risk management and standard operations management are interlinked.

11.4. Challenges in Modeling Supply-Chain Risks

In this section, some challenges in modeling the consequences of supply-chain risks will be discussed. First, five general challenges will be outlined. Then, the influence of the time horizon on the choice of models will be clarified in more detail.

11.4.1. General Challenges

11.4.1.1. Relevance of indirect effects As shown in the previous section, multiple actors have multiple possibilities to react and prepare for risks in supply chains. An unexpected event can be handled by different actors and on different levels. Taking the example of the ash cloud in Europe in 2010, conversations with company representatives have shown that the disruption was addressed by carriers rerouting the cargo, by logistics companies using existing safety stock or by production companies using different suppliers or rescheduling production. The challenge in modeling these reactions is to find all the relevant reaction paths. A key question in this is how far the impact of an event must be mapped: Is a transitive effect, for example, the effect on the second-tier supplier, still relevant? The modeler must keep the balance between mapping all relevant reactions and avoiding unnecessary complexity.

11.4.1.2. Sector and country specifics The reactions to risks and the choice of risk measures depend on the circumstances, for example on the availability of transport infrastructure and logistics services, on the organization of production processes or on product characteristics. Many of these circumstances are specific to sectors and regions. One sector-specific example is that steel-producing facilities in Germany are often accessible by multiple transportation modes (railroad, road, and inland waterways), making them more independent from disturbances specific to one transportation mode than other industries. Another example is the automotive industry, where, unlike in other sectors, supplier parks are frequently in the direct neighborhood of the purchasing original equipment manufacturer (OEM). This ensures short distances, allows quick supplying and lowers the risk of disturbances during transport. As a final example, perishable food products must be transported over long distances to some regions and are grown close to consumption in other regions.

These regional and sectoral specifics have to be incorporated to achieve realistic results. To build a model mapping all sectors for a whole country and taking into account all specifics would be much too complex for today's methods. That is why region- and sector-specific models have to be considered when modeling the consequences of supply-chain risks.

11.4.1.3. Need for explanatory models Models can be descriptive or explanatory (Bossel, 1994). A descriptive model does not explain how the system functions, but rather deduces rules connecting input and output based on observations. Since supply-chain risks are more likely to be connected to infrequent events, observations are rare. Therefore, in most cases, modeling cannot be based on observations and descriptive models are difficult to develop. Explanatory models, on the other hand, do not require a large number of observations. Instead, they explain the functioning of the system based on a system analysis. Thus, it is possible to identify the consequences of changes to system functions. That makes this kind of model suitable for analyzing supply-chain risks.

11.4.1.4. Lack of short-term and medium-term equilibria Many modeling approaches are based on the assumption of equilibrium. Examples include traffic assignment models, general computable equilibrium models or microeconomic market models. In the case of extreme events that cause supply-chain risks, there is a high proportion of dynamic factors in the overall system. If actors in markets or transportation networks change their behavior, less information is available. This causes the absence of a stable equilibrium in the short and medium term. This situation can be described as an economic disequilibrium, but disequilibrium models would only cover part of the case, because, after a certain period of time, a new stable equilibrium may emerge. For the analysis of supply-chain risks, however, the time without equilibrium is most interesting. For this reason, ways of modeling must be found that can map this dynamic in the short and medium term. Due to the higher complexity of dynamic models, calibration and verification is expected to be more extensive than for conventional equilibrium models.

11.4.1.5. Importance of time The level of supply-chain risks is highly dependent on the duration of the disturbances in question. For example, in interviews, logisticians working in the port of Hamburg assumed that a shutdown of the port would have to be longer than one week to have severe consequences for the supply of industry and consumers. Interdependencies in logistics networks are also very dependent on timing. Take, for example, the schedule of container trains. They must leave at specific points in time and there is a certain frequency of trains. That is why, when modeling supply-chain risks, the timeline has to be mapped explicitly and, for this example, the train schedule has to be included.

11.4.2. The Impact of Considered Time Horizons

The objective of the analysis determines the chosen approaches. Looking at supply-chain risks, a key question is which time horizon will be considered. The existing modeling approaches can be allocated to the following categories:

- Very short term (hours): If the event and the consequences are limited to some hours, existing processes do not change significantly. The consequences for transport are, in most cases, limited to changes in routes and the number and sequence of trips in tours. Here, traffic-flow models or process simulations may be the appropriate approach.
- Short term (days): If the events and the consequences expand to several days, reactions already become more complex. Safety stocks are consumed, transport paths change or production processes are adapted. Here, freight transportation and logistics simulations are useful.
- Medium term (weeks): If the events and the consequences expand to weeks, changes become more severe. Suppliers are changed, new warehouses are used and production may be interrupted. Again, freight and logistics simulations are useful.

- Long term (months/years): If the events and the consequences take several months or years, the effects cannot be dealt with by temporary fixes anymore and changes may even be permanent. Supplier relations change, industries disappear or production output decreases. On the macroscopic level, IO models and SCGE models address this situation.

In particular for the analysis of short-term and medium-term horizons, freight transportation and logistics simulations are useful. These models already exist for individual supply chains, but there is still a need for new ideas for overall freight-transport demand.

11.5. Example: Research Plan for Modeling Shortfalls in Food Supply

The work described in this section will be part of a three-year research project that will analyze the vulnerability of overall food distribution in Germany. The food industry is a highly vertically and horizontally integrated sector that has to fulfill customer demands for everyday supplies. It is a critical infrastructure in and of itself and it also depends on critical infrastructure such as transportation, energy supply and telecommunication. Because of these interactions and the many actors involved, it is a complex task to predict the consequences of disturbances in supply or logistics. Therefore, it is important to increase the transparency of the food-supply system through empirical work and modeling. The relevant supply-shortfall scenarios must be analyzed, and the impact of a disturbance has to be simulated to support decisions in the area of crisis management.

The research scope covers the entire German food-supply system, including imports. Flows of goods from the producer to the point of sale at the retailer will be investigated accurately. To reduce the complexity, the goods will be grouped into a limited number of relevant categories (10–20) such as meat, vegetables, frozen food, canned food, milk, and water. The simulation period will cover several weeks, in order to analyze potential severe changes. With this long simulation period, it will be possible to investigate how long the supply system takes to get back to its original state or whether there will be new equilibria after the disruption.

The analyzed scenarios will be: the breakdown of all the producing companies of a region, the breakdown of a number of warehouses, and area-wide disturbances of the transport infrastructure, for example due to heavy snowfall.

The simulation model will be based on SYNTRADE (Friedrich, 2010), a logistics simulation of the food-retailing sector in Germany. SYNTRADE was developed to artificially reproduce existing warehouse structures in food retailing. The objective of the model was to show that this is possible on a large scale (overall sector), opening new possibilities for freight-transport demand modeling. Figure 11.1 shows all existing warehouses on the left map and simulated warehouse locations on the right map. If some individual differences are not considered, the overall structures are very well met. Average distance between existing and simulated warehouses, is about 60km.

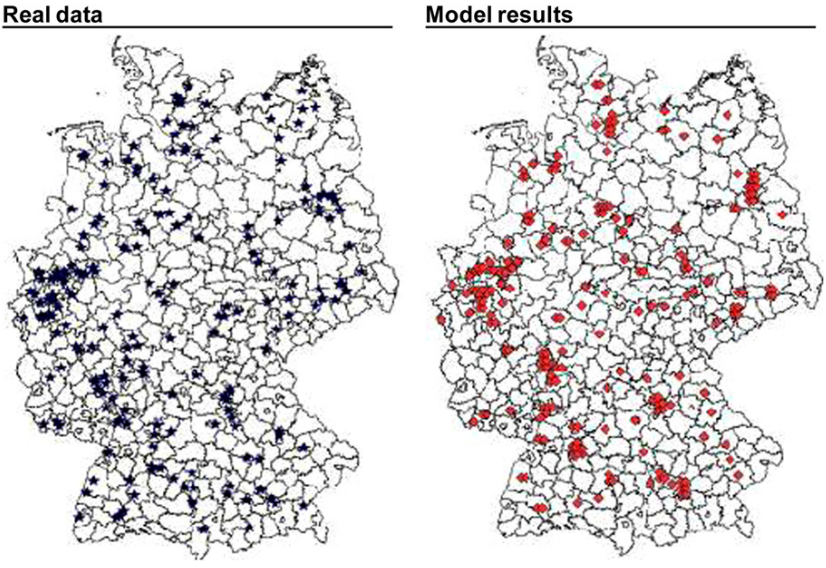

Figure 11.1: Warehouse locations – real locations versus model results of SYNTRADE (Friedrich, 2010).

It uses simulation in the sense of repeatedly executing the logistics decisions of the different actors to arrive at an equilibrium. It does not yet account for time — neither the time of day nor the duration of transports. The only aspect in which time is currently considered is the range of stock in the lot-size model.

The current scope of SYNTRADE is described briefly below: It reproduces structures in Germany, including all imports and exports of consumer products, and differentiates regions down to the NUTS 3 level. Individual commodity flows between producing companies and the points of sale of the food-retailing companies form the basis of the model. The logistics decisions that are included are delivery frequencies, supply paths, simplified tours and warehouse structure decisions. The actors that are described are mainly the food-retailing companies. The decisions of suppliers, wholesalers and logistics service providers are only modeled in a simplified manner. Compared to the new model, the SYNTRADE model differentiates more article types (about 50). Figure 11.2 shows the different supply paths and locations mapped within the model schematically.

To map the influence of risks within the project, five main extensions and simplifications will be implemented. First, time will have to be modeled explicitly (1), for example, to demonstrate the development of stock levels or lead times in case of a disturbance. The plan is to introduce steps of several hours, up to half days. Thus, a logistics simulation (in terms of mapping time) will be developed. Then, since the horizon will be weeks, only short-term decisions will be mapped (2). The warehouse structure will not be part of the simulation. As a result, the decisions of suppliers, logistics service providers and wholesalers will be mapped more precisely (3).

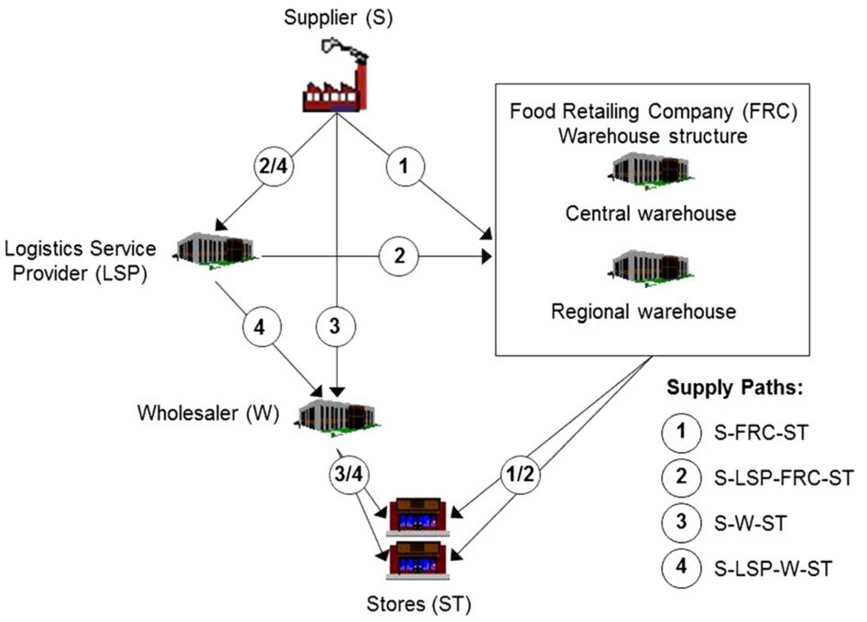

Figure 11.2: Supply-path alternatives in SYNTRADE (Friedrich, 2010).

Additional distribution channels and actors will also have to be mapped (4). This includes other retailers such as restaurants, local markets, direct sales and especially online retailers. Finally, the fluctuation of demand over time will be included, in order to map extreme reactions of consumers in case of catastrophes (i.e., panic buying) (5).

This shows some initial ideas for reusing parts of a model originally designed for freight-transport demand modeling to analyze risks to overall logistics and freight transportation. We think this is a promising path for this research area.

11.6. Synthesis

The analysis and management of supply-chain risks has become more important in a world with globally integrated production networks in which extreme events seem to occur more frequently. This is the case for companies as well as for governments that need to take measures to prepare or handle such risks. There is a lack of transparency, in particular for the overall economic, production and logistics system. Research is needed that generates knowledge on the interrelationships and impact of risks within the system. However, the analysis of effects is tricky since the possible measures that may be taken by actors are manifold, reaching from financial hedging to traditional safety stock.

Models that help understand effects already exist for individual supply chains and logistics networks, for the overall economic system, for the distribution of traffic flows and for the impact on passenger demand. There seems to be a lack of model support for the medium-term impact of risks on the overall freight-transport and logistics system. Such models would have to be explanatory, map time explicitly and probably be specific to sectors and countries.

The experience from recent efforts in transportation-research demand modeling seems to be helpful, and parts of models can be used for the analysis of medium-term impacts on the overall logistics system. One example is the plan to develop such a model based on the SYNTRADE model for the analysis of the impact of risks on food distribution in Germany.

References

Blos, M. F., Wee, H. M., & Yang, J. (2010). Analysing the external supply chain risk driver competitiveness: A risk mitigation framework and business continuity plan. *Journal of Business Continuity & Emergency Planning*, *4*(4), 368–374.

Boden, M., & Weger, H. (2008). Simulation von Verkehrsstörungen und Optimierung von Evakuierungsströmen am Beispiel der Jahrhundertflut 2002. *4. KatNet-Tagung-Netzwerk zur Katastrophenforschung*, Lutherstadt-Wittenberg.

Bossel, H. (1994). Modellbildung und Simulation. Konzepte, Verfahren und Modelle zum Verhalten dynamischer Systeme. Vieweg-Verlag Braunschweig.

Brock, M., Zhang, L., Hayden, C., Matteis, T., & Gross, W. (2012). A comparison of regular and disrupted operations for route planning in freight transportation. *Proceedings of: IDRC Davos 2012 (August 26–30, 2012)*, Davos, Switzerland.

Bröcker, J. (1998). Operational spatial computable general equilibrium modelling. *Annals of Regional Science*, *32*(3), 367–387.

Burgholzer, W., Bauer, G., Posset, M., & Jammernegg, W. (2012). Analysing the impact of disruptions in intermodal transport networks: A microsimulation-based model. *Decision Support Systems*, doi:10.1016/j.dss.2012.05.060.

Centre for Research on the Epidemiology of Disasters (CRED). (2006). *Annual disaster statistical review: Number and trends 2006*. Retrieved from http://www.cred.be/sites/default/files/ADSR_-2006.pdf. Accessed on September 11, 2011.

Chow, J. Y. J., Yang, C. H. Y., & Regan, A. C. (2010). State-of-the art of freight forecast modeling: lessons learned and the road ahead. *Transportation*, *37*(6), 1011–1030.

Christopher, M., & Peck, H. (2004). Building the resilient supply chain. *International Journal of Logistics Management*, *15*(2), 1–13.

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

De Jong, G., Vierth, I., Tavasszy, L., & Ben-Akiva, M. (2012). Recent developments in national and international freight transport models within Europe. *Transportation*, doi: 10.1007/s11116-012-9422-9.

Elbert, R., Bogusch, C., & Özsucu, Ö. (2012). Risk management for air freight forwarders: Analysis of flexible price agreements and financial hedging. *Proceedings of 9th international meetings on logistics research*, Montreal, Canada.

Erath, A., Birdsall, J., Axhausen, K. W., & Hajdin, R. (2009). Vulnerability assessment of the Swiss road network. *Transportation Research Record: Journal of the Transportation Research Board*, *2137*, 118–126.

Friedrich, H. (2010). *Simulation of logistics in food retailing for freight transportation analysis*. Dissertation. Karlsruher Institut für Technologie (KIT), Karlsruhe. Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW). Retrieved from http://digbib.ubka.uni-karlsruhe.de/volltexte/1000020602. URN: nbn:de:swb:90-206022.

Friedrich, H., Despotov, S., Zhang, L., & Kroner, P. (2012). Measures of Supply Chain Risk Management. *Proceedings of: 9th International Meeting on Logistics Research (August 15–17, 2012)* Montreal, Canada.

Holmgren, J. (2010). *On the integration of optimization and agent technology for transportation and production planning*. Dissertation. Blekinge Institute of Technology, Doctoral Series No. 2010:09, ISSN 1653-2090.

Holguín-Veras, J., Jaller, M., van Wassenhove, L. N., Pérez, N., & Wachtendorf, T. (2012). On the unique features of post-disaster humanitarian logistics. *Journal of Operations Management*, *30*(7-8), 494–506.

Jenelius, E., Petersen, T., & Mattsson, L. G. (2006). Importance and exposure in road network vulnerability analysis. *Transportation Research Part A: Policy and Practice*, *40*(7), 537–560.

Knemeyer, A., Zinn, W., & Eroglu, C. (2009). Proactive planning for catastrophic events in supply chains. *Journal of Operations Management*, *27*(2), 141–153.

Köhler, H. (2011). Supply Chain Risiken im Low Cost Country Sourcing: Reduktion von Lieferrisiken in China und der Türkei. Berlin. Erich Schmidt Verlag.

Leontief, W. (1986). Input-output economics. Oxford University Press, ISBN 9780195035278.

Liedtke, G. (2009). Principles of micro-behaviour commodity transport modelling. *Transportation Research Part E*, *45*, 795–809.

Maurer, H. (2008). *Development of an integrated model for estimating emissions from freight transport*. Ph.D thesis. University of Leeds, Institute for Transport Studies.

McKinnon, A. (2006). Life without trucks: The impact of a temporary disruption of road freight transport on a national economy. *Journal of Business Logistics*, *27*(2), 227–250.

Monti, G. (2011). *Resilience engineering #1: Robust vs. resilient*. ActiveGarage. Retrieved from http://www.activegarage.com/resilience-engineering-1-robust-resilient. Accessed on September 12, 2011.

Murata, T. (1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, *77*(4), 541–580.

Murray-Tuite, P., & Mahmassani, H. (2004). Methodology for determining vulnerable links in a transportation network. *Transportation Research Record*, *1882*(1), 88–96.

(NCHRP) National Cooperative Highway Research Program (2012). *Report 732: Methodologies to estimate the economic impacts of disruptions to the goods movement system*. Retrieved from http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_732.pdf. Accessed on November 10, 2012.

NEA. (2011). *Closure of river rhine at the lorelei rock: Estimate of impact and allocation of damage*. Retrieved from http://nea.panteia.nl/nl/Regie/~/media/7F28FB4D1321454084D7E6FAE980B9DF.ashx/. Accessed on November 25, 2012.

Norman, A., & Jansson, U. (2004). Ericsson's proactive supply chain risk management approach after a serious sub-supplier accident. *International Journal of Physical Distribution & Logistics Management*, *34*(5), 434–456.

Park, J. Y., Cho, J. K., Gordon, P., Moore, J. E., Richardson, H. W., & Yoon, S. S. (2011). Adding a freight network to a national interstate input–output model: a TransNIEMO application for California. *Journal of Transport Geography*, *19*, 1410–1422.

Pfohl, H.-C. (2008). Sicherheit und Risikomanagement in der Supply Chain, Deutscher Verkehrs-Verlag, Hamburg.

Ramstedt, L. (2008). *Transport policy analysis using multi-agent-based simulation*. Dissertation. Series No. 2008:09, Blekinge Institute of Technology, School of Engineering.

Roorda, M. J., Cavalcante, R., McCabe, S., & Kwan, H. (2010). A conceptual framework for agent-based modeling of logistics services. *Transportation Research Part E*, *46*, 18–31.

Samimi, A., Kawamura, K., & Mohammadian, K. (2011). A Behavioral Analysis of Freight Mode Choice Decisions. *Journal of Transportation Planning and Technology*, *34*(3), 857–869.

Schmidt, K., & Schaich, E. (2011). Risiko Kurzerklärung. Gabler Wirtschaftslexikon, Wiesbaden. Retrieved from http://wirtschaftslexikon.gabler.de/Definition/risiko.html. Accessed on November 17, 2011.

Schulz, C. (2011): *The identification of critical road infrastructures — The case of Baden-Wuerttemberg*. Dissertation. Karlsruher Institut für Technologie (KIT), Karlsruhe. Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW), Retrieved from http://digbib.ubka.uni-karlsruhe.de/volltexte/1000025466. URN: nbn:de:swb:90-254664.

Sheffi, Y. (2006). Worst-Case-Szenario. mi-Fachverlag, Landsberg am Lech.

Sheffi, Y. (2009). Business continuity: A systematic approach. In H. W. Richardson, P. Gordon & J. E. Moore, (Eds.), *Global business and the terrorist threat*. UK: Edward Elgar Publishing Limited.

Simchi-Levi, M. (2012). Managing Supply Chain Risks: Flexibility – From Practice to Theory. *Proceedings of: 9th International Meeting on Logistics Research (August 15 – 17, 2012)* Montreal, Canada. Retrieved from http://expertise.hec.ca/rirl/en/wp-content/uploads/2011/05/Simchi-Levi-Montreal_August_2012_Final.pdf. Accessed on August 30, 2012.

Skelton, P. (2007). Business continuity and supply chain management: How to manage logistical operations in the event of an interruption or emergency. *Journal of Business Continuity & Emergency Planning*, *2*(1), 13–20.

Snelder, M. (2010). *Designing Robust Road Networks — A general design method applied to the Netherlands*. Doctoral thesis TU Delft. Retrieved from http://www.nextgenerationinfrastructures.eu/download.php?field=document&itemID=552898. ISBN: 978-90-5986-370-5.

Tamminga, G., Tu, H., Daamen, W., & Hoogendoorn, S. (2011). Influence of Departure Time Spans and Corresponding Network Performance on Evacuation Time. Transportation Research Record: Journal of the Transportation Research Board, Issue 2234, 2011, pp. 89–96.

Tang, C. S., & Tomlin, B. (2008). The power of flexibility for mitigating supply chain risks. *International Journal of Production Economics*, *116*(1), 12–27.

Tavasszy, L. A. (2006). Freight modelling – An overview of international experience's. Paper presented at the 85th annual meeting of Transportation Research Board, Washington, DC.

Tavasszy, L., Smeenk, B., & Ruijgrok, C. (1998). A DSS for modelling logistic chains in freight transport policy analysis. *International Transactions in Operational Research*, *5*, 47–59.

Tomlin, B. (2006). On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science*, *52*(5), 639–657.

Tuncel, G., & Alpan, G. (2010). Risk assessment and management for supply chain networks: A case study. *Computers in Industry*, *61*, 250–259.

Tzes, A., Kim, S., & McShane, W. (1996). Applications of Petri networks to transportation network modeling. *IEEE Transactions on Vehicular Technology*, *45*(2), 391–400.

Wagner, S. M., & Bode, C. (2006). An empirical investigation into supply chain vulnerability. *Journal of Purchasing & Supply Management*, *12*, 301–312.

Wilson, M. (2007). The impact of transportation disruptions on supply chain performance. *Transportation Research Part E*, *43*, 295–320.

Wisetjindawat, W., Sano K., Matsumoto, S., & Raothanachonkun, P. (2007). Micro-Simulation Model for Modeling Freight Agents Interactions in Urban Freight Movement. Paper presented at the 86th annual meeting of Transportation Research Board, Washington, DC.

Chapter 12

Decision-Making Process and Factors Affecting Truck Routing

Tomer Toledo, Yichen Sun, Katherine Rosa, Moshe Ben-Akiva, Kate Flanagan, Ricardo Sanchez and Erika Spissu

Abstract

This research studies the decision-making process and the factors that affect truck routing. The data collection involved intercept interviews with truck drivers at three rest area and truck stop locations along major highways in Texas, Indiana, and Ontario. The computerized survey solicited information on truck routing decisions, the identity of the decision makers, the factors that affect routing and sources of information consulted in making these decisions. In addition, stated preferences (SP) experiments were conducted, in which drivers were asked to choose between two route alternatives. A total of 252 drivers completed the survey, yielding 1121 valid SP observations.

This data was used to study the identity of routing decision makers for various driver segments and the sources of information used both in pretrip planning and en route. A random effects logit model was estimated using the SP data. The results show that there are significant differences in the route choice decision-making process among various driver segments, and that these decisions are affected by multiple factors beyond travel time and cost. These factors include shipping and driver employment terms, such as the method of calculation of pay and bearing of fuel costs and tolls.

Keywords: Freight transportation; truck route choice; toll road

Freight Transport Modelling

ISBN: 978-1-78190-285-1

12.1. Introduction

Trucks are the dominant mode of freight transportation in the United States. In 2009, trucks carried freight at a value of 9.5 billion dollars, which is about 65% of the value of freight transported by all modes. The total annual highway miles driven by trucks have increased by 109% between 1980 and 2008 (Schmitt & Sprung, 2010). This growth rate is higher compared with that of general road traffic. The highway transportation system has not grown at a comparable rate. Its total route length has increased by only 5% during the same period (Schmitt & Sprung, 2011). This discrepancy contributes to increased congestion, energy consumption, and degradation of the environment and traffic safety.

Understanding the behavior of road users is critical in order to develop measures to improve the performance of transportation networks. However, while there have been numerous studies of the relevant passenger travel behaviors, the research on truck routing choices is limited.

Toll road operation is a useful example to demonstrate the need to better understand truck routing behavior. Heavy trucks are critically important for toll roads because of their importance in generating revenue. Bain and Polakovic (2005) found that trucks account for 10% of traffic flow on toll roads, but generate 25% of the revenue. In many cases, the use of toll roads, after they opened, was lower than originally forecasted, with an overestimation of traffic by 20–30% in the first five years of operation. Furthermore, forecasting errors for truck traffic were larger compared with those for light vehicles (Prozzi et al., 2010). This uncertainty, often over-forecasting flows and revenue, contributes to increased risks in the development of toll roads. Thus, better understanding of trucks' route choices is important to improve toll road use forecasts. It may also help road operators design measures that would make toll roads more attractive to trucks.

This research studies the decision-making process and the factors that affect truck routing. The remainder of the chapter is organized as follows: the next section reviews previous studies that addressed truck routing behavior. Then, the survey that was developed to collect data on truck routing decisions is presented. The following sections present analysis of the data and the route choice model that was estimated with the SP data. Finally, a conclusion is presented.

12.2. Literature Review

Most studies of truck route choice behavior are value of time (VOT) studies, which consider the trade-off between travel time and cost. Zamparini and Reggiani (2007) conducted meta-analysis of 46 previous studies on truck VOT. They found a mean VOT of $20/hour with a coefficient of variation of 0.66. Some of the differences among VOT values could be explained by the geographic location of the study, the GDP of the country where it was conducted and the shipping mode (five of the studies addressed rail transport). Wynter (1995) found wide variability also in VOT of French carriers. A lognormal distribution of VOT, with a coefficient of variation of 0.69, was fitted to SP responses from 408 fleet managers. The study also found that the mean VOT increases linearly with the trip length and varies considerably among various commodity types. Kawamura (2000) found even higher variability in VOT among carriers in California in the context of toll lanes. He estimated a lognormal distribution of VOT with a mean of $23/hour and a coefficient of variation of 1.37. Smalkoski and Levinson (2005) found a wide range of VOT among carriers in Minnesota, from $21/hour to $78/hour, depending on the type of facility being served. They found statistically significant higher VOT for for-hire carriers compared with private fleets ($60/hour and $42/hour, respectively). In contrast, Bergkvist (2001) found that the VOT of Swedish shippers is higher for private carriers compared with for-hire ones. With respect to trip length, Bergkvist found higher VOT for short trips (less than 3 hours) compared with longer ones. This result contradicts that of Wynter (1995). de Jong (2000), in a study of UK carriers, also found differences between the VOT of for-hire and private carriers. However, the results depended on the way the scenarios were presented: VOT were lower for private fleets in abstract scenarios, but higher in scenarios defined in a route choice context. Miao, Wang, and Adams (2011) recognized the importance of the specific conditions relative to the delivery schedule. They estimated VOTs between $26/hour to $68/hour, depending on the geographic location (Wisconsin and Texas) and on the relations to the scheduled arrival time. In addition, they found higher VOT for drivers for private carriers compared with owner-operators (OO) and for-hire drivers, and for drivers paid by miles compared with other drivers. As expected, drivers who paid the tolls themselves were less willing to use toll roads.

VOT studies are very limited in that they only consider travel time and cost and ignore the effects of any other factors. The wide range of freight VOTs across studies or within one study for various segmentations suggests that additional factors affect routing decisions. However, few studies linked truck route choices to other factors beyond time and cost. Small, Noland, Chu, and Lewis (1999) showed that carriers in California were highly sensitive to late schedule delays. When accounting for the schedule delay, the travel time itself was not significant in predicting route choices. Knorring, He, and Kornhauser (2005) found that long-haul truckers are willing to trade an increase of 1% in their travel distance for a speed gain of 0.4 mph in situations in which they have a choice between a route passing through a metropolitan area and a bypass route. Hyodo and Hagino (2010) found an effect for the road type, in addition to tolls and travel times, on truck route choices in Japan.

In the context of toll alternatives, Hunt and Abraham (2004) found that the attributes of travel time, toll cost, primary road type (freeways or surface streets), and the probability and magnitude of delays had significant effects on truck route choices in SP data collected in Montreal, Canada. The value of delay they estimated was greater than the VOT. Wood (2011) studied the factors that affect toll road usage. In most cases, only familiarity with the scenario described in the question (i.e., tolled turnpike, bypass road, or bridge) was associated with an increased willingness to pay tolls. Prozzi, Rutzen, Robertson, Prozzi, and Walton (2009) conducted a survey of carriers in Texas on their use of toll roads in the state. The main reasons to use toll roads that respondents provided were time savings and reduced congestion. Some respondents also noted better road quality, safer travel, and shorter distances. The main reason to avoid toll roads was the price.

These studies suggest that the not only travel time and cost but also risk of delays, the delivery schedule constraints, and the ultimate bearer of costs affect route choices and can help explain some of the large variability in estimated VOT values. Studies related to the choice of carrier service (e.g., Bolis & Maggi, 2001; Danielis, Marcucci, & Rotaris, 2005; de Jong, Bakker, Pieters, & Donselaar, 2004; Fowkes & Whiteing, 2006; Fuller, Rockliffe, Wigan, Tsolakis, & Thoresen, 2003; Jovicic, 1998; Kurri, Sirkia, & Mikola, 2000) also show the importance of the risk of delays and late deliveries to shippers and that the value placed on these attributes varies for different shipments, such as truckload (TL) or less-than-truckload (LTL), and by commodity types and values.

12.3. Survey

Data on the decision-making process related to truck routing and the factors that affect it, was collected using a computerized survey. The survey included two parts. The first part collected information on the routing decision making for the shipment they were transporting at the time of the interview. In addition, information on the driver and carrier characteristics, the contractual or employment terms for the driver (i.e., basis for calculation of compensation and terms related to the costs of fuel and tolls) was collected.

The second part included a stated preferences (SP) experiment. Respondents were asked to choose between two hypothetical route alternatives. The alternatives were defined by the values of the factors shown in Table 12.1. The SP questions were developed around two typical toll road scenarios:

- Bypass scenario: A choice between an urban freeway passing through the downtown of a metropolitan area and a bypass alternative, which has longer distance, but less congested and so may be faster. The bypass may also be tolled.
- Turnpike scenario: For a long section of a trip passing through a rural area, a choice between a tolled highway and a free parallel road, which offers a lower design level (e.g., includes signalized intersections).

With both scenarios the questions were set in the context of a future trip with the same origin, destination and delivery (or pick-up) schedules as the one the drivers were transporting at the time of the interview. A design with 40 cases in ten blocks of four cases was developed using the AlgDesign package in R (Wheeler, 2004). This procedure uses Fedorov's algorithm applied to a randomly selected subset of the possible set of candidate cases to obtain the D-optimal design and blocking. Each respondent was randomly assigned with one block for each scenario. Thus, each respondent was presented with a total of eight cases.

Table 12.1: Factors and their levels in the SP experiment

Scenario	Factors	Levels
Bypass	Difference in travel distance (miles)	5, 10, 15, 20
	Difference in expected travel time (min.)	0, 10, 20, 30
	Frequency of delays that exceed 30 min (in 10 trips)[a]	0, 1, 1, 2 (bypass-v1) 0, 1, 2, 4 (bypass-v2) 0, 2, 4, 6 (downtown -v2)
	Toll amount ($)	0, 5, 10, 15
	Toll payment method	Cash, electronic
	Toll bearer	Driver, other
	Toll reimbursement method (if applicable)	Prepaid, reimbursed
Turnpike	Difference in travel distance (miles)	−20, −10, 0, 10, 20
	Difference in expected travel time (min.)	0, 15, 30, 45, 60
	Toll amount ($)	10, 15, 20, 25, 30
	Toll payment method	Cash, electronic
	Toll bearer	Driver, other
	Toll reimbursement method (if applicable)	Prepaid, reimbursed
	Free road type	2 lane undivided, 4 lane divided

[a]In the first version of the SP experiment, drivers were asked to report delay probabilities that they have experienced for the downtown route and to use these in their SP responses. They were only given the values for the bypass alternatives. Later, they were given values for both alternatives.

The surveys were implemented on Apple iPad tablets using the iSurvey application (iSurvey, 2012). Questions were read to participants, and responses were recorded by the interviewer. Participants were not compensated. The survey was administered on several days between February and June 2012 to drivers at rest areas and truck stops on or near three highway corridors:

- I-35 near Salado, north of Austin, Texas
- Ontario Highway 401 near Ayr, west of Toronto, Canada
- Lake Station on the west end of the Indiana Toll Road.

The collected data set includes responses from 252 drivers (118 in Texas, 53 in Ontario and 81 in Indiana) and 1121 valid SP choices.

12.4. Results

The results presented below are derived from the responses in all three locations. For some items, there were differences (questions were added) between the questionnaires used. Therefore, the sample sizes relevant to each analysis differ.

12.4.1. Sample Composition

Seventy-five percent of the drivers in the sample are hired drivers. Within these, 56% worked for for-hire carriers and 19% for private fleets. The remaining 25% are OO. Seventy-eight percent of the drivers were transporting TL shipments when they were interviewed. Ten percent were LTL shipments and 12% were either parcels, empty trips or others. Most trips (72%) did not involve any special shipping service. Sixteen percent involved temperature control and 5% involved shipment of Hazmats. Overall these figures are consistent with figures published by the Census Bureau (USCB, 2002).

Some aspects of the driver's employment terms, especially those related to compensation and bearing of various costs, may affect routing decisions. The employment terms for the overall sample and for the hired and OO segments are summarized in Table 12.2. The majority of drivers are paid a fixed amount for a specific trip, which does not depend on their routing. Most commonly, drivers are paid by book miles. The only two payment calculation methods in which that relate to the actual travel time and distance are drivers paid by hours (12%) and to lesser extent drivers paid by actual miles (12%). Some hired drivers are paid by actual miles or hours (14% and 15%, respectively). These methods are less frequent for OOs (3% and 6%, respectively). The terms are very different for hired drivers and OOs with respect to fuel and toll costs. For 92% of hired drivers, but only 5% of OOs, the company is responsible for fuel costs. The situation with respect to toll is similar. Eighty-nine percent of hired driers, but only 24% of OOs reported that their company is fully responsible for tolls. OOs are also less likely compared with hired drivers (50% and 68%, respectively) to have electronic toll collection (ETC) tags.

Table 12.2: Employment terms

Characteristic		**Overall ($N = 252$)**	**Hired ($N = 192$)**	**OO ($N = 64$)**
Pay calculation method	Book miles	47%	48%	38%
	Actual miles	12%	14%	6%
	Hours	12%	15%	3%
	Others	29%	23%	53%
Bearer of fuel costs	Company	69%	92%	5%
	Driver – partially	15%	2%	54%
	Driver	16%	7%	41%
Bearer of toll costs	Company	74%	89%	24%
	Driver – partially	2%	68%	50%
	Driver	16%	5%	14%
	Other/no answer	8%	3%	56%
Electronic toll tag	With tag	65%	68%	50%
	Without tag	35%	32%	50%

12.4.2. Routing Decision Maker

In identifying the routing decision makers, a distinction was made between pretrip route planning and en-route adjustments. In the route planning phase, drivers may be assigned a route or choose on their own. An assigned route may be mandatory, or a recommended one that they can ask for approval to change or freely choose another one. Drivers that choose their routes may be required to do so from a set of preapproved alternatives, get their chosen route approved, or be to make their own choice. En-route drivers may not be allowed to change routes at all, may ask and be assigned a new route, or they may change their route on their own freely or after getting approval for the change. Table 12.3 shows the distribution of responses for both planning and en-route decision making for the overall sample and various segments within it.

The majority of drivers report that they are responsible for routing decisions. At the planning stage 65% of drivers were free to choose their own routes. Only 16% were assigned a route that they had to follow. While en route, drivers have even more flexibility to change their routes. Eighty-four percent reported that they could change their routes freely. Only 2% cannot change at all or will be reassigned a route by their company. This result indicates that drivers have substantial responsibility in managing their routes. OOs, almost always, decide their own routes, both at the planning stage and en route. In contrast, only 53% of hired drivers freely choose their own routes. The rest experience different levels of supervision, with 21% taking required follow routes assigned to them. Still, 96% of hired drivers can change their routes while driving, either freely (80%) or after obtaining approval. Drivers carrying LTL shipment play lesser roles in deciding routes. Only 50% of LTL drivers choose their own route freely, compared with 65% of TL drivers. At the other extreme, 25%

Table 12.3: Planning and en-route routing decision-making by driver and shipment type

			Driver type		**Shipment type**	
		Overall ($N = 153$)	**Hired ($N = 114$)**	**OO ($N = 39$)**	**TL ($N = 119$)**	**LTL ($N = 16$)**
Planning	Assigned – must follow	16%	20%	5%	16%	25%
	Assigned – approval	2%	3%	0%	1%	6%
	Assigned – freely	8%	11%	0%	9%	13%
	Choose – alternatives	7%	10%	0%	7%	6%
	Choose – approval	2%	3%	0%	2%	0%
	Choose – freely	65%	54%	95%	65%	50%
En route	Not allowed	3%	3%	0%	1%	6%
	Reassigned	1%	1%	0%	1%	0%
	Approval	12%	16%	0%	13%	19%
	Freely	85%	80%	100%	85%	75%

of LTL drivers must follow an assigned route, as opposed to only 16% of TL drivers. While the sample size for LTL is rather small, these patterns are consistent in all decision-making options. Similarly, 85% of TL drivers may change their route freely while driving, compared with only 75% of LTL drivers.

Table 12.4 shows the routing decision makers for various driver segments in terms of the bearing of fuel and toll costs and the method of pay calculation. Drivers may be fully, partially or not at all responsible for the cost of fuel and tolls. Drivers that are fully or partially (e.g., receive surcharges) responsible for the cost of fuel overwhelmingly have the right to choose routes on their own. Drivers that are not responsible for fuel costs at all are more restricted in their routing. Only 53% can choose their routes freely. Twenty percent are assigned routes that they must follow. Eighty-one percent can change their route while driving, compared with almost all drivers that pay for fuel themselves. A similar pattern is observed for toll costs. Eighty-nine percent of drivers that are fully or partially responsible for tolls select their own routes, and 100% can freely change their routes while driving. In contrast drivers freely choose routes, pretrip, and en route only in 57% and 82% of the cases, respectively, when they are not responsible for tolls. With respect to the method used to calculate the drivers' pay, the other category, which combined payment options that are unrelated to routing (i.e., fixed amounts or depending on the load weight, value or the freight charges) have the highest level of freedom in choosing their routes (81% pretrip and 91% en route). Drivers paid by hours, whose pay depends the most on routing decision had the least flexibility in making decisions (47% and 71% for pretrip and en route, respectively).

12.4.3. Sources of Information

Information about the sources of information that drivers use when planning their routes and the way they learn about delays on their routes while driving was also collected. Drivers were asked to rate the frequency at which they use various information sources on a 5-point scale. Drivers mainly base routing choice on their own prior experience. All drivers indicated that they rely on it at least half the time. Maps and navigation systems are also useful sources (62% and 65%, respectively use it at least half the time). En route, other drivers are the most frequent source of information (72% use it at least half of the time). The company is not perceived as a significant source of information at any stage. Only 27% and 18% receive information from it at least half of the time, pretrip and en route, respectively (Table 12.5).

12.4.4. Factors that Affect Route Choices

Respondents were also asked about the frequency at which several factors affect their routing decisions. Four factors were considered: travel time predictability,

Table 12.4: Planning and en-route routing decision making by employment terms

		Driver bears fuel cost			**Driver bears tolls**			**Pay calculation method**			
		No	**Partly**	**Yes**	**No**	**Partly**	**Yes**	**Book miles**	**Actual miles**	**Hours**	**Others**
		(*N* = 118)	**(*N* = 23)**	**(*N* = 18)**	**(*N* = 32)**	**(*N* = 4)**	**(*N* = 24)**	**(*N* = 66)**	**(*N* = 20)**	**(N = 17)**	**(N = 53)**
Planning	Assigned – must follow	20%	9%	6%	12%	0%	8%	21%	20%	23%	9%
	Assigned – approval	3%	0%	0%	3%	0%	0%	1%	0%	12%	0%
	Assigned – freely	12%	0%	0%	16%	0%	4%	18%	5%	6%	0%
	Choose – alternatives	9%	0%	0%	6%	0%	0%	5%	15%	12%	6%
	Choose – approval	3%	0%	0%	0%	0%	0%	2%	0%	0%	4%
	Choose – freely	53%	91%	94%	63%	100%	88%	53%	60%	47%	81%
En route	Not allowed	3%	0%	0%	3%	0%	0%	2%	5%	6%	2%
	Reassigned	1%	0%	0%	0%	0%	0%	1%	0%	0%	0%
	Approval	15%	4%	0%	9%	0%	0%	14%	10%	23%	7%
	Freely	81%	96%	100%	88%	100%	100%	83%	85%	71%	91%

Table 12.5: Sources of information used in making routing decisions

		Never 1	Seldom 2	Half 3	Usually 4	Always 5	Avg.	Std.
Planning	Prior experience ($N = 11$)	0%	0%	9%	73%	18%	4.1	0.5
	Navigation ($N = 58$)	26%	9%	20%	21%	24%	3.1	1.5
	Map ($N = 58$)	29%	9%	17%	21%	24%	3.0	1.6
	Other drivers ($N = 11$)	18%	46%	9%	27%	0%	2.5	1.1
	Company ($N = 11$)	37%	36%	18%	0%	9%	2.1	1.2
En route	Navigation ($N = 146$)	53%	7%	6%	13%	21%	2.4	1.7
	Highway Radio ($N = 146$)	40%	8%	15%	20%	17%	2.7	1.6
	Other drivers ($N = 148$)	21%	7%	16%	28%	28%	3.3	1.5
	Company ($N = 149$)	67%	15%	8%	6%	4%	1.7	1.1
	No information ($N = 149$)	21%	21%	23%	22%	13%	2.9	1.3

Table 12.6: Factors that affect routing decisions

	Never 1	Seldom 2	Half 3	Usually 4	Always 5	Avg.	Std.
Predictable travel time ($N = 57$)	9%	7%	9%	24%	51%	4.0	1.3
Parking ($N = 58$)	12%	7%	17%	17%	47%	3.8	1.4
Fuel stations ($N = 58$)	7%	5%	10%	16%	62%	4.2	1.2
Fuel consumption ($N = 11$)	46%	27%	27%	0%	0%	1.8	0.8

availability of parking locations, fuel stations that the driver can use and the effect on fuel consumption. The results are presented in Table 12.6. Drivers were most concerned with having fuel stations that they could use (88% at least half the time), followed by having predictable travel times (84%) and by being able to find truck parking (81%). In contrast, the effect of the route on fuel consumption did not factor in their responses. None of the respondents stated that they consider it usually or always.

12.5. Route Choice Model

The SP data was used to develop truckers' route choice model. A utility function is associated with each alternative:

$$U_{\text{int}} = V_{\text{int}}(X_{\text{int}}, \beta_n) + \alpha_i \varepsilon_n + \varepsilon_{\text{int}} \tag{12.1}$$

where U_{int} is the utility of alternative (route) i to individual n in choice experiment t. V_{int} is the systematic part of the utility function. X_{int} and β_n are the explanatory variables in the utility function and the corresponding parameters, respectively. ε_n is an individual-specific error term. α_i is the corresponding parameter for alternative i. $\varepsilon_{\mathrm{int}}$ is a generic error term. The error terms are assumed to be independently and identically drawn from a Gumbel distribution. Under these assumptions, the predicted probability that driver n chooses route i in experiment t is given by:

$$P_{nt}(i|\beta_n, \varepsilon_n) = \frac{\exp(V_{\mathrm{int}}(X_{\mathrm{int}}, \beta_n)}{\sum_{j=1}^{J} \exp(V_{jnt}(X_{jnt}, \beta_n)} \tag{12.2}$$

where $J = \{1, 2\}$ is the set of alternatives.

The utility parameters are defined as individual-specific in order to capture the heterogeneity in tastes within the driver population. In the model estimation, a random coefficients approach is used and the distributions of these parameters in the population are estimated (Ben-Akiva, Bolduc, & Park, 2008). In the current model, two coefficients are assumed to be distributed in the population: the coefficients of the toll amount and of a toll dummy (which takes value of 1 if the road is tolled and 0 otherwise). The coefficients will be formally defined below. Both are assumed to follow log-normal distributions:

$$ln\beta_{Toll,n} \sim N(\beta_{Toll}, \sigma^2_{\beta_{Toll}}) \tag{12.3}$$

$$ln\beta_{TollD,n} \sim N(\beta_{TollD}, \sigma^2_{\beta_{TollD}}) \tag{12.4}$$

where $\beta_{Toll,n}$ and $\beta_{TollD,n}$ are the coefficients of toll amount and toll dummy for individual n, respectively. β_{Toll} and β_{TollD} are the corresponding mean parameters of the lognormal distributions. $\sigma_{\beta_{Toll}}$ and $\sigma_{\beta_{TollD}}$ are the corresponding standard deviations.

Table 12.7 defines the variables that were used in the final model specification. The specification includes variables that capture carrier/driver characteristics (TollCompany and Hourly), shipment attributes (Temp), and route attributes (Urban downtown/Rural freeway, Travel Time, Toll, and Risk of Delay).

The utility functions given by:

$$\begin{aligned} U_{int} = {} & \beta_{downtown} + \beta_{free} + \beta_{time}\mathrm{Time}_{int} + \beta_{toll,n}\mathrm{Toll}_{int} + \beta_{tollD,n}\mathrm{TollDummy}_{int} \\ & \times (1 + \beta_{tollcompany}\mathrm{TollCompany}_{int}) + \beta_{delay}\mathrm{Delay}_{int} \\ & \times (1 + \beta_{delayHourly}\mathrm{DelayHourly}_{int} + \beta_{DelayTemp}\mathrm{DelayTemp}_{int}) + \alpha_i\varepsilon_n + \varepsilon_{int} \end{aligned} \tag{12.5}$$

The model was estimated with BIOGEME (BIOGEME, 2012) using simulated maximum likelihood with 5000 Halton draws. The model estimation results are presented in Table 12.8.

Table 12.7: Definitions of variables used in the estimated model

Variables	Definition
Downtown	Downtown constant: 1 if downtown route in bypass scenario, 0 otherwise
Free	Free route constant: 1 if free route in turnpike scenario, 0 otherwise
Time	Travel time (hours)
Toll	Toll amount (2012 US$)
TollDummy	Toll road dummy: 1 if the route involves tolls, 0 otherwise
Delay	Number of trips with delays that exceed 30 minutes (out of 10 trips)
TollCompany	Toll paid by company: 1 if company is responsible for tolls, 0 otherwise
DelayHourly	Number of trips with delays that exceed 30 minutes (out of 10 trips) if driver is paid by the hour, 0 otherwise
DelayTemp	Number of trips with delays that exceed 30 minutes (out of 10 trips) if shipment is temperature controlled, 0 otherwise

Table 12.8: Estimation results

Parameters	Estimated values	*t*-statistics
Downtown	−1.29	−5.90
Free	−0.965	−2.79
Time	−0.874	−2.84
Toll – mean	−4.56	−5.26
Toll – standard deviation	1.53	2.37
Toll dummy	−0.565	−0.98
Toll dummy – standard deviation	0.430	1.31
Toll dummy – company	−1.08	−19.4
Delay	−0.0227	−0.67
Delay – hourly pay	0.123	3.07
Delay – temperature controlled	−0.204	−1.85
$\alpha_{downtown}$	0.976	4.11
α_{free}	1.13	4.65
$\sigma_{Toll,TollD}$	−2.11	−2.62
Number of observations:	1121	
Number of individuals:	143	
Number of Halton draws:	5000	
Final log-likelihood:	−630.86	
Rho-square:	0.188	
Adjusted rho-square:	0.170	

Overall, the estimated values of the parameters are in agreement with prior expectations. As expected, the signs for the coefficients of travel time, tolls, and delays are all negative. These imply that increases in the values of these variables for a specific route alternative reduce the utility of that route and the probability that it will be chosen.

The constants in the model capture the preference of drivers to the specific types of routes described in the two experiment scenarios. In both cases, they imply preference to higher quality and level of service roads. The constant for the downtown route in the urban bypass scenario is negative. This implies that, everything else being equal, drivers prefer the bypass route to the downtown route. Similarly, the negative constant for the free route in the rural highway alternative implies that, everything else being equal (including zero tolls), drivers prefer the toll route.

The coefficients of the toll cost and the toll dummy variables were estimated as random parameter with log-normal distributions. The estimated distribution of the toll cost parameters is given by:

$$\ln\beta_{toll,n} \sim \mathcal{N}(\beta_{Toll}, \sigma^2_{\beta_{Toll}}) = \mathcal{N}(-4.56, 1.53^2) \tag{12.6}$$

Similarly, the estimated distribution of the toll dummy parameters is given by:

$$\ln\beta_{tollD,n} \sim \mathcal{N}(\beta_{TollD}, \sigma^2_{\beta_{TollD}}) = \mathcal{N}(-0.565, 0.43^2) \tag{12.7}$$

The toll dummy variable is also interacted with a dummy variable for the case that the company (and not the driver) is responsible for the toll cost. The estimate value for this variable is -1.08. This means that the negative impact of the toll road on the route choice when the driver is responsible for the toll cost is reversed when the company is responsible for the toll cost.

Other characteristics of the shipment and employment terms were interacted with the delay variable. The compensation for drivers that are paid by hours may increase when they experience delays on their trips. The estimation results show this effect, as they were much less sensitive to the risk of delays on the route. In contrast, drivers that were transporting temperature-controlled goods, were more sensitive to travel delays. This may reflect the higher time-sensitivity that may be associated with these shipments (often perishable) or the higher energy costs of keeping the required temperatures.

The estimated parameter values suggest significant trade-offs among travel time, the use of toll roads, toll costs and the frequency of delays. The estimation of random toll coefficients leads to a distribution of toll values of time. The VOT for the mean toll coefficient is 30 \$/hour. This value is consistent with figures reported in the literature. However, the range of VOT is wide with values from 30 \$/hour and 235 \$/hour between the first and third quintiles. This wide range reflects two extreme attitudes of drivers that were observed in the sample. On one extreme, one group stated that they will not use toll roads in any case. At the other extreme, drivers stated that they will always use the fastest route disregarding any tolls they may incur.

The wide range of attitudes toward toll roads is also apparent when considering the toll road dummy variable. This variable captures the attitude toward using the toll road itself, regardless of the toll amount. Drivers that pay for the tolls themselves, at the first quartile of the distribution would be willing to accept a 29 minutes additional travel time in order to avoid a toll road (before considering the toll cost itself). Drivers at the third quartile would be willing to accept additional 52 minutes of travel time to avoid the toll road. As noted above, this behavior is reversed when the driver is not responsible for the toll costs. In this case drivers are willing to incur additional travel times between 2 minutes (1st quartile) and 4 minutes (3rd percentile) in order to use the toll road.

Two characteristics of the shipment and employment terms were found to affect the disutility associated with the risk of unexpected delays: drivers that are paid by the hour favor delays compared with other drivers. Drivers that transport temperature-controlled shipments are more sensitive to the risk of delays. Other drivers are insensitive to delays, willing to trade-off only 2 minutes of travel time for a 10% reduction in the risk of delays that exceed 30 minutes. Drivers paid by hours are willing to accept 7 minutes longer travel times in order increase their risk of travel delay by 10%. While this result is not expected, it should be noted that the pay for these drivers increases when they are delayed in traffic. In contrast, drivers with temperature-controlled shipments are willing to increase their travel times by 16 minutes to reduce their risk of delays by 10%. This may reflect higher time sensitivity of these goods (perishables) and the additional energy cost for refrigeration associated with travel delays.

The choice between the toll bypass and free downtown routes is used in order to demonstrate the effects of the tolls on route choices. Figure 12.1 shows the estimated probabilities of choice of the tolled bypass as a function of the toll amount for drivers in the 1st, 2nd, and 3rd quartiles of the probability distribution, for the cases that the driver or the company is responsible for tolls. The figure is based on an assumption of equal travel times and delay frequencies in the two routes.

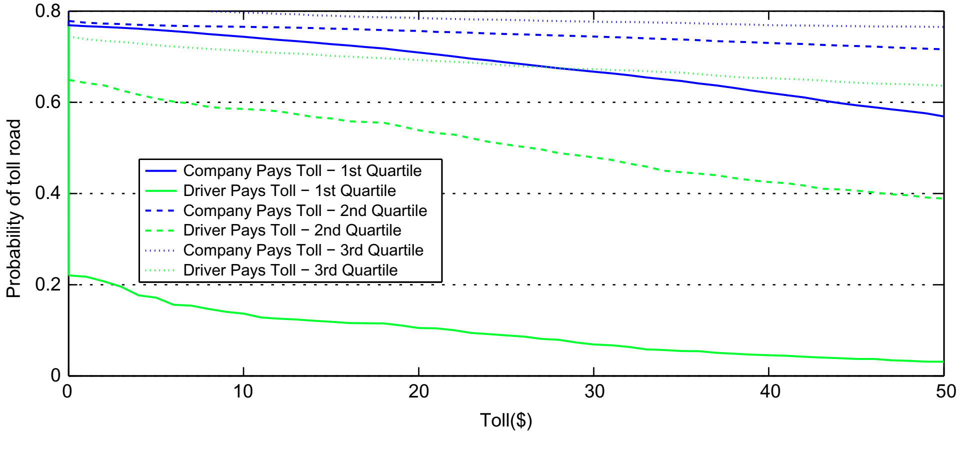

Figure 12.1: Effects of tolls.

For drivers that are responsible for tolls, the introduction of tolls (at toll value zero) sharply reduces the probability that they will choose the toll road. This captures their preference to avoid toll roads. In contrast, when drivers are not responsible for the toll cost, the introduction of tolls does not affect their route choices. Further increases in the toll amounts negatively affect the probability of toll road choice in all cases.

The figure also shows the wide variability in drivers' preferences toward the toll road. The choice probabilities are much lower for drivers that are responsible for the toll cost. But, even within the same segment, and in particular for drivers who are responsible for the toll cost, there are very large differences in the toll road choice probabilities between drivers in the 1st and 3rd quartiles of the distribution (e.g., between probabilities of 0.03 and 0.62 for $50 tolls).

12.6. Conclusion

This research studies the decision-making process and the factors that affect truck routing. Using data collected in intercept interviews with truck driver, the identity of routing decision makers was investigated. The results show that in most cases the driver has the power to choose routes. This is especially the case for OOs and for drivers that are responsible, even if partly, for the cost of fuel and tolls. Furthermore, the sources of information that drivers consult in making routing decisions are limited. They receive little support from their companies. The results also show that drivers consider additional factors beyond travel time and travel cost in deciding their routes. In the survey, drivers mentioned travel time predictability and availability of parking and fuel stations as relevant considerations. Estimation results of an route choice model based on SP data also show that there are significant differences in the route choice decision-making process among various driver segments, and that these decisions are affected by factors that include shipping and driver employment terms, such as the method of calculation of pay and bearing of toll costs. The model also showed a strong preference to avoid toll roads when the driver is responsible for the cost, but indifference to tolls when the driver is not responsible for the cost.

These findings suggest that simple VOT studies that have been used as a basis to predict truck route choices and flows, and in particular in the context of toll roads, may not be adequate. Nevertheless, the current results are based on SP data that represent simplified situations and decision protocols. In ongoing work, research with GPS devices that will collect data on actual routes that driver use is being conducted.

References

Bain, R., & Polakovic, L. (2005). *Traffic forecasting risk study 2005: Through ramp-up and beyond*. London, UK: Standard & Poor's.

Ben-Akiva, M., Bolduc, D., & Park, J. Q. (2008). Discrete choice analysis of shippers' preferences. In M. Ben-Akiva, H. Meersman & E. van de Voorde, *Recent developments in* *transport modeling: Lessons for the freight sector* (1st ed., pp. 135–155). Bingley: Emerald Group.

Bergkvist, E. (2001). *Freight transportation: Valuation of time and forecasting of flows, Umea economic studies 549*. Sweden: Umea University.

Biogeme, Version 2.2, 2012. Retrieved from http://biogeme.epfl.ch/

Bolis, S., & Maggi, R. (2001). Evidence on shippers' transport and logistics choice. In *Proceedings of the 1st Swiss transport research conference*, Ascona, Switzerland.

Danielis, R., Marcucci, E., & Rotaris, L. (2005). Logistics managers' stated preferences for freight service attributes. *Transportation Research Part E*, *41*(3), 201–215.

De Jong, G. (2000). Value of freight travel time savings. In D. A. Hensher & K. J. Button (Eds.), *Handbook of transport modeling* (2nd ed., pp. 553–564). New York, NY: Elsevier.

De Jong, G., Bakker, S., Pieters, M., & Donselaar, P. W. (2004). New values of time and reliability in freight transport in the Netherlands. In Proceedings of the European transport conference 2004, Strasbourg, France.

Fowkes, T., & Whiteing, T. (2006). The value of freight travel time savings and reliability improvements – Recent evidence from Great Britain. In *Proceedings of the European transport conference 2006*, Strasbourg, France.

Fuller, T., Rockliffe, N., Wigan, M., Tsolakis, D., & Thoresen, T. (2003). Economic evaluation of road investment proposals: Valuing travel time savings for freight. Austroads Publication No. AP–R230/03, Sydney Australia.

Hunt, J. D., & Abraham, J. E. (2004). Stated response examination of factors influencing commercial movement route choice behavior. In Proceedings of the 8th international conference on applications of advanced technologies in transportation engineering, Beijing, China, pp. 72–77.

Hyodo, T., & Hagino, Y. (2010). Modeling truck route choice behavior by traffic electronic application data. Paper presented at the Transportation Research Board SHRP 2 Symposium: Innovations in Freight Demand Modeling and Data, Washington, DC.

iSurvey – iPad Survey App Software, Version 2.8.0, 2012. Retrieved from https://www.isurveysoft.com/

Jovicic, G. (1998). Application of models based on stated and revealed preference data for forecasting Danish International Freight Transport. In Proceedings of the annual transport conference at Aalborg University, Aalborg, Sweden.

Kawamura, K. (2000). Perceived value of time for truck operators. *Transportation Research Record*, *1725*, 31–36.

Knorring, J. H., He, R., & Kornhauser, A. L. (2005). Analysis of route choice decisions by long-haul truck drivers. *Transportation Research Record*, *1923*, 46–60.

Kurri, J., Sirkia, A., & Mikola, J. (2000). Value of time in freight transport in Finland. *Transportation Research Record*, *1725*, 26–30.

Miao, Q., Wang, B. X., & Adams, T. M. (2011). Assessing the value of delay to truckers and carriers. Final Report CFIRE 03-15 and UTCM 09-00-45, US Department of Transportation, Washington, DC.

Prozzi, J., Flanagan, K., Loftus-Otway, L., Porterfield, B., Persad, K., Prozzi, J. A., & Walton, C. M. (2010). Toll roads: What we know about forecasting usage and the characteristics of Texas users. TxDOT Project 0-6044, Texas Department of Transportation, Austin TX.

Prozzi, J., Rutzen, B., Robertson, C., Prozzi, J., & Walton, C. M. (2009). Characteristics of the truck users and non-users of Texas toll roads. Report FHWA/TX-10/0-6044-2 Vol. 2, Texas Department of Transportation, Austin TX.

Schmitt, R., & Sprung, M. (2010). Freight facts and figures 2010. Publication FHWA-HOP-10-058. FHWA, US Department of Transportation, Washington, DC.

Schmitt, R., & Sprung, M. (2011). Freight facts and figures 2011. Publication FHWA-HOP-12-002. FHWA, US Department of Transportation, Washington, DC.

Smalkoski, B., & Levinson, D. (2005). Value of time for commercial vehicle operators. *Journal of the Transportation Research Forum*, *44*(1), 89–102.

Small, K. A., Noland, R., Chu, X., & Lewis, D. (1999). Valuation of travel time savings and predictability in congested conditions for highway user cost estimation. NCHRP Report No. 431, Transportation Research Board, Washington, DC.

USCB. (2002). Economic census: Vehicle inventory and use survey. US Census Bureau, Washington, DC.

Wheeler, R. E. (2004). AlgDesign. The R project for statistical computing. Retrieved from http://www.r-project.org/

Wood, H. P. (2011). Truck tolling: Understanding industry tradeoffs when using or avoiding toll facilities. Publication NCFRP Web-Only Document 3/NCHRP Web-Only Document 185. National Cooperative Freight Research Program, National Cooperative Highway Research Program; Transportation Research Board, Washington, DC.

Wynter, L. M. (1995). Stated preference survey for calculating values of time of road freight transport in France. *Transportation Research Record*, *1477*, 1–6.

Zamparini, L., & Reggiani, A. (2007). Freight transport and the value of travel time savings: A meta-analysis of empirical studies. *Transport Reviews*, *27*(5), 621–636.

Chapter 13

Logistics Managers' Stated Preferences for Freight Service Attributes: A Comparative Research Method Analysis

Edoardo Marcucci

Abstract

This chapter reviews a set of articles, based on stated preferences techniques, focusing on logistics managers' preferences for freight transport attributes with the intent of assessing the quality of knowledge applied research has produced, its reliability and transferability. Transport choices are relevant in a globalized economy where aggressive marketing strategies are often used to acquire a competitive advantage in the market. Freight transport modeling has been derived out of the four-step classical approach developed for passenger transport. A new behavioral approach is progressively gaining popularity thanks to the specific behavioral focus it is based upon. The review performed indicates that there are some evident shortcomings in the way research has been performed so far but, at the same time, there is also a high potential if corrective actions are taken. In particular, substantial improvements could be obtained by (1) clearly defining research and reporting protocols, (2) defining and circumscribing who has to be interviewed in the different freight contexts studied, (3) reporting the contractual relationships governing freight movements, (4) reporting freight details (e.g., volume, value, and weight), (5) motivating the attribute selection method used. In summary, there is a need for systematizing the procedures and reporting adopted in applied research by introducing a much higher level of detail and rigor in both defining the object of measurement and in the experimental design protocols employed.

Keywords: Transport demand; stated preferences; logistics

Freight Transport Modelling

ISBN: 978-1-78190-285-1

13.1. Introduction

This chapter critically reviews recent literature on the application of stated preferences (SP) techniques to study logistics managers' preferences for freight service attributes.

The issues investigated are relevant under various points of view. The motivations of the interest are both of general nature (e.g., overall environmental sustainability of the transport sector) and specific (e.g., concerning transport mode choice and its impact on economic competitiveness).

The chapter first recalls the motivations for studying transport related activities, in general, and, subsequently, delves on those justifying the study, in particular, of logistics managers' preferences that represent one of the main determinants of freight related transport choices. In freight transport, in fact, contrarily to what is assumed when dealing with passengers, where the single agent is usually assumed to be the relevant decision-making unit, the definition of the appropriate decision-making unit to consider is both more critical and controversial (Danielis, Rotaris, & Marcucci, forthcoming). Different transactions generating freight decisions involve different agents.

To summarize the main differences between passenger and freight, one could recall that goods neither "think" nor "speak" as opposed to passengers. In the case of freight transport, therefore, it is crucial to precisely define, for each choice situation, the agent actually taking the decision. In this case, in fact, the decision-making unit is not univocally determined and is contingent upon the specific choice situation considered. This crucial issue represents one important focus of the review proposed.

Own-account retailers, for instance, decide autonomously when and how to perform the restocking trips needed whereas the logistic managers of a supermarket chain might buy freight transport services from a third-party logistic service provider (LSP) who has to guarantee prespecified service characteristics; the buyer of the service, however, has no say on how the service is practically produced (e.g., mode choice, routing, and frequency). The two hypothetical cases described suggest that if one is interested in knowing who decides how the transport service is actually performed different people need to be interviewed in different situations. Unfortunately, the case reported does not represent just a speculative sophistication but, on the contrary, is deep-rooted in freight transport. In other words, one has to be very clear about who is interviewed and how relevant is the case analyzed in the specific context considered. In fact, data and experience tell us that own-account transport, for instance, generally plays a much greater role at an urban or local rather than at an international scale. It is, therefore, preferable to specify the geographical scale one is particularly interested in and, consequently, assume different decision makers as relevant in the different contexts. However, in most of the nonurban freight decisions it is reasonable to assume that logistic managers' preferences play an important role in shaping how freight transport supply is offered in the market and this represent the main justification for the chapter focusing on this issue.

In order to asses the quality and reliability of the knowledge acquired on logistics managers' preferences the chapter proposes a comparative evaluation of a selected number of representative papers according to a set of methodologically and practically relevant issues such as: experimental design adopted, attributes used, sample selection methods and others. The chapter aims to provide a critical assessment of what is known, how reliable and transferable it is.

The chapter is structured as follows. Section 13.2 discusses why and to what extent logistic managers' preferences, in a globalized economy, are relevant in determining freight transport related choices. Section 13.3 briefly illustrates the classical approach used to model freight transport so to clarify how the results obtained from investigating logistics managers' preferences can be translated into more valuable knowledge and better predictions. Section 13.4 introduces and briefly discusses a new and fast-rising approach to freight modeling where behavioral considerations are crucial and interaction among operators is explicitly taken into account. Section 13.5 illustrates the potential for using SP techniques and their various applicative nuances in assessing logistics managers' preferences for freight transport attributes. Section 13.6 reports a brief retrospective analysis of freight demand modeling so to clearly define where the SP&RP studies reviewed are conceptually situated and what is the underlying knowledge they rest upon. Subsequently, Section 13.7 reports and comparatively discusses a set of structural parameters of the SP&RP studies focusing on logistics managers' preferences for freight transport attributes along with the results obtained. Finally, Section 13.8 reviews still unsettled issues and controversial results with the intent of suggesting future research directions. Section 13.9 concludes.

13.2. Globalization, Freight Transport, and Logistic Managers' Preferences

There is a strong relationship between economic growth and transportation (Manheim, 1979). This relationship has been studied in two different and integrated ways. One can, in fact, focus the analysis on the contribution freight transport provides to economic growth and, on the other, investigate the impact economic activity has on freight transport demand. The two approaches are, quite clearly, interrelated and endogenously interlinked. The preference for one of the two is mainly due to the final goals the researcher is pursuing (Meersman & Van de Voorde, 2008).

Transport plays a major role in firms' supply chains. The last thirty years have been characterized by a reduction of deep-sea transportation costs thanks to the development of container transport that, allowing the extension of supply chains, consequently induced an increased competition in the global market (Levinson, 2006). The substantial and steady growth of freight flows is linked and, often, induced by the structural changes the world economy is undergoing. These changes mainly involve the restructuring of the manufacturing system and the geographical dislocation of production with consequent operational modifications also for freight distribution systems (Hesse & Rodrigue, 2004). The changes are induced by a pursuit of greater efficiency sought via production delocalization, functional specialization, progressive use of high-quality\low-cost semi-finished goods made possible thanks to a preplanned, integrated and modularized productions system. Some implications for freight logistics are:

1) inventory reduction with production targeted to orders received;
2) reduction of lead times from product order to product delivery;
3) faster rotation of capital invested;
4) customization of production;
5) development of postmanufacturing the closest possible to end-consumers (Holter, Grant, & Ritchie, 2010).

The weight attributed to the various attributes characterizing freight transport services is not determined once and for all. In fact, the relative importance is fundamentally determined by specific market characteristics. World-wide economic developments impact on market characteristics that, in turn, modify the way in which companies compete and the type of logistic and transport services they prefer and buy. The description of the process just illustrated is valid notwithstanding the heterogeneity characterizing the different supply chains that are differentiated both by, as usually considered relevant, freight transport characteristics linked to the goods transported (e.g., frozen food, fresh fish, cut flowers, and gold and diamonds) as well as to the distribution chains along which the goods are moved from producers to end consumers (e.g., direct delivery, great organized distribution, and international importers) that are, usually and unduly, not considered relevant.[1] The specific freight distribution channels, as recently empirically demonstrated in the case of Rome and Trieste (Danielis, Rotaris, & Marcucci, forthcoming), are characterized by different freight transport preferences for specific attributes due to the differentiated allocation of market power along the distribution chain.

Free market forces are one among the sources of change impacting on freight distribution along with regulatory decisions taken by public decision makers.[2]

Assessing the value various firms assign to freight transport characteristics (i.e., willingness to pay-WTP) is of great importance for many actors in a globalized economy characterized by increasing competition, extended logistic chains, complex logistics contracts, high and rising infrastructure costs, unpredictable final demand coupled with long-term infrastructural commitments of use. In fact, carriers could

1. This occurrence is partially due to the insufficient attention researchers pay to these issues and partially to the lack of relevant data given the high cost of acquiring them coupled with confidentiality issues in this realm.
2. An emblematic case, of the some time schizophrenic approach adopted in this sector, is represented, on the one side, by the increased deregulation and harsh competition characterizing European road haulage where the policies implemented have *de facto* favored a sustained growth of road freight demand (McKinnon & Woodburn, 1996) and, on the other, by the speculation, and in some cases, implementation of road pricing systems, truck bans, and intermodal transport subsidizations to induce a modal shift from road to alternative and less polluting modes, with the intent to overcome the excessive pressure on road infrastructures and reduce the negative externalities caused by road freight movements (Liedtke, 2009).

customize their services according to firms' WTP for specific service characteristics; public agencies, transport authorities and local governments could target both their investments and regulatory decisions so to influence those characteristics that are most valued by firms and induce the desired behavior; researchers, could use the estimated WTP measures as inputs in the simulation models used to predict the effects of policy changes (Danielis, Rotaris, & Marcucci, 2005).

13.3. The Classical Approach to Freight Modeling

In order to motivate and clarify why it is useful to assess logistics managers' preferences this section briefly reviews, by citing a list of nonexhaustive scientific examples, the classical approach to freight modeling. Freight models have predominantly been developed for forecasting freight transport volumes, flows, and modal split (de Jong, Gunn, & Walker, 2004; Regan & Garrido, 2002; Zlatoper & Austrian, 1989; Harker, 1985) but when designed by or for public authorities, are also used for testing transport policy measures, and predicting the impacts of new infrastructure provision on traffic. The models applied in freight transport forecasting originate from passenger transport where they were first conceived and developed (de Jong et al., 2004). The possibility of fruitfully adapting these types of models to freight is still controversial. Notwithstanding the similarities, the main differences in the application of the same modeling approach to the two different contexts mainly relate to the diversity in:

1) production and attraction;
2) distribution;
3) modal split;
4) assignment.

The four step model, when applied to freight, involve:

1) decision makers (e.g., shippers, carriers, LSPs, and drivers);
2) items being transported (e.g., parcels, containers, and bulk);
3) limited availability of data.

Furthermore, the adaptation of the model structure usually implies additional transformations that, for instance, imply the conversion of trade flows expressed in monetary terms into physical flows in tones so to determine production and attraction.

Four types of models have been practically applied to operationalize *production and attraction:*

1) trend and time series (e.g., Garrido, 2000);
2) system dynamics (e.g., ASTRA Consortium, 2000);
3) zonal trip (e.g., FHWA, 1999);
4) input–output (e.g., Cascetta, 1997).

All these models use aggregate data. *Distribution* models developed for freight modeling also use aggregate data only. In this case trade flows between origin and destination zones are determined on the basis of production and attraction and a measure of transport generalized cost. Gravity models are commonly used (e.g., Tavasszy, 1994; Tavasszy, Van Der Vlist, & Ruijgrok, 1998). *Modal split* models use both aggregate and disaggregate data. Among the most frequently used one can recall:

1) elasticity-based;
2) aggregate modal split;
3) neoclassical economic;
4) direct econometric demand estimation;
5) disaggregate modal-split;
6) microsimulation;
7) multimodal network.

Aggregate modal split models are usually binomial logit or MNL models estimated via shares of different modes across zones. On the other hand, disaggregate modal split models use shippers' data surveys and/or SP surveys. In this case the decision maker is the firm and the indirect utility function is re-stated as a profit function giving rise to a ''random profit maximization'' problem.[3]

Most of the modal split models estimated limit themselves to mode choice only (e.g., De Jong, Vellay, & Moue, 2001; Fosgerau, 1996; Jiang, Johnson, & Calzada, 1999; Nuzzolo & Russo, 1995) even if there are some disaggregate models that simultaneously deal with mode and logistic choices (e.g., Abdelwahab & Sargious, 1992; Blauwens, 2001) as well as mode and supply chain structure (Arunotayanun, 2009). Trip *assignment* models[4] estimate how given freight trips are allocated to routes representing links of the modal network considered (e.g., truck, rail, sea, and inland waterways). When dealing with logistics managers' preferences we focus on mode choice alone (i.e., third step).

13.4. A Behavioral Framework for Modeling Freight Distribution: The Role of Interaction

Before moving on to the analysis of the experimental design, first, and to logistics' managers preferences for freight transport attributes, second, it is worthwhile recalling a new stream of literature that aims at expanding the elegant extensions of

3. More about this specific and critical issue will be said in the next sections especially with reference to new research on this topic (e.g., Danielis, Rotaris, & Marcucci, forthcoming).
4. Not all four-step models include assignment models and those including it usually limit themselves to assigning trucks to network links.

straightforward passenger four-step models to the significantly more complex context of freight transportation (Hensher & Puckett, 2005). The differences, in terms of the underlying structures and difficulties in sourcing relevant data between passenger and freight transportation, are so relevant that one might even question if it is worthwhile to try and adapt passengers' models to freight. Authoritative researchers (de Jong et al., 2004) underline that the current state of freight transport modeling is generally not capable of capturing sensitivities to policy variables relevant for mode choice but concentrate only on modal share elasticities. Furthermore, when it comes to freight choice modeling one cannot underestimate the relevance of its interdependent nature unless one is willing to accept biased estimates of marginal utilities and elasticities (McFadden, Winston, & Boersch-Supan, 1985). In fact, freight decision making implies optimization over different interrelated dimensions where, for instance, frequency of deliveries might influence the shipment size as well as the time of day dispatch decision that could, in turn, also impact route choice decision.

What is usually lost in the present approach to freight modeling is both the in-depth analysis of the underlying economic behavior driving either individual agents purchasing freight services in the market or self-producing them, as well as the explicit accounting for the interaction between decision makers when choosing among different freight transport service options.[5] In other words, preferences among different agents may vary depending on the role they play in the supply chain (e.g., buyers or sellers of freight services) and, consequently, on the power and influence they exert in the context they operate in.

Departing form urban freight analysis Hensher and Puckett have initiated and progressively developed a new generation of economic-behavior-based freight models (Hensher & Puckett, 2005). In particular, it is important to recall that freight transport demand is a derived demand implying that goods are moved for very specific reasons and an ideal modeling framework (Ogden, 1992) should be able to capture these reasons as well as how the changes in the motivation and consequent ability to move goods affect the subsequent movements. Furthermore, if interactions are important, a modeling approach should also capture the process that explains the different influence each agent has in defining a contract. SP methods are particularly useful when inquiring potential interaction effects since they offer the opportunity to dynamically simulate the interaction between supply and demand; RP data, on the contrary, are not suitable to this end given they only represent the final result of the interaction that took place and provide no information on its dynamics.

5. While interaction can safely be ignored in many choice situations when the analyst is specifically interested in investigating single decision-making unit's preferences in passenger transport, this issue has to be considered when the analyst studies group decisions (e.g., where to go on holidays, which mode to choose for traveling to a destination, and which route to take). In the case of freight, on the other hand, interaction is more the rule than the exception and cannot be ignored without paying a high price in terms of biased estimates.

13.5. Considerations Concerning SP and Logistics Managers' Preferences

As previously recalled, most freight models are based on the four-step methodological approach where the third step comprises the determination of the modal split functions, that is, the distribution of the total freight volume to the alternative transport modes. This is performed by first identifying the main elements determining transport mode choice (e.g., role played by service attributes such as transport cost/price, time, reliability, and frequency) and, subsequently, estimating the demand elasticities of the decision makers considered to modified values of the relevant attributes (Fries & Patterson, 2008). SP methods are usually adapted to this end.

The decision-making process for freight is more complex than for passenger. In particular, probably, the most relevant difference between the two sectors relates to the decision maker. In fact, while the passenger chooses himself the best modal alternative, in the case of a freight shipment this has to be organized and steered by a logistic manager that must know the logistic requirements of each shipment in relation to the commodity shipped, the production process it is needed for and the budget available for that purpose. The shipper has, first of all, a *make* or *buy* decision to take, in other words he needs to decide either to use its own logistics managers or contact a LSP to organize the transport chain.[6] This difference is quite relevant for the implications it bears on the treatment of the mode attribute in applied research. In fact, existing freight demand studies can be subdivided in two groups of different size; the first, and larger, explicitly includes mode as a relevant attribute and focuses on mode choice (*buy*-in this case it is implicitly assumed that the choice of mode is a relevant issue and most probably the logistic manager does not dispose of its own rolling stock) the second, and smaller, group does not include transport mode as a relevant attribute (*make*-in this case it is assumed that mode is not primarily relevant for the shipper's choice) (Marcucci, 2005).

In other words, the choice valuation process can be treated as a maximization problem. In the case of freight transport the firm maximizes a profit function whereas in passenger transport the agent maximizes a utility function. This difference has strong methodological implications connected to the identification of the interviewee who should be considered representative of the firm profit maximization function. Although the mode and route choices can be also taken by the shipping firm, the preferences for the attributes characterizing freight service attributes might be part of the profit function of either the sending, receiving consignee firm or others. The relevance of this issue grows with the growth of integrated logistic processes (Zamparini & Reggiani, 2007).

6. This decision usually hinges on whether the shipper owns the specific means of transport considered appropriate for the given freight movement. We would like to acknowledge the contribution of an anonymous referee who pointed out the fact that large shippers own various means of transportation.

Before discussing in detail the role the decision maker plays in freight analysis, in general, and in mode choice, in particular, it is useful to recall which are the possible sources of errors when performing a SP survey. To summarize and systematize the possible error sources it is convenient to recall which are the main components of a generic choice model. In fact, one can safely say that anytime there is a mismatch between the desired and actual measurement one has an error introduced in the modeling process. The framework for a discrete choice model can generally be reconnected to the following set of structural elements:

1) decision maker: defining the decision-making entity and its characteristics;
2) alternatives: determining the options available to the decision maker;
3) attributes: measuring the benefits and costs of an alternative to the decision maker;
4) decision rule: describing the process used by the decision maker to choose an alternative.

Whenever we make an assumption (e.g., perfect compensatory rule among attributes considered) that is not in line with actual behavior (e.g., choice based on noncompensatory principles-lexicography) we are introducing a source of error in our modeling process (Ben-Akiva & Bierlaire, 1999).

13.6. A Retrospective on SP Freight Demand Studies

Freight demand models date back to the 1960s where both aggregate and disaggregate models were developed (Winston, 1981). Disaggregate models reflect the underling behavioral driving forces explaining freight transport decision making including who actually takes the relevant decisions. The first SP studies concentrating on freight demand research date back to the early 1990s (de Jong, Gommers, & Klooster 1992; Swait, Louviere, & Williams, 1993; Transek, 1990; Widlert & Bradley, 1992). The early studies investigated the inclusion of relevant attributes for studying freight demand given the general agreement on adopting the manufacturing firm as the relevant decision maker in determining the organization of the transport chain thus also influencing/determining freight mode choice. It was soon suggested to differentiate the analysis between shippers and receivers due to their different role and, possibly, preferences in determining a mode transport choice (Winston, 1981). In most cases, this issue can be overlooked considering that the shipper is well aware of the logistic requirements the receiver has and assumes them as relevant when making a transport choice.

The papers investigating shippers' preferences for freight service attributes that can aptly be divided in two groups, those explicitly considering transport mode as a relevant attribute (*between mode models*) and those who do not (*within mode models*). The first group represents the majority. To mention just a few, one could recall Bolis and Maggi (2003), Fosgerau (1996), Jiang et al. (1999), Shinghal and Fowkes (2002), Maggi et al. (2005) and Vellay and de Jong (2003). The second group include

Danielis et al. (2005) and Wigan, Rockliffe, Thoresen, and Tsolakis (2000). By not explicitly considering the transport mode used for freight transportation, these papers try to avoid potential correlation between mode and other attributes characterizing a specific mode.

Before proceeding with a detailed analysis of a specific set of papers selected to analyze and discuss the most important methodological characteristics and results obtained by the papers focusing on logistic managers' preference elicitation it is important to address some methodological issues of SP and RP when dealing with freight. The relevant and growing number of studies dealing with freight service attribute elicitation indicate a consistent variability both in term of model specification and model estimation. Consequently, the estimates of the attributes considered are not always reliable or comparable for practical issues such as, for instance, cost-benefit analysis especially if there is not a clear definition of what is exactly being measured. The object of measurement is strictly linked to who is actually interviewed. In fact, one assumes that the firm's preferences can, *de facto*, be correctly represented by those of the interviewee that holds a specific position in the company.

When comparing studies we are implicitly assuming that each person interviewed holding a managerial position (logistics manager) has the same authority in making decisions. Furthermore, one is also assuming that the logistic manager interviewed always and correctly represents either the shipper or the receiver. In fact, should not this be so, we are implicitly estimating different preferences that cannot and should not be compared.[7] The theoretical underpinnings of this relevant distinction date back to early 1980s (Winston, 1981) when the linkage between interviewing the shipper/receiver logistic manager and the Cost Insurance and Freight (CIF)/Free on Board (FOB) type of contract governing the freight service involved was clearly discussed. In fact, given a CIF contract most probably the relevant preferences that should be elicited are those of the shipper whereas in the case of a FOB denominated contract the receiver presumably plays a more relevant role.[8]

There are cases (see e.g., Bergantino & Bolis, 2008) where transport operators (e.g., freight forwarders) have been interviewed; in general, this approach does not seem totally appropriate if one assumes, as it seams reasonable, that transport operators are most likely interested in minimizing their own operation costs rather than the costs of the goods transported even if one can infer that the specific contractual agreements tend to transfer the final customer's preferences onto those of the forwarder. The lack of clarity concerning the issues raised in the last section might have a relevant impact on model estimates.

A further element inducing confusion and impeding comparability among results is the substantial heterogeneity in both SP and RP studies concerning the effective

7. Few are the studies that explicitly consider this distinction and report the differences between two groups in the estimates. A problem having a similar nature has recently been analyzed in a paper focusing on household location choice by Marcucci, Stathopoulos, Danielis, and Rotaris (2011).

8. These considerations are of minor importance in the case of own-account transport.

alternatives in relation to the shipments considered that are, usually, very diverse. In fact, while shipment can differ in terms of size, weight, value, distance, etc., the models estimated to predict choice generally assume that the attributes are uncorrelated to these aspects and are generically referred to a *typical* shipment that is defined by the interviewee in a preliminary part of the questionnaire.[9]

13.7. A Critical Review of SP/RP Freight Mode Studies

This section reports the results of a selection of papers dealing with the estimation of logistic managers' preferences for freight service attributes. The papers reviewed are selected according to a principle of relevance, data quality and geographical coverage. The list of the papers considered is not exhaustive and is biased toward the personal knowledge and likings of the author. Nevertheless, it should also be somewhat representative. The intent of the reviewing effort is to provide a common evaluation scorecard as to express an evaluation of the experimental design adopted, the models estimated, and the main results obtained. The aim is to provide an overall evaluation of the work done so far, of the clarity, completeness and transferability of the reported results and, finally, to indicate future research efforts needed to increase the knowledge of a highly complicated issue.

Table 13.A.1 (see Appendix 13.A) reports the results of the review performed on a set of 15 papers analyzed with the intent of evidencing the most important similarities and dissimilarities among the studies performed in terms of: attributes considered, attribute selection method, type of design, type of data elicited (i.e., SP, RP, and SP&RP), data collection period, sample dimension, localization of the study, number of alternatives presented to respondent, number of exercise repetitions, response format, questionnaire administration method, who was interviewed, freight sector considered, type of model estimated, and main results. All the articles considered have been published in well respected internationally rated academic journals.

With reference to the *attributes* used in the papers analyzed one sees that 3 is the minimum number of attributes considered and 51 the maximum with a mean of 8.4 attributes and a median of 6. With just one exception all papers include a monetary attribute (2 price and 12 transport cost). The most frequently used attributes are: time (12), frequency (7), reliability (7), flexibility (6). Other frequently used attributes are: risk of damage, risk of delay, safety, and inter-modality. From this analysis no homogeneous definition of the attributes considered clearly emerges. Furthermore, it is also quite difficult to compare apparently similar attributes among studies. To make an example: we have both a generic "reliability" attribute, a "time reliability" and "risk of delay"; in this case it is quite difficult to meaningfully compare the

9. A segmentation by shipment characteristics can potentially reduce the distortions induced by this generalization even if it seems that the solution proposed can only attenuate rather than solve the problem raised.

different estimates not only because we are not completely sure about the comparability of what is actually measured but also because the attributes are expressed in different unit of measurement.

In one third of the cases (5) the *experimental design* adopted is not clearly described and discussed and in 3 cases the experimental design is not a relevant information that needs to be provided (n.a.). For the 7 cases where the *experimental design* is clearly stated and discussed the prevailing format adopted is (3) a fractional factorial of one kind or the other (orthogonal, partial profile) and in only one case a d-efficient design was adopted. Given the high cost of acquiring information, which usually translates into a limited sample, this is not good news given the desirable properties of efficient designs. All the studies reviewed use SP data with two exceptions were RP data alone are used in one case and SP&RP data in another. The studies have been performed mostly in Europe (8), followed by Asia (3), North America (2) and Middle East (1).[10]

The papers reviewed have proposed choice sets with a number of alternatives ranging from 2 to 7 with an average of 2.9 and a median of 2. The most frequently used alternative is the road mode (13). Note that in some cases not only is the road mode specified but also the type of vehicle is reported (e.g., small truck *vs.* large truck) thus confirming the prevalence of this mode for transporting freight and the level of detail that in some cases is considered appropriate. Rail is used in 5 cases, combined or multimodal in 6, sea transport in 4 and intermodality in 2. The heterogeneity in the number of modes considered is generally greater than the number of alternatives presented to the interviewees. Given the heterogeneity in the alternatives presented the results obtained are not easily, if at all, comparable.

In the SP experiments only in 7 cases the number of repetitions proposed was reported. A minimum of 8 repetitions and a maximum of 40 was detected with an average of 15,9 and a median of 10. The relatively high number of repetitions is most probably due to the high cost, both monetary and in terms of time, of getting an interviewee to participate to the research. Therefore once the cooperation has been secured it is worthwhile to maximize the number of repetitions as to acquire a sufficient number of observations given the relatively small sample of participants.

The analysis of the response format adopted shows that in 6 cases choice was used while rating and ranking were respectively employed in 5 and 2 cases.[11] In the case of rating and ranking the information content of the data acquired is greater and can always be converted in choice equivalent information assuming the first rating or ranking alternative would be the chosen one. The benefits related to ranking and rating exercises do not come without costs. In fact, one has to recognize that ranking and/or rating rather than choosing among alternatives is usually more cognitively complicated. Furthermore one has also to recognize that the behavioral assumptions

10. There is one case where the geographical location of the study was not clearly specified.
11. In two cases, the response format was not clearly reported.

underlying both ranking and rating are more demanding with respect to a choice exercise.

Confirming the complexity of the questionnaire administration process and need for accurate information derived from personal interaction it is important to underline that in 7 cases the researchers adopted a Computer Aided Personal Interview (CAPI) interview and in only 2 cases Computer Aided Web Interview (CAWI) was used.[12]

With respect to who was actually interviewed one has to underline that in only 1 case this was not clearly stated. In 7 cases only 1 type of agent was interviewed, 2 types were interviewed in 4 cases and 3 types in 3 cases. Shippers are the agent-type most frequently interviewed. In other cases the specific role the interviewee plays in the decision-making process is not clear and if there is a logical compatibility between the person interviewed and the type of contract signed between the parties involved (i.e., FOB or CIF contracts). The general impression gained is that in most cases there is no clear-cut definition of the type of agent to be interviewed and this most likely induced a considerable level of heterogeneity in the data.

Great heterogeneity can also be found in the number and type of freight considered. Given that the estimates of the attributes considered are strictly influenced by type of freight transported the heterogeneity of the sectors considered does not help in claiming a potential generalization of the results obtained.

Finally, with reference to the models estimated one can safely assert that the most popular and frequently used model, no surprise, is the Multi-Nomial Logit (MNL) model (7 cases) followed by the Mixed Logit (MXL) model (4). In general there is a sufficient level of sophistication in the analysis of the data performed measured as the degree of flexibility in the treatment of both the deterministic and stochastic part of utility (see Marcucci & Gatta, 2012).

The main results of the papers analyzed overall stress the relevant role that qualitative attributes play in influencing choice thus underscoring the attention policy makers should pay to these attributes if more sustainable modal split are to be pursued effectively. Given the high heterogeneity in the studies analyzed no formal comparison of the Willingness to Pay (WTP) measures for different attributes is reported since they frequently refer to differently defined attributes, noncomparable contexts, and differ in the level of sophistication of the models estimated.

13.8. Unsettled Issues and Future Research

This section summarizes some still unsettled issues and proposes some future research themes to be addressed. The comparative evaluation of the papers selected (see Appendix 13.A) has shown that there is no overall consensus concerning the

12. In six cases, the paper did not explicitly report the administration procedure adopted.

experimental design adopted, the attributes used and the sample selection methods employed to mention just the most important issues. Some still unsettled issues emerge from the analysis and, in particular, it is worth recalling that:

1) there is no clear definition of the object of measurement;
2) most studies produce results that are very context specific and cannot easily be generalized;
3) samples, due to budget and participation constraint, are generally small and scarcely representative;
4) the high heterogeneity in shipment size, volume, etc. is usually ignored by assuming a generic reference to a "typical shipment" that is jointly defined together with the interviewee as the shipment that best represents, under all shipment characteristics, the most frequent shipment performed;
5) no explicit account is usually given of the either inward or outward movement of freight nor of the CIF or FOB nature of the contract;
6) the relative general obscurity concerning the experimental design adopted contributes to the non comparability of results whose differences can be attributed to a long list of nonauto-exclusive sources.

The general impression that emerges from the review conducted suggests that this field of research, notwithstanding the relevant efforts and high quality of the research performed, is still in its infancy. The investigation of logistics managers' preferences is both relevant and complex so it is worth focusing on this issue in future research endeavors. As for freight transport in general there are some structural difficulties that the applied researcher has to tackle. There are some steps researchers need to take quickly and jointly if a deeper and more reliable knowledge is to be pursued in order to have a wider and more stable information base to use for effective and efficient policy interventions. In particular, specularly to the critical and still unsettled points previously mentioned, one could suggest the following to foster substantial improvements:

1) univocally define research and reporting protocols;
2) motivate and define who has to be interviewed in the specific freight context studied;
3) clearly report the contractual relationships governing freight movements (e.g., CIF or FOB);
4) clearly report freight details (e.g., volume, value, and weight);
5) motivate the attribute selection method used.

In summary, a general suggestion is to systematize the procedures and reporting adopted in performing applied research by introducing a much higher level of detail and rigor in both defining the object of measurement and in the experimental design protocols employed. A standardization of the fundamental methodological decisions to be taken will help to acquire a wider, more reliable and transferable knowledge. In fact, a general agreement on the basic methodological choices would help in simultaneously pursuing two very important objectives; on the one side, it would help to consolidate research endeavors throughout the world and, at the same time, leave enough room of maneuver to the researcher so to taylor research to the specific contextual needs. Concerning the first point, it is important to recall that, given the financial and participation constraints often characterizing this stream of research, it is extremely important to guarantee the basic comparability of results obtained in different circumstances to acquire more reliable and statistically representative information. In conclusion, it seems reasonable to suggest launching a vast research project where the some time contrasting objectives previously discussed could be adequately treated. Notwithstanding the difficulties inherently characterizing such a project the potential benefits would, by far, outweigh the costs.

13.9. Conclusions

After defining and justifying the relevance of studying logistics managers' preferences of freight service attribute in a globalized economy, the chapter has described the classical approach to freight modeling and a new emerging behavioral approach to clarify the potential improvements deriving from an improved knowledge in this field. Subsequently, a succinct retrospective analysis on SP freight demand analysis is carried out in order to conceptually define the research environment within which the specific literature review is found. The review of the literature on logistics managers' preferences for freight service attributes has provided a comparative analysis of a set of (nonexhaustive) papers based on a number of indicators with both methodological and practical implications. The picture that emerges is not completely reassuring but, at the same time, it also suggests, given the relevance of the research stream, a great potential for improving our knowledge and policy implementation capabilities. In fact, the review has underlined some weaknesses due to:

1) lack of clarity and homogeneity in the definition of the object of research;
2) lack of rigor in delimiting the contextual characteristics of research;
3) limited representativeness of the samples used;
4) uncomparable and unclear research reporting thus hindering potentially useful generalizations and transferability of results.

But, specularly, should these issues be adequately tackled and solved the potential benefits would be substantial. In particular, it seems appropriate to suggest launching an all encompassing project so to assure a commonly shared research and reporting protocol in this field with the intent of guaranteeing a minimum level of comparability while allowing for the customization of research given the specific contextual needs. Participation to such a research project can only be imagined on a voluntary base while it is advisable to have the largest possible number of influential researchers participate and collaborate. Notwithstanding the difficulties inherently characterizing such a project the potential benefits would, most likely, outweigh the costs.

Acknowledgments

The author would like to thank Alessandro Sforna for bibliographical research assistance and for helping with the typesetting of this article.

References

Abdelwahab, W., & Sargious, M. (1992). Modelling the demand for freight transport. *Journal of Transport Economics and Policy*, *26*(1), 49–70.

Arunotayanun, K. (2009). *Modelling freight supplier behaviour and response*. Ph.D. thesis, Imperial College of London.

Arunotayanun, K., & Polak, J. (2011). Taste heterogeneity and market segmentation in freight shippers' mode choice behaviour. *Transportation Research Part E*, *47*(2), 138–148.

Ben-Akiva, M., & Bierlaire, M. (1999). Discrete choice methods and their applications to short-term travel decisions. In R. Hall (Ed.), *Handbook of Transportation Science* (pp. 5–34). Norwell, MA: Kluwer.

Bergantino, A., & Bolis, S. (2008). Monetary value of transport service attributes: Land versus maritime roro transport. An application using adaptive stated preference. *Maritime policy and management*, *35*(2), 159–174.

Beuthe, M., & Bouffioux, C. (2008). Analysing qualitative attributes of freight transport from stated orders of preference experiment. *Journal of Transport Economics and Policy*, *42*, 105–128.

Blauwens, G. (2001). De keuze tussen vervoerswijzen op basis van het total cost concept. *Vervoerswetenschap*, *3*(10), 28–32.

Bolis, S., & Maggi, R. (2003). Logistic strategy and transport service choices: An adaptive stated preference experiment. *Growth and Change*, *34*(4), 490–504.

Cascetta, E. (1997). National modelling in Italy: simulation and evaluation models for the Italian dss. Paper presented at the seminar on national transport models: the state of the art, Noordwijk.

Consortium ASTRA. (2000). *Final report: assessment of transport strategies*. Technical report, University of Karlsruhe, Karlsruhe.

Danielis, R., Rotaris, L., & Marcucci, E. (2005). Logistics managers stated preferences for freight service attributes. *Transportation Research part E: Logistics and transportation review*, *41*(3), 201–215.

Danielis, R., Rotaris, L., & Marcucci, E. (Forthcoming). Whom to interview when studying freight transport mode choice? Evidence from a road vs. rolling motorway study. In E. Marcucci & Puckett, S. (Eds.), *Special issue of the Journal Research in Transportation Economics*.

De Jong, G., Gommers, M., & Klooster, J. (1992). *Time valutation in freight transport: Methods and results*. In Proceeding of PTRC summer annual meeting, Manchester.

De Jong, G., Gunn, H., & Walker, W. (2004). National and international freight transport models: An overview and ideas for future development. *Transport review*, *24*(1), 103–124.

De Jong, G., Vellay, C., & Moue, M. (2001). *A joint sp/rp model of freight shipments from the region nord-pas de Calais*. In Proceedings of the *2001* European transport conference, Cambridge.

FHWA. (1999). *Guidebook on statewide travel forecasting*. Report for the FHWA, Centre for Urban Transportation Studies, University of Wisconsin-Milwaukee and the Wisconsin Department of Transportation.

Fosgerau, M. (1996). *Freight traffic on the storebaelt fixed link*. In Proceedings of seminar, *24*th European transport forum, London, UK.

Fries, M., & Patterson, Z. (2008). *Carrier or mode? The dilemma of shipper choice in freight modelling*. In Proceedings of the *8*th Swiss transport research conference, Monte Verit, CH.

Garrido, R. (2000). Spatial interaction between truck flows through the Mexico-Texas border. *Transportation Research part A*, *33*, 23–33.

Harker, P. (1985). The state of the art in the predictive analysis of freight transport system. *Transport Reviews*, *5*(2), 139–153.

Hensher, D., & Puckett, S. (2005). Refocusing the modelling of freight distribution: Development of an economic-based framework to evaluate supply chain behavior in response to congestion charging. *Transportation*, *32*, 573–602.

Hesse, M., & Rodrigue, J. (2004). The transport geography of logistics and freight distribution. *Journal of transport geography*, *12*, 171–184.

Holter, A., Grant, D., & Ritchie, J. (2010, October). *Long-range transport: Speeding up the cash-to-cash cycle International journal of logistics: Research and applications*, *13*(5), 339–347.

Jiang, F., Johnson, P., & Calzada, C. (1999). Freight demand characteristics and mode choice: An analysis of the results of modelling with disaggregate revealed preference data. *Journal of transportation and statistics*, *2*, 149–158.

Jovicic, G. (1998). *Application of models based on stated and revealed preference data for forecasting Danish international freight transport*. Retrieved from http://www.trafikdage.dk/td/papers/papers98/trafikmodeller/g_jovicic/g_jovicic.pdf

Kofteci, S., Ergun, M., & Serpil Ay, H. (2010). Modeling freight transportation preferences: Conjoint analysis for Turkish region. *Scientific Research and Essays*, *5*(15), 2016–2021.

Levinson, M. (2006). *The box: how the shipping container made the word smaller and the word economy bigger*. Princeton: Princeton University Press.

Liedtke, G. (2009). Principles of micro-behavior commodity transport modelling. *Transportation Research part E: Logistics and Transportation Review*, *45*, 95–809.

Lobo, B. (2010). Cross-boundary container truck congestion: The case of the Hong Kong-pearl river delta region. *Transportation*, *37*, 257–274.

Maggi, R., Bolis, S., Rudel, R., Stoppa, R., Ruesch, M., & Abel, H. (2005). *Evolution of quality attributes in freight transport*. Report, Swiss Federal Roads Office, Bern, Switzerland.

Maier, G., & Bergman, E. (2002). Conjoint analysis of transport options in Austrian regions and industrial clusters. In R. Danielis (Ed.), *Freight transport demand and stated preference experiments*. Milan: Franco Angeli.

Manheim, M. L. (1979). *Fundamentals of transportation system analysis*. Cambridge, MA: The MIT Press.

Marcucci, E. (Ed.), (2005). *Introduzione ai modelli a scelta discreta. Teoria, metodi e applicazioni*, Carocci Editore, Roma, Italy.

Marcucci, E., & Gatta, V. (2012). Dissecting preference heterogeneity in consumer stated choices. *Transportation Research Part E*, *48*, 331–339.

Marcucci, E., Stathopoulos, A., Danielis, R., & Rotaris, L. (2011). Comparing single and joint preferences: A choice experiment on residential location in three member householders. *Environment and Planning Part A*, *43*, 1209–1225.

McFadden, D., Winston, C., & Boersch-Supan, A. (1985). Joint estimation of freight transportation decisions under nonrandom sampling. In A. Daughety (Ed.), *Analytical studies in transport economics* (pp. 137–157). Cambridge: Cambridge University Press.

McKinnon, A. C., & Woodburn, A. (1996). Logistical restructuring and road freight traffic growth. *Transportation*, *23*, 141–161.

Meersman, H., & Van de Voorde, E. (2008). The relationship between activity and freight transport. In M. Ben Akiva, H. Meersman & E. Van de Voorde (Eds.), *Recent development in transport modelling* (pp. 69–92). Bingley: Emerald Group.

Menendez, L. G., Martinez Zarzoso, I., & Pinero De Miguel, D. (2004). Determinants of mode choice between road and shipping for freight transport. *Journal of Transport Economics and Policy*, *38*(3), 447–466.

Norojono, O., & Young, W. (2003). A stated preference freight mode choice model. *Transportation Planning and Technology*, *26*(2), 195–212.

Nuzzolo, A., & Russo, F. (1995). *A disaggregate freight modal choice model*. In Proceeding of the 7th WCTR conference, Sydney.

Ogden, K. W. (1992). *Urban goods movements: A guide to policy and planning*. Aldershot, UK: Ashgate.

Patterson, Z., Ewing, G., & Haider, M. (2007). Shipper preferences suggest strong mistrust of rail: The first sp carrier choice survey for the Quebec City-Windsor Corridor. *Transportation Research Record*, *67–74*.

Puckett, S., Hensher, D., Brooks, M., & Trifts, V. (2011). Preferences for alternative short sea shipping opportunities. *Transportation Research Part E*, *47*(2), 182–189.

Regan, A., & Garrido, R. (2002). Freight demand and shipper behavior modeling: State of the art, directions for the future. In D. Hensher & J. King (Eds.), *Travel behavior research: The leading edge* (pp. 185–216). Oxford: Pergamon press.

Shinghal, N., & Fowkes, A. (2002). Freight mode choice and adaptive stated preferences. *Transportation research, Part E*, *38*(5), 367–378.

Swait, J., Louviere, J., & Williams, M. (1993). A sequential approach to exploiting the combined strengths of sp and rp data: Application to freight shipper choice. *Transportation*, *21*, 125–152.

Tavasszy, L. (1994). *Characteristics and capabilities of Dutch freight transportation systems models*. Report Report MR-382-EAC/VW, RAND Europe.

Tavasszy, L., Van Der Vlist, M., & Ruijgrok, C. (1998). *Scenario-wise analysis of transport and logistic system with a smile*. In Proceeding of 8th WCTR conference, Antwerp.

Transek. (1990). *Godskunders varderingar*. Report 9, Banverket Report, Göteborg, Sweden.

Vellay, C., & de Jong, G. (2003). *A simulation sp/rp analysis of mode choice in freight transport in the region nord–pas de Calais*. Report, French Ministry for Equipment, Housing, Transport and Tourism, Paris, France.

Widlert, S., & Bradley, M. (1992). *Preferences for freight services in Sweden*. In Proceeding of WCTR conference, Lyon.

Wigan, M., Rockliffe, N., Thoresen, T., & Tsolakis, D. (2000). Valuing long-haul and metropolitan freight travel time and reliability. *Journal of Transportation and Statistics*, *3*(3), 83–89.

Winston, C. (1981). A disaggregate model of the demand for intercity freight transportation. *Econometrica*, *49*(4), 981–1006.

Witlox, F., & Vandaele, E. (2006). Determining the monetary value of quality attributes in freight transportation using a stated preference approach. *Transportation Planning and Technology*, *28*(2), 77–92.

Zamparini, L., & Reggiani, A. (2007). Freight transport and the value of travel time savings: A meta-analysis of empirical studies. *Transport Reviews*, *27*(5), 621–636.

Zlatoper, T. J., & Austrian, Z. (1989). Freight transportation demand: a survey of recent econometric studies. *Transportation*, *16*, 27–46.

Appendix 13.A

Table 13.A.1: Articles reviewed.

Articles	Arunotayanun and Polak (2011)	Beuthe and Bouffioux (2008)	Bolis and Maggi (2003)
Attributes considered	Delivery time, transport cost, measures of logistical and transport services	Cost service frequency, transport time, reliability, flexibility, safety	Price, time reliability, mode, frequency, flexibility
Type of design	n.s.	Fixed orthogonal fractional factorial	n.s.
Attribute selection method	Literature analysis and researchers' s own decision	Literature review, authors' autonomous choice	n.s.
Type of data (SP, RP, SP&RP)	SP	SP	SP
Data collection period	1998–1999	n.s	n.s
Sample dimension	186	133	22
Geographical reference	Indonesia (Java)	Belgium	Italy and Switzerland (Frejus, Gotthard, Brenner)
Number of alternatives	3	35	2
Which alternative were considered	Small truck, large truck, rail	Rail, road waterway, short-sea shipping and their inter and multimodal combinations	Hypothetical road based, hypothetical rail based
Number of repetitions	8	n.a.	40
Response format	Choice	Ranking	Rating (LASP)
Questionnaire administration method	n.s.	Face to face	Face to face CAPI

Who was interviewed	Shippers	Transport managers (logistic operators and forwarders)	Logistics or distribution managers
Goods sectors considered	Non-perishable food, leather, textile, electronics	Foodstuff, minerals, metal products, chemicals pharmaceuticals, miscellaneous	Perishable goods, chemical, machinery, metallurgy, miscellaneous
Types of models estimated	MNL, MXL, LC	Conditional logit	MNL, Tobit
Main results	Significant taste heterogeneity among shippers for service attribute	Different results for different portions of the sample and substantial dispersion of results with respect to other published studies	Firms do not like the hypothetical alternatives tested; preference cannot be easily classified by type of goods

Table 13.A.1 (*Continued*)

Articles	**Danielis et al. (forthcoming)**	**Danielis et al. (2005)**	**Jiang et al. (1999)**
Attributes considered	Travel time, departure time, cost, day of the week departure, number of sleeping places per compartment, highway toll	Cost, travel time, punctuality, damage and loss	Quantitative and qualitative—subdivided in shipper's or receiver's characteristics; goods' physical attributes; spatial and flow characteristics of shipments
Type of design	Deficient design	Fractional factorial partial profile (ACA)	n.a.
Attribute selection method	Literature review, authors autonomous choice	Literature analysis	n.a.
Type of data (SP, RP, SP&RP)	SP	SP	RP
Data collection period	2011	n.s.	1988
Sample dimension	60 pilot	65	3473
Geographical reference	Italy (Trieste), Ukraine (Chop)	Italy (Friuli Venezia Giulia, Marche)	France (whole country)
Number of alternatives	3	2	4
Which alternative were considered	Small truck, large truck, rail	No mode specification is provided	Private, road, combined, rail
Number of repetitions	10	n.s.	n.a.
Response format	Choice	Rating	n.a.

Questionnaire administration method	Face to face CAPI	Face to face CAPI	n.s.
Who was interviewed	Truck drivers, forwarding and transport operators	Logistic managers	Industrial and commercial firms
Goods sectors considered	n.s.	Chemical and fibers; machine; furniture; wood, textile and clothing, construction, paper and paper products, food and tobacco, electric and electronic equipment, metal parts	20 economic sectors
Types of models estimated	MNL, ECM, MXL	Ordered Probit	MNL, NL
Main results	Similar preference for cost and travel but dissimilar for departing day, highway toll, attitude toward RoMo	Strong preference for reliability and safety; modal shift is not foreseeable	The attributes explaining rail and combined transport are: distance, shipper's accessibility to infrastructure, availability of own transport facilities and shipment packaging

Table 13.A.1 (*Continued*)

Articles	**Jovicic (1998)**	**Kofteci, Ergun, and Serpil Ay (2010)**	**Lobo (2010)**
Attributes considered	Cost, transport time, risk of damage, risk of delay, frequency, information system, flexibility	Cost, time reliability, damages, and losses	Road toll, time reliability, route length, customs time
Type of design	n.s.	Fixed orthogonal fractional factorial	n.s.
Attribute selection method	Literature analysis	Literature analysis	Researcher's choice
Type of data (SP, RP, SP&RP)	RP/SP	SP	SP
Data collection period	1996–1997; 1993	n.s.	2005–2006
Sample dimension	1012(RP)	50	230
Geographical reference	n.s.	Turkey (Antalya)	China (Hong Kong)
Number of alternatives	4	n.s.	7
Which alternative were considered	Lorry transport, conventional rail, combined road-rail, sea	Road intermodal	Different customs check points
Number of repetitions	10	n.s.	n.s.
Response format	n.s.	Rating	Choice
Questionnaire administration method	n.s	Face to face CAPI	Face to face

Who was interviewed	n.s.	Transport managers	Container truck drivers, goods owners, others.
Goods sectors considered	Vegetable products, animal products beverages, feeding stuffs, metal products, paper and wood products	Cement	n.s.
Types of models estimated	Hierarchical MNL	MNL	MNL, NL
Main results	Transport cost is the most important choice variable; transport time is more important that the rest of the remaining variables	Road vs. intermodal transport is perceived as better with the exception of cost; reliability of service is also very important	Preference of great heterogeneity; need to focus on factory owners and logistic suppliers

Table 13.A.1 (*Continued*)

Articles	**Maier and Bergman (2002)**	**Menendez, Martinez Zarzoso, and Pinero De Miguel (2004)**	**Norojono and Young (2003)**
Attributes considered	Cost, time, reliability, frequency, flexibility, mode	Cost, transit time, damage, delay, frequency, distance, environment, consolidation	Transport cost, delivery time, quality, flexibility
Type of design	n.a	n.s.	Orthogonal four level factor set, hierarchical stated preference.
Attribute selection method	Researcher's choice	n.s.	Preliminary survey and literature review
Type of data (SP, RP, SP&RP)	SP	SP	SP
Data collection period	n.s.	1998	1998–1999
Sample dimension	n.s.	157	186
Geographical reference	Austria	Italy and Spain (Valencia Community)	Indonesia (Java)
Number of alternatives	2	2	3
Which alternative were considered	Road, rail, water	Road, sea	Small truck, large truck rail
Number of repetitions	20	n.s.	8
Response format	Rating	Choice	Choice

Questionnaire administration method	n.s.	Face to face	n.s.
Who was interviewed	Logistics managers	Exporting firms, freight forwarders, other transport firms	Shippers
Goods sectors considered	Motor vehicles, chemical, pharmaceutical, construction material, electronics, food, metal working, wood paper, manufacturing	Wood manufacturing and furniture, ceramics, textiles, agroindustry	Nonperishable food, leather, textile, electronics
Types of models estimated	n.s.	CL	HEV
Main results	Logistic managers show different preferences in different regions and industry clusters; the most important attributes is reliability	Cost, travel time and frequency are significant, sea-transport is more sensitive than road transport to variation in its own cost changes	Quality and flexibility of service highly influence the mode chosen, safety and reliability also important

Table 13.A.1 (*Continued*)

Articles	**Patterson, Ewing, and Haider (2007)**	**Puckett, Hensher, Brooks, and Trifts (2011)**	**Witlox and Vandaele (2006)**
Attributes considered	Cost, on-time reliability, damage risk, security, intermodality	Price of service, total transit time, frequency	Cost time, loss, frequency, reliability, flexibility
Type of design	Random factorial design	n.s.	Orthogonal
Attribute selection method	Literature review, focus group (65 face to face)	n.s.	Researchers' personal Choice
Type of data (SP, RP, SP&RP)	SP	SP	SP
Data collection period	2005 (August–December)	2006	n.s.
Sample dimension	392	42	88
Geographical reference	Canada	North America (Canada and USA)	Belgium (Antwerp and Ghent)
Number of alternatives	3	2	25
Which alternative were considered	Truck, intermodal	Truck, short sea	Multimodal, inland navigation, road
Number of repetitions	n.s.	15	n.a.
Response format	Choice	Rating	Ranking
Questionnaire administration method	Web based	Internet	n.s.

Who was interviewed	Shippers	Shippers	Transport managers (industrial firms and logistics service providers)
Goods sectors considered	All sectors	n.a.	Textiles, steel, cooling, machines, plastics
Types of models estimated	MNL, MXL	MNL, SMXL	MUSTARD (Multi Criteria Utility Based Stochastic Aid for Ranking Decisions)
Main results	Reduced bias against intermodal shipping on longer-distance shipments but still strong unexplained bias against using rail	WTP a premium for frequency, weak preference for integrated short sea shipping mode	Shipper's modal choice is influenced not only by time and costs but also by frequency, reliability, flexibility, transport duration and risk of loss and damages

Appendix 13.B

List of abbreviations and acronyms

- ACA = adaptive conjoint analysis
- CAPI = computer-assisted personal interview
- CAWI = computer-assisted web interview
- CIF = cost insurance and freight
- CL = conditional logit
- ECM = error component logit
- FOB = free on board
- GMXL = generalized mixed logit
- HEV = heteroscedastic extreme value
- LASP = Leeds adaptive stated preferences
- LC = latent class
- MXL = mixed logit
- MNL = multinomial logit
- NL = nested logit
- RP = revealed preferences
- SMXL = scaled mixed logit
- SP = stated preferences

Chapter 14

Capacity Utilisation of Vehicles for Road Freight Transport

Megersa A. Abate and Ole Kveiborg

Abstract

This chapter discusses a central aspect of freight transportation – capacity utilisation with a link to empty running of commercial freight vehicles. It provides an overview of the literature on these topics and groups the contributions into two segments according to their analytical approach and origin of research. The first approach looks at utilisation based on economic theories such as the firms' objective to maximise profitability and considers how various firm and haul (market) characteristics influence utilisation. The second approach stems from the transport modelling literature and its main aim is analysing vehicle movement and usage in transport demand modelling context. A strand of this second group of contributions is the modelling of trip-chain and its implication on the level of capacity utilisation. A key lesson from the reviewed studies is that it is important to take into account the commercial activity that initiates vehicle movements to evaluate performance. It appears that there is room for further enrichment of the modelling exercise by incorporating information regarding the operator to give a stronger behavioural basis for the vehicle movements and utilisation analysis.

Keywords: Overview; capacity utilisation; load factors

14.1. Introduction

Road freight transport is vital for most industries to ensure smooth movement of goods. Moreover, the transport sector employs a large number of the labour force

Freight Transport Modelling

ISBN: 978-1-78190-285-1

and contributes with a rather large share to GDP in most developed countries (American Trucking Trends, 2011; Transport Canada, 2003). Thus, the level of capacity utilisation of vehicles used for freight transport is an important indicator of how well economic resources are used both from the perspective of vehicle operators and other sectors of the economy, which rely on their service. Despite its positive contributions, freight transport leads to negative externalities that need to be reduced, ideally, not at the expense of economic prosperity. If capacity utilisation can be improved then it will be possible to reduce the amount of vehicle kilometres required to satisfy the demand for goods movements and the related detrimental effects. It is, therefore, interesting to investigate the factors behind capacity utilisation and how it influences overall travel demand.

The observed level of capacity utilisation is naturally the consequence of some optimisation by the transport operators due to, for example, cost minimisation, and optimal resource use. This optimisation is taken as given for the analysis of capacity utilisation. So the choices influencing capacity utilisation as is described here are implicitly the result of the underlying economic optimisation of the firm.

In this chapter we study capacity utilisation and present literature contributions on the subject from different approaches that has followed independent paths. We will shed some light on its measurement and its determinants. In particular, we focus on how the different approaches improve current methodologies within freight transport modelling as well as within economic modelling of the vehicle operators' decisions. There are several studies ranging from the economics to the transport literature that analyse the issue of vehicle capacity utilisation. These studies usually define capacity utilisation as a function of the physical dimensions of a load such as weight and volume. Each study emphasises different and yet key aspects of utilisation from which interesting insights can be gained.

Studies from the economics literature analyse the underlying determinants of capacity utilisation giving particular weight to cost minimisation and other economic incentives as the main behavioural issues behind resource allocation decisions (Beilock & Kilmer, 1986; Wilson & Beilock, 1994; Wilson & Dooly, 1993; Winston, 1985). They show how various haul (trip), vehicle and carrier (operator) characteristics affect the level of utilisation. According to the economics literature there is a persistent challenge of matching capacity with demand as a result of two factors: freight movement imbalance between regions and market access cost differential between operators. Studies from this strand of the literature show how information technology capabilities improve utilisation by facilitating the matching process (e.g. Barla, Bolduc, & Boucher, 2010; Hubbard, 2003).

Studies from the transport literature mainly focus on modelling vehicle movements as part of freight demand modelling and a vehicle routing problem with the aim of predicting directional traffic. One frontier interest of these studies is the issue of 'trip chain' or 'tour' undertaken by freight vehicles in urban areas and its implication on efficiency (Holguín-Veras & Thorson, 2003; Raothanachonkun, Wisetjindawat, & Matsumoto, 2007). Empty runs or the number of kilometres driven without a load is often taken as the main indicator of capacity utilisation in this literature. However, some studies (Figliozzi, 2007; Figliozzi, Kingdon, & Wilkitzki,

2007) show that the extent of empty running is a poor proxy for efficiency of vehicle movements, unless the purpose of commercial activity which initiated the movement is taken into considerations.

The key lesson from both strands of the literature is that capacity utilisation varies depending on the specific setting in which vehicles are used. The extraordinary heterogeneity in terms of weight and volume of loads, direction and distance of movement as well as time window constraints results in varying degree of utilisation even for rather identical vehicles. Future studies should focus on micro level vehicle utilisation rather than a general level analysis. Empirical analysis showing the scope of potential gains from improved utilisation can also be an interesting addition to this important topic.

The rest of the chapter is organised as follows. The next section gives an overview of the concept of capacity utilisation and production technology in trucking as it is discussed in the literature. This is followed by a section where determinants of capacity utilisation are discussed. 'Trip chain' and capacity utilisation issues are presented in the penultimate section followed by concluding remarks in the final section.

14.2. Capacity Utilisation and Production Technology in Trucking

14.2.1. Production Technology

Capacity utilisation is broadly defined as the ratio of actual output to the maximum potential output consistent with a given capital stock (Nelson, 1989). It may have specific definitions depending on an economic activity under consideration. However, attaining the maximum possible output out of a given set of inputs is the universal objective in all definitions. This definition focuses on physical productivity rather than economic productivity since price and cost variables are not directly considered.

For trucking activity, capacity utilisation is defined along the lines of its physical meaning. Aspects such as distance of movement as well as weight, volume and economic value of a cargo carried by a vehicle are considered independently or simultaneously to evaluate utilisation of vehicle capacity. The extent of empty running, fuel efficiency and quality of service with regard to time performance also describe how well a vehicle is used, the last two dimensions have been used in a few cases in the literature.

In trucking, the physical part of production involves a simple process of hauling a load from one point to another by a vehicle and a driver. The production technology is typically a simple Leontief type where a labour (driver) and capital (vehicle) combination is needed to produce an output (hauling a load). A driver never drives two vehicles simultaneously, or two drivers (almost) never operate a vehicle at the same time. Loading capacity of a vehicle, speed and hours of service are also limited with an implication on attainable level of productivity. As pointed out by Boyer and Burks (2009), the scope for getting more output (ton-kilometres) per truck per year by the classic manufacturing method-substituting capital for labour through automation-appears to be very limited in road freight transport. Therefore, in order to use a vehicle-driver combination to produce a maximum level of output, it has to be driven over long distance utilising the maximum loading capacity of the vehicle.

A unique feature of freight transport services is that it is by nature a joint-product. A vehicle movement from origin 'A' to destination 'B' (a front-haul direction) is usually followed by a reverse movement from 'B' to 'A' or to the base of operation of the vehicle if it is different from 'A' (a backhaul direction). Unlike passengers, freight rarely returns to its origin implying less-than optimal vehicle utilisation during backhauls (return trip). This problem is referred to as 'the back-haul problem' in the literature (see e.g. Boyer, 1998; Demirel, van Ommeren, & Rietveld, 2010; Felton, 1981; Wilson, 1987). Hence, optimal utilisation of capacity availed to distinct directions depends on the extent to which trucks are moving loaded between them. Joint transport services result in cost complementarities forcing firms to serve multiple markets to minimise cost (Beilock & Kilmer, 1986).

14.2.2. *Output Measures and Capacity Utilisation*

Measuring the capacity utilisation of a truck mainly involves consideration of two dimensions: spatial attributes of a truck movement and the physical attributes of a shipment (i.e. for a given capital and labour input of the firm). Both dimensions can be considered independently or they can be considered simultaneously. The location of origin-destination pairs (ODs) and the distance between them constitute the main component of the spatial dimension. As for the physical attributes, the weight and the volume (density) of a shipment essentially determine how well a truck can be used relative to its maximum carrying capacity. To characterise capacity utilisation of a truck we note that different pictures emerge depending on how the output of a freight transport service is defined. The following discussion illustrates this point.

Hubbard (2003) discusses two concepts of capacity utilisation. The first concept relates to the share of 'loaded kilometres', defined as the numbers of kilometres trucks are driven with a load during the periods trucks are in operation away from their base. Seen this way, the performance of a truck can be evaluated by the share of 'loaded kilometres' relative to the total kilometres the truck is driven. The second concept considers the number of times trucks are in use in a given period. For instance, trucks that are driven more weeks in a year are considered to have a higher level of capacity utilisation than those used for a fewer number of weeks due to frequent maintenance or lack of demand. It is assumed here that being away from base of operation more often implies higher level of utilisation. Such temporal consideration of truck usage is not common in the literature, but it adds an interesting perspective

Using these two concepts on Danish data from 1999 to 2009 shows developments depicted in Figure 14.1. As mentioned above, in many freight transport models the average load on vehicles is used as a measure of utilisation in order to calculate number of vehicles. This measure is also shown in Figure 14.1.

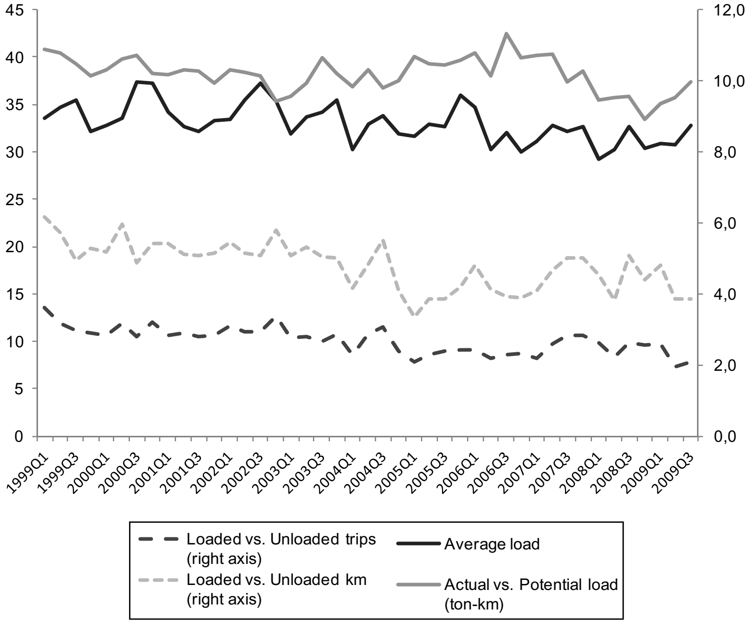

Figure 14.1: Different measures of capacity utilisation on Danish freight transport each quarter from 1999 to 2009.

An important dimension overlooked in Hubbard's discussion is the attributes of the load carried by trucks. Failing to consider weight and/or volume of the load can be misleading since utilisation does depend also on trucks' carrying capacity that is filled with cargo or the level of the load factor. According to Hubbard's first definition, two identical trucks carrying different size of otherwise identical loads are reported to have equal level of utilisation in terms of 'loaded kilometres' performance if they travel over the same distance. Their performance in terms of the level of the load factor, however, is clearly different, which can be seen from the comparison of these different performance measures outlined in Figure 14.1.

An informative comparison of utilisation can be made using ton-kilometres (TKM) as an output measure. TKM is defined as the product of cargo weight carried and the distance over which it is shipped. Boyer and Burks (2009) use TKM to define capacity utilisation (productivity) at a vehicle level as the annual ton-kilometres per truck-and-driver combination. As pointed out earlier, to have higher utilisation a vehicle has to be driven over long distance and loaded to its maximum capacity. However, unless both dimensions-distance and weight-are taken into account it is difficult to fully capture utilisation. Moreover, TKM does not consider the type of transport service undertaken since the more kilometres a vehicle is running the better the capacity utilisation is. For instance, if a truck is used for a specific type of delivery service, where loading and un-loading takes extended time then the truck will run only few kilometres and will not be able to have large TKM. But it may be fully loaded when it is operating. To overcome this problem, capacity utilisation can also be defined as the extent of the actual versus the maximum laden capacity of the vehicles on the trips that they have actually performed. Such a definition using TKM is also shown in Figure 14.1.

Another important dimension, volume of a load, is missing when one looks only at TKM performance. Light and voluminous items fill up physical space of a vehicle before its maximum laden capacity is reached. Vehicles carrying such items will appear less productive compared to those moving less voluminous and heavy item over the same distance. In relation to this, it may also be important to consider the specific type of good being carried. Some goods such as dangerous goods, and oil products require particular types of trucks and should therefore be considered specifically. The same may of course be the case for other goods although these may be carried in normal articulated trucks (e.g. food and general cargo). Ideally, one needs to consider the distance travelled by the vehicle, the specific type of good, weight and volume of the load carried to make a complete comparison of capacity utilisation of vehicles.

As indicated by Fernie and McKinnon (2003), quality of service with regard to 'time utilisation' and deviation from schedule as well as fuel efficiency should also be quantified to know how a truck's hauling capacity is effectively utilised.

Finally, note that the above definitions of capacity utilisation take into account only the physical (engineering) aspect of freight transport service since their focus is on vehicles rather than firms. In general no single capacity measure can accommodate all of these 'requirements'. It is thus important to be specific about what is the objective of the measure is (e.g. to be usable for political decisions).

14.2.3. Empty Running and Capacity Utilisation

Empty running is another important indicator of the extent of capacity utilisation. It arises when carriers provide capacity to several locations (markets segments) and choose to access only some of them due to unavailability of load or vehicle routing decision. A vehicle moving in a round trip between two locations may choose to carry a load on the front-direction and run empty during backhaul or vice-versa.

For a vehicle moving in a more complex pattern, empty runs occur at several stages of its journey. An illustration of this is included in Figure 14.2. To make it simple, we assume that all goods going from a production (P) to a consumption (C) pass through two distribution centres (DC), where the goods are consolidated and only truck is used. In the figure we consider three 'supply chains' one going from P1 to C1, another going from P2 to P3 and a third going from P3 to C3. The chain P1C1 passes through distribution centres DC1 and DC2, the chain P2C2 passes through DC3, DC4 and DC5, while the chain P3C3 passes through DC5 and DC1.

Each leg of the supply chains is performed by a distinct vehicle. However, in the figure we follow a single truck in a trip chain starting in DC1 going to DC2 (loaded), from DC2 to DC3 (un loaded), from DC3 to DC4 (loaded), from DC4 to DC5

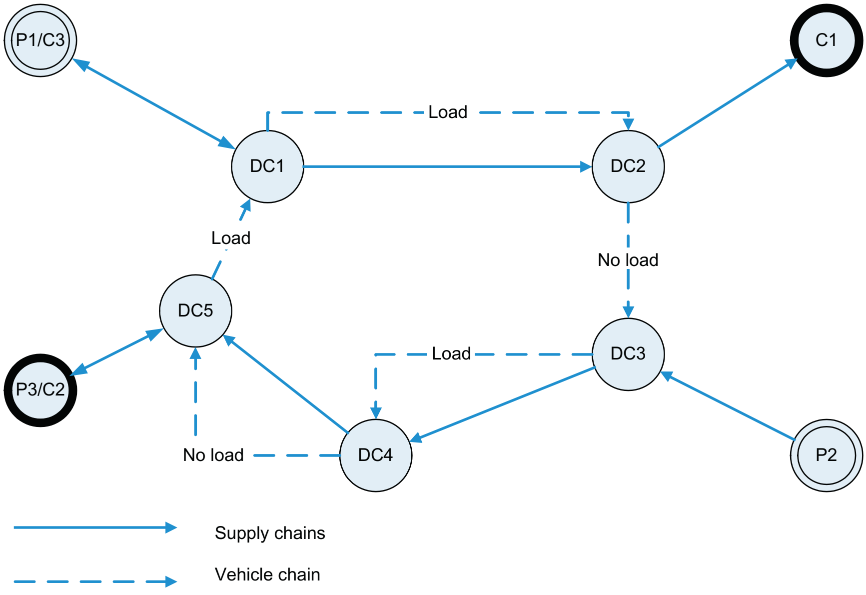

Figure 14.2: Complex trip chains and the related supply chains.

without load, and finally from DC5 to DC1 with load. The trip chain consists of 5 individual trips determined by separate supply chains and a trip back to the origin for the trip chain.

We may also observe simpler cases where there are one to one relations between supply and trip chain. For example, when a consignment is conveyed directly from P to C and where the vehicle return to its origin loaded or (more likely) empty.

The figure reveals that although vehicle trips are derived as part of the supply chain(s), the supply chains cannot describe all observed trips by trucks and they cannot always determine the actual use of vehicles. Empty trips are more likely related to the trip chain rather than the supply chains. Moreover, utilisation of vehicles is also related to the trip chain and to a large extent also to loaded and un-loaded trips. For example, if a truck owner must decided between undertaking an empty trip to go to a place, where a full load can be picked up of whether he chooses to take a smaller consignment from where the truck has made its delivery. The two choices influence the probability of either a higher load/utilisation and the probability of an empty trip.

Barla et al. (2010) identify three stages being part of the chains: first, at the initial stage of a journey empty runs occur if the vehicles' base is not close to the shipment origin; Second, empty runs arise during a backhaul trip, since demand from a client is rarely bidirectional (typical feature of freight movement); and finally, empty runs occur during a return trip if a truck is diverted in order to pick up a backload. It is important to note that in some of these situations empty running is inevitable. For example, geographical imbalances in freight movement forces trucks to run empty during backhaul from a net freight importing region. It is also impractical to find a back load for specialised vehicles such as oil tankers in to which loading edible oil is impossible (McKinnon, 1996; McKinnon & Ge, 2006). They also indicate lack of transparency in the road freight market, short haul lengths, scheduling constraints and the incompatibility of vehicles and loads as possible causes of empty running. Seeing vehicle utilisation solely based the extent of empty runs may portray an incomplete picture since utilisation also depends on vehicle routing problem faced by the operators and the activity which initiates the movement in the first place.

Generally, empty running can be considered as a reflection of sub-optimal capacity utilisation when it arises due to a matching problem between demand and supply. From an individual operator's point of view there might be cases where running empty in specific direction is optimal and hence load is not searched (e.g. to pick a load in another location). For an operator choosing to operate in this fashion, the price received for a movement of load in the accessed market reflects the cost of operating a vehicle to locations for which capacity is supplied but not utilised. From societal point of view, however, where the objective is sometimes to minimise vehicle movements due to environmental concerns, empty runs that could have been loaded are considered as underutilised capacity. In the next section, studies which analyse the underlying determinants of capacity utilisation are discussed. Finally, we note that when we compare the various truck output and utilisation measures it is evident that they each tell us something different about the level of usage, and that they cannot be interchanged for each other. This shows the importance of choosing the right capacity measure that suits the specific purpose for the analysis and that comparing capacity utilisation across different studies etc. should be done with care.

14.3. Determinants of Productivity in Trucking

14.3.1. Lessons from the Economics Literature

Studies from the economics literature aim at finding determinants of empty running and capacity utilisation of trucks giving particular emphasis to underlying behavioural issues. They give theoretical and empirical analysis for questions such as: why does a truck or a carrier to be precise, choose to access some market segments (origin-destination pairs) and not others? Why do we see empty trucks alongside fully loaded trucks starting from the same origin and going to the same destination? Answers to the first question explain the determinants of individual carrier's vehicle utilisation while answers for the second question explain carriers' market behaviour. The two common factors that determine how well trucking carriers use their capacity are distance and the respective freight movement imbalance between market segments. The effect of these two factors and regulatory environment on rate of capacity utilisation is thoroughly analysed by Wilson and Beilock (1994), Wilson and Dooly (1993) and Beilock and Kilmer (1986) for the US trucking industry. Recent studies focus more on the structure and the relative efficiency within the trucking sector arising from information technology capabilities of trucks (Barla et al., 2010; Hubbard, 2003).

Early empirical studies mainly focus on the cost structure of carriers to explain differences in utilisation performance (see Beilock & Kilmer, 1986; Wilson & Beilock, 1994; Wilson & Dooly, 1993). Their basic explanation is that if there is a systematic cost differential between carriers, then we see difference in capacity utilisation level between similar trucks even if they are used for the same haul and between the same origins and destinations. There are two implied assumptions in these studies, carriers incur more or less equal cost of operating empty trucks and they face equal freight rates. With regard to costs associated with accessing a particular freight market, however, some carriers may have cost advantages. Two main sources of cost advantages and their implication to trucking efficiency are discussed in the literature, namely market regulation and information technology (see Abate, 2013 for a detailed discussion).

The first source of cost advantages arises from government intervention in the trucking industry. The effect of market regulation on carriers' decision to access different freight markets is discussed by Wilson and Beilock (1994), Wilson and Dooly (1993) and Beilock and Kilmer (1986). Even if trucking activities are now free from government intervention in most countries, these studies identify an important source of variation in access cost. All argue, from different angles, that having a special license to haul a regulated commodity improves capacity utilisation (less empty runs) by lowering access cost; whereas not having such a license forces carriers to forego loading opportunities and increase their 'search cost' for finding new loads. There are also later examples where firms choose not to undertake costs to be allowed to carry special types of loads (e.g. permission to carry dangerous goods, which in addition can only happen at certain times during a day or week and thus leading to additional time where a vehicle and driver is not in operation.

One important limitation of these studies is that capacity utilisation is measured solely based on whether a truck is loaded or not without considering how much of the truck's capacity is utilised during a loaded trip (i.e. the load factor).

Another source that leads to different cost structure between carriers is information technology (IT) capability of trucks. Controlling for firm, truck and haul characteristics, Barla et al. (2010) and Hubbard (2003) show that on-board IT capabilities results in higher capacity utilisation by lowering search cost and by improving the matching of vehicle capacity with the available load. The IT systems considered are not clearly defined, but they provide a better communication between operator and driver. Hence, other ITS systems involved in the operation of trucks, such as the logistic operators' software used to optimise the use of vehicles and the general positioning of the vehicles are not included in these studies, although they play important roles in the trucking industry. Chakraborty and Kazarosian (2001) also find positive impact of IT capabilities on productivity by controlling for marketing objectives such as on-time-performance vs. lower rate carrier.

For in depth review on IT capabilities and transport, we refer to Banister and Stead (2004). The specific aspects of capacity utilisation considered in these studies are the level of the load factor and the number of loaded kilometres respectively. Both studies use data from the late nineties when IT adoption was relatively small in North America. Thus, IT capability may not explain capacity utilisation differences in today's activities since the technology might have diffused well by now. It is likely that other attributes of trucks (size, fuel efficiency etc.) or structural changes in the industry are playing important roles in determining how well a particular truck is used. For instance, Boyer and Burks (2009) argue that trucking in the United States has increased its proportion of traffic that is relatively cheap to handle. As a result changes in traffic composition have inflated productivity level in the transport industry. Their finding to some degree plays down the claim that real productivity gains result from systematic difference in cost structure between carriers. However, it may also be that continuously improved matching and routing software continue to induce differences across the trucking industry, but this is very difficult to verify.

Finally, note that optimal decision in a net revenue maximisation framework assumed in the studies summarised above may not be similar to other alternative objectives of firms such as fleet optimisation and/or driver optimisation. For example, drivers may have to call at their home base at a pre-specified time to pick up another load, for vehicle maintenance or to change drivers. This is forcing them to skip a loading opportunity during a backhaul. We also note that the reviewed studies focus on trucks travelling in round trips between two locations (Wilson & Beilock, 1994) or only on one leg of a truck's journey (Barla et al., 2010). This is a rather simplistic depiction of a complex freight vehicle movement which involves a journey with several segments or trips. The reasons often cited for such an approach is limited data availability and analytical tractability. A realistic analysis, however, needs to consider all segments of a truck's movement.

14.3.2. Lessons from the Transport Modelling Literature

In the transport literature the issue of vehicle capacity utilisation and the empty running problem is discussed under freight demand modelling and a vehicle routing problem (VRP). The literature on VRP is rich. An overview of recent developments can be found in Golden, Raghaven, and Wasil (2008). However, most of this literature does not take the capacity utilisation of vehicles or the empty running explicitly into account. An exception is Jordan and Burns (1984, 1987), which gives interesting discussions of truck backhauling as part of a VRP.

Here, we give a short review from the freight demand modelling studies with particular emphasis on trip chain or tours in freight movements and the inclusion of empty trips. We think this is the area where there is a potential for the transport and economics literature to complement each other. Little is known about the behavioural underpinnings resulting commercial vehicle tours. What we know already is from studies based on either simulation or limited data. Starting with general discussions on freight demand modelling and how empty trips can be included, we present a short overview of some of some of these studies in this section.

14.3.2.1. Capacity utilisation in freight demand modelling The literature of freight demand modelling is very rich. Some of the recent developments can be found in this volume on freight modelling, but also the earlier book on freight modelling (Ben-Akiva, Meersman, & Van de Vorde, 2008) as well as in a very early study by Winston (1983). Moreover, recent model developments focus on the logistic choices that can be used to determine the load of vehicles and further the number of vehicles needed to move the quantities determined by the OD matrices (see e.g. de Jong & Ben-Akiva, 2007). We refer to these volumes for more insights. Here we focus on the specific parts, where capacity utilisation and empty trips can be included.

There are two major freight demand modelling platforms, where capacity and empty trips can be included: commodity-based modelling and vehicle-based modelling (Federal Highway Administration, 2007; Holguín-Veras & Thorson, 2000).[1] The difference between them lies in the level of emphasis put on various dimensions such as weight, volume, number of vehicle trips, and economic value of the commodities being transported. In the case of commodity-based modelling, the weight and volume of the freight are the main units of analysis. The assumption is that by focusing on the characteristics of the commodity being transported, it is possible to capture the underlying economic activity which gives rise to vehicle movements. The limitation of commodity-based models is their inability to predict empty trips and the level of vehicle capacity utilisation. This limitation is ameliorated by vehicle-based models, which use vehicle trips to estimate freight demand. As pointed out by Holguín-Veras and Thorson (2000), vehicle-based models in turn overlook the characteristics of cargoes that play an important role in the vehicle selection, mode choice and routing process. Furthermore, these models have limited applicability to multimodal freight transport systems because of their focus on the vehicle trip which is an outcome of prior choice process (Abdelwahab, 1998; Holguín-Veras & Jara-Diaz, 1999; McFadden, Winston, & Boersch-Supan, 1986).

There have been two different approaches to overcome the problems that arise from focusing either on commodity or vehicle movements in the context of urban freight movement. Wang and Holguín-Veras (2008) refer to them as: 'hybrid models' (which estimate commodity flows between origin-destination pairs and delivery routes) and 'tour models' (which directly estimate tours based on logistic considerations, tour-based behavioural models, activity models or profit maximisation behaviour). The common feature of these approaches is consideration of vehicle tours directly or indirectly. But they differ on the extent of treatment of underlying behavioural support for the way individual tours are generated. What is lacking in both approaches is a deeper analysis of the determinants of capacity utilisation of vehicles used for goods-transport. Improvements may be obtained by explicit inclusion of operators' objective function disparities as a result of the characteristics of the owner of the vehicle and its impact on a type of tour undertaken can make the analyses more complete.

1. Here, the concept of demand refers to the flow of freight or the level of freight transport demanded. See Boyer (1998) for an interesting comparison of this concept to when demand is defined as a relationship between the amount of freight transport and the price paid for it.

As indicated in the second section, the level of capacity utilisation of freight vehicles depends on the distance range they are using. Interurban freight movements usually tend to be more efficient since high emphasis is put on consolidation to avoid empty running and less-than-truckload movements. In short range or urban settings, vehicles are usually used in a less efficient manner since several stops are involved. Recent studies on freight demand modelling try to characterise 'trip-chain' or 'tour' based vehicle movements involving several stops. Their main theme is incorporating empty trips and 'trip-chain' behaviour in freight demand modelling (e.g. the modelling approach used by Holguín-Veras & Thorson, 2003). These studies are analysed from a traffic modelling point of view in which the aim is estimation of directional traffic. Other than depicting directional traffic, they do not deeply analyse underlying causes of why some vehicles run empty and others do not. However, there are interesting behavioural analyses of capacity utilisation of vehicles with regard to 'trip chain' movements. In what follows, a brief review of some of these studies is given.

14.3.2.2. Trip chain and capacity utilisation 'Trip chain' or 'tour' is used interchangeably in the literature to describe vehicle movement involving several stops. According to Wang and Holguín-Veras (2008) a trip is defined as an individual vehicle movement connecting an origin to a destination for a specific purpose, and an entire journey comprising two or more trips is a 'trip chain' or 'tour'. Trip-chain is a typical feature of urban freight vehicle movement. Vehicles usually serve several destinations in succession before finally returning back to their base which usually lies in peripheries of cities or near major traffic generators such as distribution centres or ports. Previous studies from Denver (Holguín-Veras & Patil, 2007), Calgary (Hunt & Stefan, 2007), and Amsterdam (Vleugel & Janic, 2004) report an average number of stops per tour of 5.6, 6 and 6.2, respectively.

Figliozzi (2007) gives a theoretical analysis of efficiency of urban commercial vehicle tours with regard to their generation of vehicle kilometres travelled (VKT) assuming a simple scenario where several destinations in an urban area are served from a single distribution centre. He argues that the efficiency of tours depends on the requirements of commercial activity, which initiates the tours and vehicle routing constraints imposed by truck capacity, frequency of service, tour length and time window length. Similarly, Holguín-Veras and Thorson (2003) depict the number of empty trips as a function of the routing choices that the commercial vehicle operators make, which in turn are based on the commodity flows in the study area. Figliozzi's (2007) findings indicate that multi-stop tours generate more VKT than direct deliveries even for equal payloads. Another interesting finding is that the percentage of empty trips has no correlation with the efficiency of the tour regarding VKT generation. According to Figliozzi (2007) looking at the share of empty trips as a measure of efficiency can be misleading. Direct delivery tours, the most efficient tours in terms of VKT since deviations are minimised, always have a 50% share of empty trips while for the less efficient multi-stop tours the share declines with the number of stops. Using data from a single truck engaged in less-than-truckload delivery tours in the city of Sydney, Figliozzi et al. (2007) also shows that there is no clear relationship between tour distance, percentage of empty trips, and percentage of empty distance.

It has to be noted that the context of these studies is urban transport where distance travelled is already short and as such the gain from lower VKT and efficiency is limited.

An empirical analysis using a synthetic dataset of trip chaining behaviour is given by Wang and Holguín-Veras (2008). Even though their study is entirely based on simulation, it gives interesting insights into determinants of destination choice and decision to end a tour using discrete choice models. Destination choice is modelled using a multinomial logit model where the choice set is updated at every node of the trip chain. At each node, a vehicle is faced with four alternative destinations randomly selected based on criteria set by the median distance between all nodes in the network (this is similar to a 'stratified importance sampling technique' suggested by Ben-Akiva & Lerman, 2000). Wang and Holguín-Veras (2008) report that the choice of next destination is negatively affected by the distance from the current location to the potential destination, and it is positively affected by the amount of cargo available for pickup and delivery. As for tour termination decision, the benefit derived from tour termination declines with the increase of return distance and increases with the accumulation of cargo delivered.

A detailed analysis of urban tour-based vehicle movements for the city of Calgary in Canada is given by Hunt and Stefan (2007). The analysis uses a tour-based micro-simulation to model movements made by light vehicles and heavier vehicles with single unit and multi-unit configurations operating in all sectors of the economy. The overall simulation is based on models of tour generation and subsequent sub-models, which determine vehicle and tour purpose, next stop purpose and next stop location. To determine the probability of the next stop's purpose in a tour (which includes carriage of goods, service stops, return-to-establishment and other stop categories), a logit model is used. The total number of previous stops is shown to decrease the propensity to return to vehicle terminal, indicating that tours with a large number of stops are less likely to end at a given next stop. In addition, the time elapsed in travel is shown to impact the propensity to end tours more than the total time elapsed (including both travel time and stop time). Once the purpose of the next stop is known, they use another logit model to determine the location of the next stop. Accordingly, they show that the tendency for the next stop on a tour to be near the current stop is greater than the tendency for the next stop to be near the vehicle terminal. Both models are estimated for 13 different segments of commercial movements based on combination of industry category, vehicle type and tour (stop) primary purpose. The reviewed literature in this section highlights the importance of explicit consideration of trip chaining behaviour to gain insights into commercial vehicle movements. The amount of VKT generated can be an interesting aspect to look into to evaluate efficiency. As shown, however, the share of empty trips of a tour is a poor indicator of efficiency with regard to VKT. It appears that there is a big challenge with regard to finding data as some of the studies are based on simulation and information gathered from operation of a single vehicle. To test some of the theoretical findings, future studies should be based on dataset that contain both the full movement of vehicles over several weeks or months and information on the owners of the vehicles. One such approach is outlined in Abate (2013) for interurban trips. Using detailed trip-by-trip information for about 2000 vehicles for an entire week of operation, this study econometrically estimates determinants of capacity utilisation.

14.4. Concluding Remarks

In this chapter we looked at studies on capacity utilisation of freight vehicles. We classified the studies to general categories of either transport or economics literature, which address the issue from different perspectives. In Table 14.1 some of the reviewed examples of the different strands of literature are summarised. The main characteristics regarding type of utilisation measure as well as the main topic of the studies are highlighted. The list is not comprehensive, but provides an overview.

Our review shows that the strands of the literature have not fully benefited from each other. According to studies based on economic theory, vehicle capacity is underutilised as a result of constant challenge of matching capacity with demand arising from freight movement imbalances between regions and market access cost differential between operators. The problem caused by freight imbalances is considered as an external (exogenous) problem that operators can only minimise through appropriate location of their principal base of operation near major traffic generators, at least in the long run. However, operators have to make continual market access decision as part of the challenge to match specific demand with specific capacity based on net revenue considerations. Accordingly, to the extent to which there are differences in access costs which are not distance related, we see some vehicles running with a load while others running empty across similar market segments and carriers. Recent studies from this strand of literature show the effect of information technology capabilities to match capacity with demand by enabling carriers to keep their trucks on the road and loaded more often by lowering market access costs.

Our review also found that despite simple production process involved in the physical part of the transport process, measuring capacity utilisation is complex due to the extraordinary heterogeneity in terms of weight and volume of a load, direction and distance of vehicle movement. Therefore, a realistic efficiency analysis should account for such heterogeneity that may lead to productivity differences arising from the various settings in which vehicles are used.

In the transport literature, the recent focus is on the relationship between 'trip-chain' and the vehicle routing problem faced by operators in urban freight transport context where utilisation is lower compared to long haul operation. It is shown that unless the commercial activity, which initiates vehicle movements taken into consideration, measures such as the share of empty trips (or distance) can be a poor proxy as efficiency measure when utilisation is compared with regard to generation of vehicle kilometres travelled (VKT). The trip-chain approach to analyse freight movement can improve the modelling of freight transport demand greatly. An even more interesting analysis is the potential of using the information contained

Table 14.1: Main studies on capacity utilisation.

	Source	Utilisation measure	Missing dimension	Remark
Economics literature	Boyer and Burks (2009)	Ton – miles per truck	Density and value cargo; quality of service	Main theme – the effect of traffic composition on trucking productivity
	Barla et al. (2010)	Load factor	Density and value of cargo; quality of service	main theme – the impact of IT technology on capacity utilisation
	Hubbard (2003)	Number of loaded miles and number of periods trucks are in use per period	Weight, density and value of cargo; quality of service	main theme – the impact of IT technology on capacity utilisation
	Abate (2013)	Load factor and empty running	Quality of service and value of cargo; quality of service	This study uses a joint econometric model that considers both load factor and empty running
	Wilson and Beilock (1994); Wilson and Dooley (1993); Beilock and Kilmer (1986)	Empty trips	Weight, density and value of cargo; quality of service	Main theme – the effect of market deregulation on trucking performance
Transport modelling literature	Figliozzi (2007)	Vehicle miles travelled (VMT)	Weight, density and value of cargo; quality of service	The study analyzes urban commercial vehicles
	Holguín-Veras and Thorson (2000)	Empty trips	Weight, density and value of cargo; quality of service	
	Holguín-Veras, Torres Cruz, and Ban (2011)	Total travel time, Vehicle miles Traveled (VMT), and pollution levels	Density and value of a cargo; quality of service	The study analyzes urban delivery vehicle classes

in data about trip chaining in relation to the firm optimisation behaviour investigated in the economics literature. The trip chain approach also adds a spatial element to the firms' desire to match capacity with demand. The analysis of the operators' choice of optimal location of their principal base may very well be affected by the specific inclusion of trip chain information. Moreover, including trip chains and/or the relationship between loaded trips and empty trips in the analysis of the matching behaviour will increase the predictive power of such analysis.

Finally, empirical analyses of the scope of possible gains from improved capacity utilisation are limited. More studies are needed along the McKinnon and Ge (2006) study to put into perspective how much can be gained by improving empty runs. It is also interesting to know how the desire for sustainable transport can be accommodated within the objective of transport operators which is usually based on economic efficiency. The recent freight transport modelling approach, where the logistic decisions are involved to better predict vehicle movements is an obvious link between the operators' pursuit of efficiency analysed in the economics literature and the freight transport models. The determinants in the operators' decision making influence capacity utilisation and thus also the vehicles that are loaded onto the network in the assignment models. A joint estimation of firm optimising behaviour and the logistics of freight transport models will thus be a natural research objective for further investigation.

Acknowledgements

The work presented in this chapter is financed by the Danish Strategic Research Council through the Drivers and Limits project. We thank two anonymous referees for suggestions.

References

Abate, M. A. (2013). Determinants of capacity utilisation in road freight transport. *Journal of Transport Economics and Policy* (forthcoming).

Abdelwahab, W. M. (1998). Elasticities of mode choice probabilities and market elasticities of demand: Evidence from a simultaneous mode choice/shipment size freight transport models. *Transport Research. Part E: Logistics & Transport Review*, *34*(4), 257–266.

American Trucking Associations. (2011). *American trucking trends*. Alexandria, VA: Transport Topics Press.

Banister, D., & Stead, D. (2004). Impact of information and communications technology on transport. *Transport Reviews*, *25*(5), 611–632.

Barla, P., Bolduc, D., & Boucher, N. (2010). Information technology and efficiency in trucking. *Canadian Journal of Economics*, *43*(1), 254–279.

Beilock, R., & Kilmer, R. (1986). The determinants of full-empty truck movements. *American Journal of Agricultural Economics*, *68*(1), 67–76.

Ben-Akiva, M., & Lerman, S. R. (2000). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, Mass: MIT Press.

Ben-Akiva, M., Meersman, H., & Van de Vorde, E. (Eds.), (2008). *Recent developments in transport modeling: Lessons for the freight sector*. Bingley, UK: Emerald Group Publishing Ltd.

Boyer, K. D. (1998). *Principles of transport economics* (pp. 48–49). Addison-Wesley Reading: Massachusetts.

Boyer, D. K., & Burks, S. V. (2009). Stuck in the slow lane: Undoing traffic composition biases in the measurement of trucking productivity. *Southern Economic Journal*, *75*(4), 1220–1237.

Chakraborty, A., & Kazarosian, M. (2001, January). Marketing strategy and the use of information technology: New evidence from the trucking industry. *Research in Transport Economics*, *6*(1), 71–96.

De Jong, G., & Ben-Akiva, M. (2007). A micro-simulation model of shipment size and transport chain choice. *Transportation Research Part B: Methodological*, *41*, 950–965.

Demirel, E., van Ommeren, J., & Rietveld, P. (2010). A matching model for the backhaul problem. *Transport Research Part B*, *44*(4), 549–561.

Federal Highway Administration. (2007). *Quick Response Freight Manual*. Washington, DC: U.S. Department of Transport.

Felton, J. R. (1981). The impact of rate regulation upon ICC-regulated truck back hauls. *Journal of Transport Economics and Policy*, *15*(3), 253–267.

Fernie, J., & Mckinnon, A. (2003). The grocery supply chain in the UK: improving efficiency in the logistics network. *International Review of Retail, Distribution & Consumer Research*, *13*(2), 161–175.

Figliozzi, M. A. (2007). Analysis of the efficiency of urban commercial vehicle. *Transport Research Part B*, *41*, 1014–1032.

Figliozzi, M., Kingdon, L., & Wilkitzki, A. (2007). Commercial vehicle tours and road congestion in urban areas: implications for carriers' operations and public data collection and dissemination efforts. In: *Proceedings of the 86th Transport Research Board Annual Meeting* CD room – January 2007, Transport Research Board of the National Academies, Washington, DC.

Golden, B., Raghaven, S., & Wasil, E. (Eds.). (2008). *The vehicle routing problem: Latest advances and new challenges*. New York: Springer.

Holguín-Veras, J., & Jara-Diaz, S. (1999). Optimal space allocation and pricing for priority service at container ports. *Transport Research Part B*, *33*(2), 81–106.

Holguín-Veras, J., & Thorson, E. (2000). An investigation of the relationships between the trip length distributions incommodity-based and trip-based freight demand modelling. *Transport Research Record* 1707:37–48.

Holguín-Veras, J., & Thorson, E. (2003). Modelling commercial vehicle empty trips with a first order trip chain model. *Transport Research Part B*, *37*, 129–148.

Holguín-Veras, J., & Patil, G. R. (2007). Integrated origin-destination synthesis model for freight with commodity-based and empty trip models. *Transportation Research Record*, No. 2008, 2007, pp. 60–66. Transport Research Board of the National Academies, Washington, DC.

Holguín-Veras, J., Torres Cruz, C. A., & Ban, X. (2011). On the comparative performance of urban delivery vehicle classes. *Transportmetrica*, *9*, 1–24.

Hubbard, T. (2003). Information, decisions, and productivity: On board computers and capacity utilisation in trucking. *The American Economic Review*, *93*(4), 1328–1353.

Hunt, J., & Stefan, K. (2007). Tour-based microsimulation of urban commercial movements. *Transport Research Part B*, *41*, 981–1013.
Jordan, W. C., & Burns, L. D. (1984). Truck backhauling on two terminal networks. *Transport Research B*, *18*(6), 487–503.
Jordan, W. C., & Burns, L. D. (1987). Truck backhauling on networks with many terminals. *Transport Research B*, *218*(3), 183–193.
McFadden, D., Winston, C., & Boersch-Supan, A. (1986). Joint estimation of freight transport decisions under non-random sampling. In A. Daugherty (Ed.), *Analytical studies in transport economics*. Cambridge: Cambridge University Press.
McKinnon, A. C. (1996). The empty running and return loading of road goods vehicles. *Transport Logistics*, *1*(1), 1–19.
McKinnon, A., & Ge, Y. (2006). The potential for reducing empty running by trucks: A retrospective analysis. *International Journal of Physical Distribution & Logistics Management*, *36*(5), 391–410.
Nelson, R. A. (1989). On the measurement of capacity utilisation. *The Journal of Industrial Economics*, *37*(3), 273–286.
Raothanachonkun, P., Wisetjindawat, W., & Matsumoto, S. (2007). Estimating Truck Trip O-D with Commodity-Based and Empty Trip Models. *Transportation Research Record*, No. 2008, 2007, pp. 43–50. Transport Research Board of the National Academies, Washington, DC.
Transport Canada. (2003). *Les transports au Canada, Rapport Annuel 2003 et 2004*. Retrieved from http://www.tc.gc.ca/pol/fr/rapport/anre2004/tocf.htm
Vleugel, J., & Janic, M. (2004). Route choice and the impact of 'logistic routes'. In Taniguchi, E., & Thompson, R. (Eds.), *Logistics systems for sustainable cities*. Elsevier.
Wang, Q., & Holguín-Veras, J. (2008). Investigation of Attributes Determining Trip Chaining Behavior in Hybrid Microsimulation Urban Freight Models Transport Research Record: *Journal of the Transport Research Board*, No. 2066, Transport Research Board of the National Academies, Washington, pp. 1–8.
Wilson, W. (1987). Transport markets and customer behavior: The backhaul problem. *Journal of the Transport Research Forum*, *28*(1), 325–333.
Wilson, W., & Beilock, R. (1994). Market access in regulated and unregulated markets: The continuing cost of interstate motor carrier regulation. *Journal of Regulatory Economics*, *6*, 363–379.
Wilson, W., & Dooly, F. (1993). An empirical examination of market access. *Southern Economic Journal*, *60*(1), 49–62.
Winston, C. (1983). The demand for freight transport: Models and applications. *Transport Research*, *17*(a), 419–427.
Winston, C. (1985). Conceptual development in the economics of transport: An interpretive survey. *Journal of Economic Literature, American Economic Association*, *23*(1), 57–94.

Chapter 15

Valuation of Transport Time Savings and Improved Reliability in Freight Transport CBA

Inge Vierth

Abstract

The purpose of this chapter is to show how freight transport time savings (VTTS) and reduced variability in transport time (VTTV) are valuated in cost–benefit analysis (CBA) and to discuss alternatives. The chapter suggests a structure of the benefit components where VTTS comprises the reduction in (a) transport resources required for a given service VTTS (T), (b) capital tied up in goods while transported VTTS (G) and (c) goods users' opportunity costs VTTS (O). VTTV comprises the reduction in (a) transport resources required for a given service VTTV (T) and (b) users' opportunity costs VTTV (O).

Normative and behavioural approaches are applied to determine VTTS and VTTV. The VTTS definition differs between countries: Sweden, Norway, Germany and the United Kingdom calculate VTTS (T) based on market prices, the Netherlands use an all encompassing VTTS derived from SP data. Sweden and Norway are the only countries that use VTTS to denote only VTTS (G). Only the Netherlands, Sweden, and Norway apply preliminary VTTV in their CBA.

Different VTTS and VTTV constructions make the transfer of results difficult. It is found that the savings in transport resource costs VTTS (T) can be calculated with help of engineering formulae. Provided that the VTTS can be limited to VTTS (G), where market-based approaches can be used, and VTTS (O). The incorporation of VTTV in CBA requires further development of the valuation methods. The use of information of the trade-off between transport costs and inventory costs is one promising approach. Hopefully

Freight Transport Modelling

ISBN: 978-1-78190-285-1

lessons can be learned from the Dutch and Norwegian SP studies for VTTS and VTTV. A better understanding of the opportunity costs is needed.

Keywords: Transport time savings; improved reliability; freight cost–benefit analysis; CBA guidelines; valuation; buffer stocks; benefit transfer

15.1. Introduction

Transport time savings and improved reliability are key elements in the calculation of the (potential) benefits resulting from improved transport infrastructure, pricing and regulatory measures.[1] This is true for passenger transports as well as for freight transports that are addressed here. The consumer surplus is more difficult to access for freight transports (as intermediate good) than for passenger transports (as final good). The complexity of the valuation of transport time savings and improved reliability in freight is illustrated in the surveys by (Bruzelius, 2001) and (Significance, Goudappel Coffeng, and NEA, 2012).

To calculate the benefits due to reduced transport time from, for example, infrastructure measures in socio economic cost–benefit analysis (CBA), typically goods volumes, vehicle flows and delays per infrastructure link are calculated separately for the reference case and the scenario case, including the policy being evaluated, and then compared. The reduced travel time, in number of minutes or hours, is then transferred into monetary values by using official unit values, that is, e.g. 'Values of Travel Time Savings (VTTS)'.

Reduced standard travel time variability is typically measured in terms of reduced standard deviation around the average travel time, and can be transferred into monetary values by using 'Values of Travel Time Variability (VTTV)'. The term VTTV comprises too late and too early arrivals as well as earlier departures than desired. Normally all these deviations cause costs for the companies involved. Other dimensions of reliability than time, for example, damage or theft of the goods during the transport are not included below.

The reduction of the travel time and the standard travel time variability leads to time gains and monetary gains for different benefit components. The VTTV can be seen as a factor or multiplier of the VTTS.

15.1.1. Structure and Sources of the Benefit Components

To evaluate the benefits due to reduced transport time and variability all relevant benefit components for the shippers (senders and receivers of the goods) and carriers

1. Time savings and improved reliability can be positive or negative.

have to be included. The structure of the benefit components related to VTTS and VTTV for freight transport could be described in the following way.

At the highest level VTTS and VTTV are distinguished and within VTTS resp. VTTV the following components:

VTTS

a) reduction in the *transport resources* that are required to carry out certain transports (VTTS (T)), that is, reduction in time and (monetary) costs for vehicles and staff
b) reduction in the unproductive use of capital tied up in *goods* while being transported (VTTS (G))
c) reduction in the *opportunity* costs to the users of goods VTTS (O). One example is the 'unavailability' costs for not having the goods available at the point of sales (which can lead to lost sales).

VTTV

d) reduction in the *transport resources* required to ensure a given service level (VTTV (T)), that is, reduction in slack related to transport resources
e) reduction in *opportunity* costs to user of goods (VTTV (O)), that is, reduction in buffer stocks (Vierth, 2010).[2]

The benefits of time savings (VTTS) and improved reliability (VTTV) for freight transport can be obtained from the following sources:

a) Values of Travel Time Savings–VTTS (T = transport resources)
   Engineering cost models or transport models can be used to calculate the use of transport resources VTTS (T). Alternatively, carriers could be asked to estimate the value of shorter transport time related to the transport resources.
b) Values of Travel Time Savings–VTTS (G = goods)
   The VTTS (G) component can be calculated as the market value of the goods multiplied by the interest rate and the time spent in transport.
c) Values of Travel Time Savings–VTTS (O = opportunity for the users of the goods)
   The VTTS (O) component is related to specific situations of specific companies in specific markets. Due to the heterogeneity of the firms and products it is likely that any estimation method will have to be capable of handling this heterogeneity. One has to be aware that there is a risk of double counting, since some benefits may mirror savings that are already included in VTTS (T) or VTTS (G). However, there may be aspects with respect to delivery time that is offered to the users of the products.

2. The Swedish report includes an English summary; the report is in many parts based on (Swahn, H; Bates. J, 2010).

d) Values of Travel Time Variability–VTTV (T)
   The same type of approach as for VTTS (T) can be applied.
e) Values of Travel Time Variability–VTTV (O)
   The impact of a change of VTTV (O) can in principle be measured by the change of the buffer stock required (using the appropriate stock out cost formulae).

15.1.2. Organization of This Chapter

We address the question how travel time savings and improved reliability in freight transport are valuated/and how the unit values for VTTS and VTTV are applied in the societal CBA in different countries. Alternatives to the existing solutions are also discussed. In Section 15.2 we try to establish a base for understanding the CBA context in which the unit values are applied. In Section 15.3, the calculation methodology for time related benefits in the CBA is presented. In Section 15.4, it is shown how VTTS and VTTV are applied in CBA in a few European countries. In Section 15.5 alternatives for the definition, scope and methods to quantify VTTS and VTTV are discussed. Finally, conclusions are provided in Section 15.6.

15.2. Societal CBA Context

In most countries infrastructure projects (at least the larger ones) are evaluated via socio-economic CBAs. The HEATCO project has proposed harmonized guidelines for project assessment for transnational transport projects in Europe (HEATCO, 2006). These guidelines have been developed based on research on the different aspects of project appraisal and on an analysis of existing practice in European countries. There is a longer CBA tradition for passenger transports than for freight transports and benefits related to passenger transport are, except for transports on inland waterway and sea transports, generally larger than for freight transports.

15.2.1. CBA Requirements

In general, a socio-economic CBA can only provide a correct picture of the benefits if there are:

- no market failures. Though transport markets are generally competitive there exist situations where, for example, carriers have a strong market position. Monopoly pricing is most common in specialized segments and at a regional level.
- no unpriced externalities and subsidies (this is not always given but can be illustrated with help of sensitivity analyses).
- linear production functions in the transport sector and the transport using sectors (we develop in Section 15.4 how economies of scale in the transport system can be taken into account).

15.2.2. Non-Marginal Infrastructure Changes

The traditional CBA approach assumes that the policies that are evaluated imply only marginal changes to the transport system and the surrounding production/consumption system, and that costs and prices in general remain unchanged. However, especially in the long-term development of the infrastructure there will be certainly non-marginal changes. Beside the direct effects on the transport time and its variability, such non-marginal changes could also have wider impacts on, for example, prices of services within the transport sector as well as on prices of goods and services outside the transport sector.

15.3. Methodology of Freight VTTS and VTTV Components

15.3.1. Methodology

The value related to transport time savings and reduced transport time variability in general and in the context of freight transport, is due to the *benefits to the consumers*. However, under certain assumptions about the function of markets (see above) we may assume that the net benefit to one or more firms could be used to approximate the benefit to consumers. This gives the rationale for approaches to the estimation of the unit values by taking a *company perspective*.

(Bruzelius, 2001) notes that there are two basic approaches for determining unit values: (a) a normative market price approach and (b) a behavioural approach which is based on inferring values from choices made between available alternatives for transporting a shipment. Alternatives for different modes and/or routes may each be associated with different transport times and quality levels. Typically logit models are used for the estimation of the unit values. They express the probability of the choice of an alternative in terms of the generalized costs of that alternative as well as of the other available alternative(s). The behavioural approach can be based on actual choices (revealed preference—RP), hypothetical choices (stated preference—SP) or combinations of RP and SP.

15.3.2. VTTS and VTTV for Freight and Passenger Transport

In the CBA context, a person performing a leisure trip by public transport ‘corresponds’ to the transported goods (if the transport is not an own account transport). In both cases, transport resources, organized by public transport operators resp. freight transport carriers, are used to perform a transport. The benefits relating to both the *transport object* and the *transport resources* need to be taken into account in the CBA.

For *the public transport traveller* (who is doing a leisure trip), the VTTS is the amount he/she would be prepared to pay to reduce the scheduled time of the journey. This is typically determined with help of SP studies where passengers are asked.

However, a reduction in travel time can also impact on the public transport operator. The operator's vehicle operating costs and staff costs are typically measured by market prices.

In principle the situation is analogous for freight transport, with the *goods on the freight vehicle* taking the place of the passenger on the public transport vehicle. The alternative to ask the passengers about their VTTS that is used in passenger transport is not available for freight transport. But unlike the passengers, the goods cannot be interviewed to determine their VTTS (G) and VTTS (O). In principle it is possible to ask the shippers and carriers involved in the freight transport, but this approach comprises problems that are discussed below.

15.4. Application of Freight VTTS and VTTV in CBA in a Few Countries

In this section, we show how the time savings associated with the vehicle, the staff and the goods are dealt with in practice in the United Kingdom, the Netherlands, Germany, Sweden and Norway.

15.4.1. United Kingdom

15.4.1.1. Vehicle (VTTS (T)) According to the UK Department's Guidance (DFT(UK), 2012) travel time savings associated with the *vehicle* contribute to vehicle operating costs (VOC). There is only explicit guidance for road vehicles, though there would be no objection to including operating savings for vehicles on other modes. The *fuel costs* are calculated by means of a fuel consumption formula (litre per km) which is sensitive to speed; in this way time savings impact on fuel operating costs. To get the total fuel costs, the fuel consumption is multiplied by the distance, the fuel price and by a factor representing the changes in fuel efficiency over time. The *non-fuel vehicle operating costs* include costs for oil, tyres, maintenance, depreciation and allowances for the purchase of new vehicles. The marginal resource costs of oil, tyres, mileage and maintenance related depreciation, are assumed to be fixed costs per kilometre. The time component of depreciation is excluded since it does not vary with distance or speed. Non-fuel vehicle operating costs are assumed to remain constant in real terms over the forecast period.

15.4.1.2. Staff (VTTS (T)) The *staff time* savings are evaluated in the same way as any time savings in the course of work, by assuming that savings in travel time convert non-productive time to productive use and that, in a free labour market, the value of an individual's working time to the economy is reflected in the wage rate paid. Composition of the staff per vehicle (including wage rate for each person) are needed as input in the calculations.

15.4.1.3. Goods (VTTS (G)) Currently no allowance is made for time savings to the goods in the United Kingdom. While there has been discussion as to whether any aspect of the 'freight value of time' should be related to the goods itself, no recommendations have been made in this respect.

15.4.1.4. VTTV Improved reliability for freight transports is neither included in the CBA in the United Kingdom.

15.4.2. Germany

The German CBA guidelines (BMVBS, 2005) specify how to calculate the freight transport costs (VTTS (T)) for road, rail or on inland waterway. These transport costs comprise costs related to the vehicle and the staff. The costs are specified for several vehicle types within the modes. Benefits related to time savings for the goods transported (VTTS (G)) are not incorporated the German CBA guidelines.

Benefits due to improved reliability (VTTV) are not part of the existing CBA guidelines either. The ministry has commissioned a pre-study to assess the possibilities to include reliability for passenger and freight transport in the CBA in the future, (Significance, Goudappel Coffeng, and NEA, 2012).

15.4.3. The Netherlands

The existing Dutch CBA Guidelines (Rijkswaterstaat, 2012) comprise the following three items related to VTTS and VTTV for freight transport:

– *Vehicle operating costs (VTTS (T))*
  The variable operating costs are calculated only for the vehicle. They are primarily distance based (and not speed-related).
– *Time savings for vehicle, staff and goods (VTTS (T) + VTTS (G))*
  The VTTS for the vehicle, the staff and the goods is derived from a SP study (RAND Europe, SEO, Veldkamp/NIPO, 2004) and should in principle include staff costs, depreciation on the vehicle and goods-related costs.
– *Reliability (VTTV)*
  If the effect on reliability of a policy can be measured ex ante in terms of more or less standard deviation, it is recommended to work with the *Value of Reliability* (VoR) that is calculated as the *Value of time* (VoT) * *Reliability ratio* (RR). The VoT corresponds to the time savings for vehicle, staff and goods. The reliability ratio is the ratio of the value of one minute of standard deviation to the value of one minute of average travel time. Provisional RRs, based on (Kouwenhoven, de Jong, & Rietveld, 2005), are recommended. The provisional RR for road freight is 1.2.

When no information about the (potential) reliability effects of a measure is available, a rule of thumb approach is applied. This means that the travel time savings are increased by 25% to include reliability related benefits. However, this mark-up may only be applied in road projects that lead to less congestion. There is though a 'tenuous link' between reduced congestion and increased reliability.

15.4.4. Sweden

The Swedish CBA guidelines (Trafikverket, 2012a) differentiate between the time savings for the staff and the vehicle VTTS (T) and the time savings for the goods itself VTTS (G).

15.4.4.1. VTTS (T) The VTTS (T) calculation for road and rail is carried out with help of engineering approaches (Trafikverket, 2012b).

So far the aggregate-disaggregate-aggregate (ADA) freight model system described in Chapter 4 in this book *(M. Ben-Akiva and G. de Jong, The Aggregate-Disaggregate-Aggregate (ADA) Freight Models System)* that has originally been developed for Norway and Sweden has not been used to calculate the VTTS and VTTV.[3] This is true both for Norway and Sweden; there are though plans to do this in both countries.

15.4.4.2. VTTS (G) The VTTS (G) component is calculated based on the market value of the goods, (Trafikverket, 2012c), an interest rate of 20% is applied and it is assumed that the logistics system is available 3600 hours (out of 8760 hours) per year. The so called logistics factor (of 2) is also applied. This factor was calculated by (Bruzelius, 1986) as indication of the size of the logistics benefits that can be achieved in the material handling system as a result of reduced transport times. The logistics factor is calculated based on the possible cost reductions in buffer stocks. The market based approach has been chosen due to the fact that the results in the SP study (INREGIA, 2001) indicated that the value of the goods is essential for the VTTS (G). Otherwise, it was difficult to use the within the SP study collected mode specific responses to estimate commodity specific VTTS (G).[4]

15.4.4.3. VTTV The Swedish guidelines recommend taking into account improved reliability for rail freight with help of a provisional VTTV (G) that is two times the VTTS (G). Reliability is an important issue for rail freight. However, the provisional VTTV (G) have only been applied in very few rail infrastructure appraisals due to limited information about how specific measures influence travel time and/or reliability (Vierth, 2012).

3. These models describe economies of scale for transport costs and the trade-off between transport and inventory costs.

4. In the past, mode-specific VTTS for road and rail were applied based on two SP studies (Banverket, 1990) and (Vägverket, 1992. See (Bruzelius, 2001) for a description of the studies in English.

15.4.5. Norway

In Norway the CBA guidelines published by the Road Administration (Statens Vegvesen, 2006) describe the calculation of VTTS (T) for road transports and the guidelines published by the Rail Administration (Jernbaneverket, 2011) the values for VTTS (T) for rail transports. The latter include also values for the goods VTTS (G) transported by rail. Average values per train load are used and not commodity specific values. The rail guidelines include also values for VTTV (G) that are about 20 times the VTTS (G). The guidelines do not include information about the basis of the values of NOK 0,65 and NOK 12,5 per tonne hour.

A first version of a calculation tool has been developed that facilitates the appraisal of freight transport related measures based on the ADA model system, (Madslien & Minken, 2011). This tool delivers logistics costs, tonne-km and vehicle-km in the reference and base scenario. By applying the ADA model system some effects of economies of scale on transport cost can be captured and thus potentially influence VTTS (T) and VTTV (T).[5]

15.4.6. Summary

15.4.6.1. VTTS We find that current practice in relation to CBA and the implied definition of the VTTS for freight transport differ between the United Kingdom, the Netherlands, Germany, Sweden and Norway. This is summarized in Table 15.1.

For Sweden, Norway, Germany and the United Kingdom, the vehicle related costs are derived by market price calculations, though relating to different prices. According to the existing guidelines in the Netherlands, an all encompassing VTTS is used, derived from SP data. The use of different VTTS definitions can lead to misunderstandings when comparing unit values between countries. See for example in Bruzelius survey (Bruzelius, 2001).

Sweden appears to be the only country that uses VTTS to denote only the benefit component VTTS (G) for all modes, that is, the component relating to the goods. In Norway a VTTS (G) is used for rail transports. It is notable that the discussion of VTTS for commercial transports in (HEATCO, 2006) seems to concentrate on vehicle and staff operating costs, which according to the Swedish definition has nothing to do with the VTTS for the goods. There is though a passing reference to the goods themselves and it is notable that HEATCO argues that heterogeneity of goods should be accounted for within the modelling process, but the discussion

5. The ADA model system comprises different vehicle sizes per mode (resp. vehicle type) and can take into account economies of scale in the logistics choices (choice of shipment size and vehicle type/size, level of consolidation etc.)

Table 15.1: Current practice in relation to CBA and the implied definition of freight VTTS in the United Kingdom, Germany, the Netherlands, Sweden and Norway

		UK	Germany	The Netherlands	Sweden	Norway
Vehicle operating costs (VTTS (T)	Km-based	Elsewhere in CBA	Elsewhere in CBA	"VTTS" (any effects due to vehicle operating costs will be removed)	Elsewhere in CBA	Elsewhere in CBA
Vehicle purchase costs (inch depreciation)	Time-based	Elsewhere in CBA	Elsewhere in CBA		Elsewhere in CBA	Elsewhere in CBA
Staff costs VTTS (T)	Time-based	"VTTS"	Elsewhere in CBA		Elsewhere in CBA	Elsewhere in CBA
goods costs (VTTS (G)	Time-based	Not used	Not used		"VTTS" incl. allowance for variability	"VTTS" used for rail transports
Opportunity costs VTTS (O)	Time-based	Not used	Not used		Not used	Not used

leaves the issue open on what to do if transport demand models with the required level of sophistication are not available.[6]

One possible explanation for why the VTTS (G) component is often excluded is that the freight transports are complex and involve several decision makers (shippers, forwarders, carriers) and strategies (own account or outsourced transports) and commodities (high/low value goods, perishable goods etc.) Another possible explanation is that the benefits related to the goods are excluded as they are assumed to be very low (and not worth to be calculated).

15.4.6.2. VTTV Improved reliability for freight transport is not included in the CBA in the United Kingdom and Germany. In the Netherlands it is recommended to work with the 'Value of Reliability' that is calculated as the 'all-encompassing VTTS' *'Reliability ratio', if the effect on reliability of a policy can be measured in terms of more or less standard deviation. If not, the travel time savings for road transports are increased by 25% to include reliability related benefits. In Sweden and Norway provisional values are used to include reliability benefits for rail transports, the VTTV(G) values are 2 respectively 20 times the VTTS (G) values.

15.4.6.3. Composition of link/node flows As VTTS and VTTV normally are specified for a number of commodity groups[7] (load units or modes). As infrastructure (including the links that are affected by an infrastructure project) typically is used by many different firms, it is necessary to obtain information about the composition of the flow on the links affected, that is, by a specific infrastructure project by the same disaggregation criteria as the VTTS and VTTV. An alternative but less accurate approach is to apply national or regional averages.

15.5. Alternatives to Existing VTTS and VTTV

15.5.1. VTTS

Below alternatives to today's applied methods and definitions of VTTS and VTTV are discussed.

15.5.1.1. Methodological considerations and definitions of VTTS We assume that the savings in vehicle operating and staff costs *VTTS (T)* can be calculated elsewhere within the CBA framework (whether they derive from time savings or due to other reasons).

6. Infrastructure holders are typically not able to observe the contents of the freight vehicles that use their infrastructure (i.e. if the vehicles are loaded or not).

7. In the Norwegian/Swedish model about 30 commodity types are used.

Provided that, it would be sufficient to limit the VTTS to the goods (VTTS (G)). This is in line with (Bruzelius, 2001) and the current Swedish VTTS definition. In principle, the costs can be calculated as the product of the value of the goods multiplied by an appropriate rate of interest. The estimation of the VTTS (G) might though cause 'practical' problems related to which market prices (e.g. due to heterogeneity within commodities), interest rate to use and which period of time to use. The door-to-door time (transit time) comprising times for loading, unloading, transhipment (inch waiting times) is typically longer than the pure transport time (on the links addressed). The impact of a change in transport time on the total door to door time may be small, which is related to the questions about the additive nature of small time savings and indivisibilities in the transit process. The impact may be large when schedules have to be maintained.

The opportunity costs (VTTS (O)) are not *per se* related to the value of the goods itself but rather to the added value at the place of use/consumption. The opportunity costs are normally counteracted by keeping goods available in stock. In a situation with no uncertainty relating to demand and lead time for supply, the stock that is required to effectively eliminate stock-out situations can easily be calculated. But the buffer stocks approach does not work in cases where there is no prior knowledge about the demand for a certain product (i.e. fashion, seasonal products or important spare parts). The potential scope and definition of the VTTS (O) depends on how the overall CBA evaluation is arranged, for example, if wider impacts are included.

It is essential to recognize that none of the effects above relate to the value of competitive gains for one company in relation to the competing companies.

15.5.1.2. Approaches to estimation of VTTS (G)

Normative approach In the logistics model that is part of the ADA model system the transport time-related costs for the goods are described as 'capital costs due to unproductive use of the goods'. Information about the value of the goods can be taken from commodity flow surveys, foreign trade statistics etc. Nonetheless, some efforts are required to make a clear specification in order to calculate an appropriate set of (commodity specific) unit values, taking account geographical variations for the commodity groups (e.g. distinction between domestic flows, imports and exports) and allocation to mode (e.g. distinction between goods that are transported by air or by land based modes).

Behavioural approach While the use of RP data in principle can be preferred to hypothetical SP data, the effort involved collecting RP data is often high. In practice most VTTS (G) in different countries have been obtained by using SP experiments. In his literature review (Bruzelius, 2001) describes the results of the studies with scepticism. It seems to be essential how the questions in the SP experiments are put and to whom they are addressed. Also the recently performed review (Significance, Goudappel Coffeng, and NEA, 2012) shows that all unit values for the VTTV have been obtained from SP studies. Most of the studies described include VTTS as well.

Three SP studies have been carried out lately in the Netherlands and in Norway.

Dutch SP study for all modes The Dutch study measures the value to society of travel time benefits and travel time reliability benefits in passenger and goods transport (Ministerie van Verkeer en Waterstaat, 2012). The VTTS will be updates of the existing unit values, while the VTTV will be the first of their kind for the Netherlands. At the time writing the results from the study are not yet available.[8]

In the SP experiment for VTTS, where the trade-off between delivery time and costs is analysed, it is recognized that not all SP respondents can provide the full information on all aspects and the following assumptions are made. The opportunity costs (VTTS (O)) are included (see Table 15.2).

In line with our notation the expectation is, that with *shippers* one would get an estimate of VTTS for typical shipments that are the combined VTTS (G) + VTTS (O), while for *carriers* one would get VTTS (T). Shippers are generally interviewed in relation to the door-to-door chains, while carriers (often) are asked on parts of the chains. One aim of the Dutch SP study is to analyse the data with a view to identifying which elements appear to have been included. It is in particular assumed to be possible that some respondents may include vehicle operating costs that are already considered elsewhere in the CBA (see Table 15.1). To avoid double counting, these need to be removed.

Norwegian SP study for all modes The 'Dutch methodology' has also been used in the Norwegian SP study for VTTS and VTTV for passenger and freight transport. The results related to freight transport can be found in (Halse, Samstad, Killi, Flügel, & Ramjerdi, 2010). Separate SP experiments have been carried out with carriers and shippers according to the assumptions in Table 15.2. With approximately 9% for

Table 15.2: Assumptions of which actor should answer which questions in Dutch SP study

	VTTS (G + O)	**VTTS (T)**
Carrier	0	Factor cost
Shipper that contracts out	Interest, deterioration, disruption of production, out of stock	0
Shippers with own account transport	Interest, deterioration, disruption of production, out of stock	Factor cost

Source: de Jong et al. (2012).

8. According to the Dutch Transport Ministry, the new VTTS and VTTV values will be released before then end of 2012.

shippers and around 6% for carriers, the response rates are generally low. Large and medium sized companies are overrepresented.

The authors recommend using the VTTS (G) and VTTV (G) for *road freight*. The values have been derived from the *shippers'* valuations between fast/reliable transport services and prices for these services. The calculated unit values for the goods should be added to the VTTS (T) and VTTV (T). The quality of the data collected for the other modes is not good enough to be able to calculate the corresponding VTTS and VTTV.

It is however not recommended to apply results for the responding *carriers*, as the results indicate that the carriers have included as well the shippers' costs in their answers. The transport cost estimates in the Norwegian CBA-guidelines (Statens Vegvesen, 2006) that are based on factor costs should be used (instead of the answers form the carriers in the SP study). This implies for large lorries that the share of the VTTS (G) of the total VTTS is 17% (if the loaded lorries are included) resp. 11% (if the loaded and unloaded lorries are included).

Due to the heterogeneity of the results, the authors see a need for studies with more homogenous samples, before applying the values in CBA.

Norwegian SP study for rail A 'follow-up' study for rail (Halse & Killi, 2010) with a similar survey design as in (Halse et al., 2010) has been carried out in order to obtain better data for the rail mode. It was possible to analyse the data collected in this study. For results, see under VTTV below.

15.5.2. VTTV

Reliability hardly features in social CBA despite the intentions of the authorities in several countries, as the theoretical basis for valuing reliability is generally less developed than that for valuing time savings (OECD/ITF, 2010). The studies that have been carried out are typically related to expected delays. Unexpected delays are however nearly nowhere in the world taken into account systematically, (Significance, Goudappel Coffeng, and NEA, 2012).

15.5.2.1. Methodological considerations and definitions of VTTV In general, the appropriate response to uncertainty for the companies is to incorporate 'slack'. Essentially as developed by Bruzelius (1986), this is a trade-off between the additional cost of holding inventory and the possible risk of lost sales or other contingencies. The exact consequences differ depending on if the goods are delivered to final consumers, trading or producing companies. In all cases there is potentially a penalty function at both the sending and receiving end. Regardless of the actual organization, the benefit of reduced variability in travel time is related to the impact on utility of the user of the goods. The approach relies on some knowledge of the distribution of the demand. The two sources of uncertainty– demand and travel time – can both interact and reinforce each other.

As for VTTS it is essential to recognize that none of the effects above relate to the value of competitive gains for one company in relation to the competing companies.

15.5.2.2. Approaches to estimation of VTTV

Normative approach (Bruzelius, 2001) suggests that VTTV (G) values that are based on market prices can be derived from an investigation of the buffer stocks (and their implicit additional costs). According to Bruzelius the capital value approach can be taken further by also accounting for reliability (and damage). A theoretical analysis of the issue of reliability has been presented by Minken (1997), and a simplified version is to be found in (Bruzelius, 1986). The approaches used by Minken and Bruzelius both presume that the variability in transport time can be described in terms of a probability density function and that logistics planners build up stocks to ensure that stock out will not occur or will occur rarely on account of the variability in transport time.

The value of improved reliability is determined from the reduction in the buffer stock made possible by a reduction in the variability in transport time. The method may be used to obtain values of reliability through simulation. An analytical approach for VTTS and VTTV for road freight is developed in (Minken & Samstad, 2006). The authors identify which data on demand and incidents that cause delays are needed; they add that this type of information is necessary independent from which calculation method is chosen.

Use of ADA model Both in Norway (Madslien & Minken, 2011) and Sweden (Edwards & Ramstedt, 2011), it is discussed to use the ADA model system to calculate the benefits due to less devitions in travel time. The idea is to make use of the trade-off between transport costs and inventory costs (that is modelled in the ADA:s logistics model). If the reliability in the transport system is improved, the need for buffer stocks and the inventory costs, both costs related to the ware house building, staff etc. and costs related to the capital tied in the stocks, can be reduced. In principle the benefits due to less variability in transport time can be calculated as the reduction of the inventory costs. Provided that information on demand, standard deviations for lead times and costs that are generated when the goods are not delivered in time is available, calculations can be carried out with help of the logistics model of the ADA system. It has to be taken into account that there are also sources for the uncertainty.

Information on deviations for freight trains In Sweden the spatial, temporal and size distribution of rail freight train delays that influence the transport time has been described by (Krüger, Vierth, & FakhraeiRoudsari, 2012). The fact that the delays are far from normally distributed can indicate a high share of unexpected long delays. (It is important to have in mind that the delays of the trains at the terminals have been measured and not the delays that are experienced by the receivers of the goods.) Another result in that study is that there is a high share of too early trains, due to ‘slack in the time table’.

Behavioural approach Much less SP studies obtain VTTV than VTTS. On the other hand consequences of unexpected very large delays are described in some cases. The delays caused by the accident of a tanker and closure of the river Rhine at the Lorelei Rock in 2011 (Weenen, Quispel, & Visser, 2011) or the delays in the Swedish rail system caused by the extreme winter 2009/2010 (SOU2010:69, 2010) are examples for cases where the costs for carriers, shippers etc. were estimated. The primary sources of unreliability infrastructure, traffic and weather/climate are often inter-related.

Dutch SP study In the Dutch SP study (see above) the main surveys contain SP experiments for VTTV. Special attention is paid to how 'reliability' should be presented to the survey respondents (Tseng, Verhoef, de Jong, Kouwenhoven, & Hoorn, 2008). In-depth interviews with different respondents have been carried out in order to identify the best choice design in the SP games.

Norwegian SP studies In the Norwegian SP study covering all modes (Halse et al., 2010), it is recommended to use standard deviations, if possible the statistical distribution of the transport times should be taken into account. The authors describe, as for the VTTS, the problem that carriers probably have taken into account shippers' costs.

In the study that was limited to rail freight (Halse, & Killi, 2010) a reduction in average delay of one hour (VTTV) is calculated to be worth five to six times more to the companies using the rail freight system than the corresponding reduction in the regular travel time (VTTS). Most of the shippers are more interested in later departures than earlier arrivals of the trains.

15.6. Conclusions

In the context of the socio-economic CBA the values of transport time savings (VTTS) and reduced variability of transport time (VTTV) for freight transport are calculated by taking a company perspective. All relevant benefit components for the companies involved in a transport should be included.

A structure of the benefit components related to VTTS and VTTV for freight is suggested. At the highest level, VTTS and VTTV are distinguished and within VTTS:

- VTTS (T) related to the reduction in the transport resources required to perform certain transports,
- VTTS (G) related to the reduction in the unproductive use of capital tied up in goods while being transported and
- VTTS(O) related to the reduction in the opportunity costs to the users of goods, for example, in form of reduced time to market.

and within VTTV:

- VTTV (T) related to the reduction in the transport resources required to ensure a given service level and

– VTTV (O) related to the reduction in the opportunity costs, for example, in form of reduced stocks.

Typically two types of approaches are applied for determining VTTS and VTTV:

a) a normative approach based on market prices and
b) a behavioural approach which can be based on revealed or stated preferences.

Current practice in relation to CBA and implied VTTS definition differ between countries. For Sweden, Norway, Germany and the United Kingdom the VTTS (T) is derived by market price calculations, though relating to quite different prices. For the Netherlands, an all encompassing VTTS is used, derived from SP data. Sweden (for all modes) and Norway (for rail freight) appear to be the only countries that use VTTS to denote only VTTS (G). There is a risk that different VTTS and VTTV constructions in different countries lead to misunderstandings (double counting etc.) and make the transfer of results between countries or the use of unit values in border crossing projects difficult.

We find that the savings in vehicle operating costs and staff costs (VTTS (T)) can be calculated with help of market based approaches using engineering formulae, transport models etc. Provided that it would be sufficient to limit the VTTS to the goods the VTTS (G) and VTTS (O) can be calculated using a marked based approach. Hopefully lessons can be learned from the SP studies for VTTS and VTTV in Norway and the Netherlands.

Up to now only a few countries have incorporated the benefits according to improved reliability in their CBA guidelines. Provisional recommendations are available in the Netherlands, Sweden (all modes) and Norway (rail freight). The incorporation of realibility of transport times in CBA requires further development of the methods to measure VTTV. The use of information of the trade-off between transport costs and inventory costs (buffer stocks approach) in combination with the logistics model of the ADA model system can be one promising development strategy.

A better understanding of the potential opportunity costs is also needed, for example, whether such effects are relevant for VTTS and VTTV and how they should be dealt with in the context of CBA should be further investigated.

Finally, the ex ante quantification of the expected impacts of policies on transport time and reliability as well as the interaction with passenger transport need to be developed.

Acknowledgements

This chapter is, to a large extent, based on projects carried out for the Swedish Transport Administration (since 1 April 2010) and the former Road Administration and Rail Administration and the ASEK group for CBA guidelines. Henrik Swahn and John Bates were heavily involved in these projects. The author would like to thank Gerard de Jong of Significance, ITS Leeds and CTS Stockholm for comments on an earlier version.

References

Banverket. (1990). *Godskundernas värderingar (Transek, Staffan Widlert och Willy Andersson)*. Borlänge: Banverket.

BMVBS. (2005). *Die gesamtwirtschaftliche Bewertungsmetodik des Bundesverkehrswegeplan 2003*. Bonn: Druckerei des BMVBW (Bundesministerium für Verkehr, Bau und Stadtentwicklung).

Bruzelius, N. (1986). *Företagens MA-kostnader och Företagens kapitalkostnad för fordon, Två ansatser om samhällsekonomiska effekter av vägåtgärder*. Borlänge: Vägverket.

Bruzelius, N. (2001). *The valuation of logistics improvements in CBA of transport investments-a survey*. Stockholm: SIKA (SAMPLAN).

De Jong, G., Kouwenhoven, M., Bates, J., Koster, P.,Verhoef, E., Tavasszy, L, & Warffemius, P. (2012). SP surveys, models and results for the VOT and VOR for freight transport in NL. *Presentation at the European Transport Conference*, 8–10 October, Glasgow.

DFT(UK). (den 8 7 2012). *DFT(UK) Transport Analysis Guidance (Unit 3.5.6)*. Hämtat från http://www.dft.gov.uk/webtag/

Edwards, H., & Ramstedt, L. (2011). *Företagsekonomiska effekter med Samgodsmodellen*. Stockholm: Vectura.

Halse, A. H., & Killi, M. (2010). *Verdsettning av pålitelighet i samfunnsøkonomiska analyser-PUSAM teorigrunnlag*. Oslo: Transportøkonomisk Instiutt.

Halse, A. H., Samstad, H., Killi, M., Flügel, S., & Ramjerdi, F. (2010). *Versetting av famførningstid og påletelighet i godstransport*. Oslo: Transportøkonomisk institutt.

HEATCO. (2006). *HEATCO Developing Harmonised European Approaches for Transport Costing and Project Assessment, Deliverable 5, Proposal for Harmonised Guidelines*. Hämtat från http://heatco.ier.uni-stuttgart.de/deliverables.html

INREGIA. (2001). *Tidsvärden och transportkvalitet-INREGIA:s studie av tidsvärden och transportkvalitet förgodstransporter 1999. Underlagsrapport till SAMPLAN 2001:1*. Stockholm: SIKA.

Jernbanverket. (2011). *BIBLIOGRAPHY Metodehåndbok JD 205 Samfunnsøkonomiske analyser for jernbanen, versjon 3.0 juli 2011. Oslo: Jernbaneverket*.

Kouwenhoven, M., de Jong, G., & Rietveld, P. (2005). *Reliability Ratio's voor het Goederenvervoer (Rapport voor Rijkswaterstaat Adviesdienst Verkeer en Vervoer*. Leiden: RAND EUROPE.

Krüger, N. A., Vierth, I., & FakhraeiRoudsari, F. (2012). Spatial, Temporal and Size Distribution of Rail Freight Train Delays: Evidence from Sweden (unpublished working paper). CTS.

Madslien, A., & Minken, H. (2011). *Dataverktøy for beregning av samfunnsøkonomisk nytte av godstiltak-Forprosjekt*. Oslo: Transportøkonomisk institutt (TøI-rapport 1140/2011).

Ministerie van Verkeer en Waterstaat. (den 6 7 2012). *The Value of Travel Time and Travel Time Reliability: Survey Design*. Hämtat från http://english.verkeerenwaterstaat.nl/english/kennisplein/page_kennisplein.aspx?id=361151&DossierURI=tcm:249-30843-4

Minken, H., & Samstad, H. (2006). *Virkningsberegning av tiltak for raskera og mer påtelig godstransport-en ny metode*. Oslo: Transportøkonomisk Institutt (TøI rapport 825/2006).

OECD/ITF. (2010). *Improving reliability on surface transport networks*. OECD.

RAND Europe, SEO, Veldkamp/NIPO. (2004). *Hoofdonderzoek naar de reistijdwaardering in het goederenvervoer*. Leiden: RAND EUROPE.

Rijkswaterstaat. (den 6 7 2012). *Rijkswaterstaat/kenniscentrum/economische_evaluatie/kengetallen*. Hämtat från http://www.rws.nl/kenniscentrum/economische_evaluatie/kengetallen/

Significance, Goudappel Coffeng and NEA. (2012). *Erfassung des Indikators Zuverlässigkeit des Verkehrsablaufs im Bewertungsverfahren der Bundesverkehrs wegeplanung: Schlussbericht Erstellt für: Bundesministerium für Verkehr, Bau und Stadtentwicklung*. SIGNIFICANCE.

SOU2010:69. (2010). Förbättrad vinterberedskap inom järnvägen.

Statens Vegvesen. (2006). *Konsekvensanalyser, Veiledning, Handbok 140*. Oslo: Statens Vegvesen.

Swahn, H., & Bates, J. (2010, 23 February). *Time and quality in freight transport, a pre-study for VTI, final version*.

Trafikverket. (2012a). *Samhällsekonomiska principer och kalkylvärden för transportsektorn: ASEK 5*. Borlänge: Trafikverket.

Trafikverket. (2012b). *Kapitel 14 Fordonskostnader och transportkostnader godstrafik*. Trafikverket. Retrieved from http://www.trafikverket.se/PageFiles/73641/samhallsekonomiska_principer_och_kalkylvarden_for_transportsektorn_asek_5_kapitel_14_fordonskostnader_och_transportskostnader_godstrafik.pdf

Trafikverket. (2012c). *Samhällsekonomiska principer och kalkylvärden för transportsektorn: ASEK 5, Kapitel 8 Tid och kvalitet i godstrafik*. Trafikverket (http://www.trafikverket.se/PageFiles/73641/samhallsekonomiska_principer_och_kalkylvarden_for_transportsektorn_asek_5_kapitel_8_tid_och_kvalitet_i_godstrafik_.pdf).

Tseng, Y.-Y., Verhoef, E., de Jong, G., Kouwenhoven, M., & Hoorn, T. v. (2008). A pilot study into the perception of unreliability of travel times using in-depth interviews. *Journal of Choice Modelling*, 2(1), 8–28. http://www.jocm.org.uk/index.php/JOCM/article/viewFile/69/29

Weenen, R. d., Quispel, M., & Visser, J. (2011). *Closure of river rhine at the lorelei rock: Estimate of impact and allocation of damage*. Zoetermerr: NEA.

Vierth, I. (2010). *Värdering av minskad transporttid och minskad variation i transporttid för godstransporter*. Stockholm: VTI (Rapport 683).

Vierth, I. (2012). *Värdering av tidsvinster och högre tillförlitlighet för godstransporter, Kunskapsinventering på uppdrag av ASEK 5-gruppen (kommande VTI-Notat)*. VTI.

Vägverket. (1992). *Godskundernas transportmedelsval (Transek, Staffan Widlert & Ulrika Lindstedt)*. Borlänge: Vägverket (VV 1992:25).

Chapter 16

Freight Transport Pricing Models

Siri Pettersen Strandenes

Abstract

Freight transport prices typically increase proportionally with distance and with weight or size. Trade models traditionally apply iceberg transport costs, where the value of the good transported upon arrival is reduced by a percentage of the initial value to reflect transport and other trade costs, whereas freight charges in transport and logistics models reflect physical characteristics such as weight and distance. We also show that non-linear pricing models can be applied to build priority pricing schemes where shippers pay a surcharge for priority handling of their cargoes. Basing point pricing and unified delivered prices often used in freight transport pricing imply spatial price discrimination. Yield management developed by passenger airlines upon deregulation of the airline industry in USA has spread to air cargo. EU recent prohibition of liner conferences makes yield management more probable also in container shipping.

Keywords: Freight transport; freight rate structure; pricing models; non-linear pricing; spatial price discrimination; two-part tariff; iceberg transport costs

JEL codes: R40; L91; D4

16.1. Introduction

Traditionally most transport modes operators charged users per kilometre. Customers might choose different quality, typically first or second class in passenger transport and express or basic in freight transport. Passenger transport in addition

Freight Transport Modelling

ISBN: 978-1-78190-285-1

offered discounts based on social criteria such as lower prices for youth or elderly people. These discounted prices followed the same structure as the ordinary prices and where typically percentage discounts on the base or standard price. In freight transport the operators offered discounts for regular customers sending large volumes.

Neither of these traditional pricing models fully follows the recommendation from economic theory to use marginal cost pricing. For presentations and discussion of marginal cost pricing in transport see Button (2010) and De Borger and Proost (2001). Even though distance-based pricing implies a certain element of marginal cost pricing, and volume discounts may reflect some economies of scale effects of big unit loads, this pricing model disregards external costs from transport either to other users or to the environment. Pricing according to distance furthermore disregards capacity utilization and capacity limits.

In passenger transport, especially in air transport, deregulation induced large changes in pricing structures. Instead of a set of percentage discounts based on social criteria, air fares now more closely reflect marginal costs of an extra passenger and the fact that this extra cost varies across day, week and year. Air transport similar to most other transport modes; enjoy economies of scale. Adding an extra passenger to a flight reduces unit costs on that flight. In an airline network adding a new link may reduce unit costs, for example, by higher exploitation of passenger and cargo handling facilities at the airports. To cover costs and create a profit the airline industry introduced price discrimination, not based on social criteria such as age, but by trying to capture the consumer surplus. That is, to set prices according to passengers' willingness to pay above the marginal costs-based airfare.

Even where deregulated, the traditional pricing models used in freight transport with charges proportional to distance and weight remained. Price discrimination based on the value of the goods transported, however, came under pressure from standardized packaging, especially with the introduction of containers. The classical value-based price discrimination only implicitly reflected the assumed higher willingness to pay. It echoed the higher time costs for high value goods in transit, but not the individual cargo owner's willingness to pay. Even low value cargoes may pose high time costs for the cargo owner and high value cargos may pose a lower than assumed time costs, when, for example, transit time reduces owners' storage costs. Fries et al. (2008), for example, finds that price and on-time reliability are more relevant to cargo owners than transport time. Cargo owners alternatively may choose to transport their goods for their own account either by lorry or ship, to control the quality of transport both regarding punctuality, duration and cargo handling. A transport provider might be able to attract such cargoes by differentiating the services and prices.

The structure of this chapter is as follows. The next section presents freight pricing structures employed in economic models, both trade and transport models. The following section gives an example of a pricing model in use by a transport service provider. Following this is a section on non-linear pricing models for transport infrastructure exemplified by port pricing. The last section before summing up, comments on spatial price discrimination and yield management in freight transport.

16.2. Transport Service Pricing Used in Trade, Transport and Logistics Models

Trade models typically disregarded the structure of transport costs setting trade and transport costs either to zero (for traded goods) or to infinity (for non-traded goods). Samuelson (1954, p. 268) introduces transport costs in trade models stating that 'To carry each good across the ocean you must pay some of the good itself. Rather than setting up elaborate models of a merchant marine, invisible items, etc., we can achieve our purpose by assuming that just as only a fraction of ice exported reaches its destination as unmelted ice, … '. This later was coined iceberg trade costs. Iceberg trade costs thus are proportional to the volume transported and to time or distance, but do not reflect the costs to the transport provider from moving the cargo. In Samuelson's notation domestic price ratios for each of the trading countries in a 'two goods – two countries' model with iceberg transport costs, differ by these transport costs;

$$\begin{aligned}&\text{(a) } \left(\frac{P_c}{P_f}\right)_A = \frac{X}{a_y Y} \text{ and (b) } \left(\frac{P_c}{P_f}\right)_E = \frac{a_x X}{Y}\\&\text{thus } \left(\frac{P_c}{P_f}\right)_E = a_x a_y \left(\frac{P_c}{P_f}\right)_A < \left(\frac{P_c}{P_f}\right)_A \text{ and } \left(\frac{P_f}{P_c}\right)_A = a_x a_y \left(\frac{P_f}{P_c}\right)_E < \left(\frac{P_f}{P_c}\right)_E\end{aligned} \tag{16.1}$$

where $X =$ the country A's food (f) exports, $Y =$ country E's clothing (c) exports, $(P_c/P_f)_A =$ domestic price ratio of cloth and food in country A, $(P_c/P_f)_E =$ domestic price ratio in country E, $a_y < 1$ is the fraction of clothing export (Y) from country E that reaches country A and $a_x < 1$ is the fraction of food export (X) from country A that reaches the destination in country E. That is, because of transport costs, goods are relatively cheaper in their country of origin.

Iceberg trade costs became the standard modelling approach for trade and transport costs in international trade models. New economic geography (Krugman, 1991) introduced imperfect competition and increasing returns into international trade theory, but kept the traditional assumption of iceberg transport costs.

Some examples exist using other assumptions on the transport costs. One example is Ottaviano, Tabuchi, and Thisse (2002) who assume transport costs absorb resources that differ from the transported good itself. Thus, transport costs do not necessarily reflect the value of the good transported as do iceberg transport costs, but they assume that transport costs still are proportional to the traded volume. Ottaviano et al. (2002) in their model assumed, however, that transport costs for front and back haul are the same. Behrens & Picard (2011) refined the model by opening up for asymmetry in transport costs between the countries, but they also model transport costs proportional to the traded volume. Thus,

$$\pi_i = p_{ii}Q_{ii} + (p_{ij} - t_{ij})Q_{ij} - fr_i \quad \forall i,j \tag{16.2}$$

where $\pi_i =$ profit for a manufacturing firm in country i, $p_{ii}Q_{ii} =$ operating profits from local sales, $(p_{ij} - t_{ij})Q_{ij} =$ operating profits from exports from country i to country j, t_{ij} = transport costs for exports from country i to country j, and fr_i = the firms fixed production factor. Hence, when t_{ij} differ from t_{ji} there is an asymmetry in the transport costs between the countries.

Transport firms are usually not explicitly modelled in trade models. One exception is Francois and Wooton (2001, 2008) who model the transport sector as an intermediate industry providing services to exporters and importers. They focus on the effects of market power in the transport sector for trade and find that the more imperfect the competition in the transport industry, the higher is the impact of the industry on producers and consumers. They model transport costs or the shipping margin as the difference between the CIF and FOB prices after separating out an ad valorem import tariff. In one version of the model they also take the cost of domestic distribution and transport activities into consideration. The assumption on ad valorem transport costs remains, however.

Hummels (2008) points at an interesting difference between trade and transport specialists. Trade specialists think of transport costs and other impediments to trade in ad valorem terms. That is, they model the necessary transport costs relative to the value of the goods transported. Iceberg transport costs have this characteristic. Transport specialists on the other hand according to Hummels (2008) value transport costs per unit of cargo transported either per tonne-km or the cost of moving one TEU. Both specialist groups generally model transport costs as linear and proportional to distance disregarding the cost structure of the transport providers, however.

Transport models also by and large focus on other aspects than pricing structure. Most models specify the price or freight rate in terms of cargo weight and transport distance and assume a proportional relationship between the freight rate and distance and weight. In their review of freight models Wigan and Southworth (2006) considers the range of applications and issues associated with different categories of freight models. Their main concern is the different models' suitability in planning and forecasting. They aim at highlighting gaps in the coverage of existing models relative to the analytical needs. They point out that among other aspects, the current models do not enlighten planners either on the shipping, carriers or logistics agents' response to regulatory, pricing of infrastructure investment policies. But, pricing of the transport service is not put forward explicitly as one of the gaps between existing models and analytical needs.

Wigan and Southworth (2006) list the modelling techniques to estimate volume, cost, and environmental and safety related impacts associated with physical freight flows. Among those mentioned are engineering cost-based, statistical (regression)-based, least cost-based single and multiple path freight traffic routing models, network-based spatial pricing equilibrium models, and hedonic freight pricing models. Hedonic pricing implies that both transport service characteristics and factors related to the customers' valuation of the service are included in the pricing model. Hence, the hedonic pricing method should help reveal customers' willingness to pay for the transport service and thus approximate the values reflected in customer choices or revealed preferences.

Discussions of transport pricing issues, both for transport service and for transport infrastructure, usually link them to investment in and capacity of infrastructure. Since transport services and infrastructure both feature economies of scale and government induced price distortions, transport does not fulfil conditions for first best marginal cost pricing, and second-best pricing Ramsey type rules have emerged. Ramsey pricing implies that price mark-ups set by firms offering transport services to different shippers should be set inverse to shippers' price elasticity of demand:

$$\frac{(P_i - C_i)}{P_i} = \frac{\lambda}{e_i} \tag{16.3}$$

where $P_i =$ price charged shippers (i) per unit of transport service, $C_i =$ marginal costs of transport service for shippers (i), $e_i =$ the elasticity of demand for shippers (i) and $\lambda =$ is a constant < 1, similar for all shippers and determined by the requirements for recovering fixed costs. This implies equal mark-up weighted by the shippers' price elasticity, for example, for express and standard services.

$$\frac{(P_i - C_i)}{P_i} e_i = \frac{(P_j - C_j)}{P_j} e_j \tag{16.4}$$

De Borger and Proost (2001) introduce an externality-corrected Ramsey price rule for freight charges in their interregional passenger and freight model developed to analyse optimal pricing and regulatory policies for interregional transport. Winston (1985) points out that Ramsey pricing poses equity problems and that charging customers with relatively inelastic demand more, introduces dynamic problems since this creates incentives among customers to change their underlying demand elasticities. Blauwens, De Baere, and Van de Voorde (2012) while agreeing that economies of scale necessitates a deviation from marginal cost pricing, point out that the joint product situation common in transport does not require deviation from marginal cost pricing. Front haul and back haul transport legs represent joint products and when the cargo flows differ between these legs, the freight rates should reflect the different marginal costs and thereby the capacity utilization rates.

Transport is an integrated part of logistics and Lamberts et al. (1998) provide an overview of different factors that influence transport cost and prices. They review several price models used in transport. Such influence partly comes from product-related factors such as stowability, weight–volume ratio, potential liability for high cargo value and handling requirements, and partly from market-related factors such as competition, distance, border crossing, transport flow imbalances, regulation and seasonality. The third group of factors is service quality and mirrors quality factors like punctuality, reliability, flexibility, market coverage, and loss and damage performance.

Focusing on pricing models in logistics Lamberts et al. (1998) separate 'Cost-of-service' pricing models and 'Value-of-service' pricing models. Cost-of-service pricing gives transport costs that vary with distance and volume and that are set to cover fixed and variable costs and some profit. Value-of-service pricing on the other hand implies charging what the market will bear. Within these two pricing approaches transport providers use different freight rate categories, such as line-haul rates, accessorial charges and rates for special instances. The assessorial charges cover payments made for non-line-haul services needed to fulfil the transport assignment. Pricing models set to cover line-haul-costs may be either (1) class rates, which groups products in classes for pricing purposes, (2) exception rates as when offering a discount to stay competitive, (3) commodity rates charged for specific point-to-point transport services typically used on dense routes, or (4) miscellaneous rates. By miscellaneous rates they mean contract rates specially negotiated with cargo owners and freight of all kind (FAK) rates that apply to shipments and not products. Such contract and FAK rates offset freight rates based on classes of products.

The above discussion illustrates the complex freight rate structure and the many differing pricing models employed in freight transport. The contract and FAK rates furthermore show that standardization by containerization reduces the freight differentiation traditionally seen in freight pricing models.

In addition to this pricing complexity the direct payment for transport charges and handling cost can be split differently between the transport firm and the buyer of the freight services conditional on whether they agree on free on board FOB or cost insurance freight (CIF) pricing. FOB in essence is a delivered pricing model and may be calculated using zone pricing, basing point pricing or uniform delivered pricing. These three models differ on the distance element in the freight charges. In zone pricing each zone gets a specific delivery charge, whereas basing point pricing calculates freight charges from a fixed geographical point to the location of the buyer. The basing point may be another location than the origin of the goods sold and represents a simplification of distance based fare calculations.

16.3. A Real-World Pricing Model Example

We have seen that transport models typically apply pricing structures with simple linear prices that are proportional to distance and increasing in weight or number of units. Hence, the modelled prices do not reflect the cost structure transport providers face. The following example from a freight transport provider, also show prices that increase in distance and weight, but the firm charges two-part tariffs.

Two-part tariffs usually have a fixed and a variable price element, were the variable element is proportional to quantity and/or distance. The fixed element sometimes may function as a block element that shifts the transport price schedule according to transport zone or distance, as an alternative to letting the variable element reflect both weight and distance. Thus,

$$T(q) = A_z + pq \quad \forall z = 1, \dots n \tag{16.5}$$

where T = total payment, q = weight class, p = unit price and z = zone.

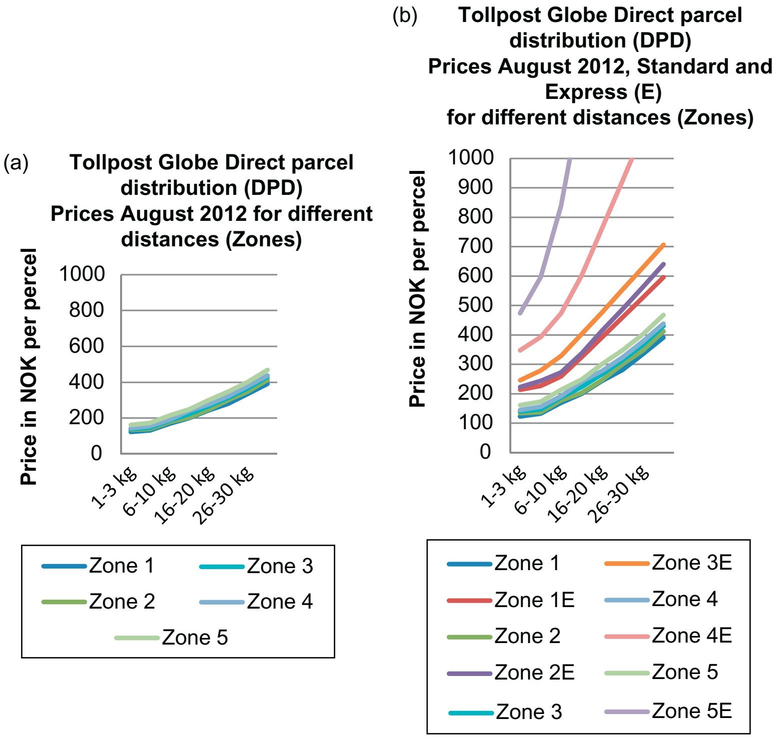

Figure 16.1: Panel (a) Tollpost Globe price structure. Panel (b) Tollpost Globe express parcel tariffs.

Tollpost Globe is a Scandinavian transport provider. Figure 16.1 panel (a) illustrates their tariffs for direct parcel distribution for standard forwarding over different distances standardized into different zones in Norway. We see that Tollpost Globe offers a two-part tariff, first an initial minimum charge changing slightly with distance or zone, plus a tariff that increase with parcel weight class. The shift in the transport price schedule reflecting zone is modest, whereas the proportional increase with weight is stronger in Tollpost Globe pricing schedule.

The firm also provides express parcel service and Figure 16.1 panel (b) illustrates the express parcel tariff structure. Express parcels face a similar two-part tariff structures, but the weight-based tariff correction is stronger for express than for standard parcel forwarding. The tariffs for express parcels also shift more with distance or zone and especially so for the furthest zones.

16.4. Non-Linear Pricing Models

Above we showed an example of a currently used two-part freight pricing model. There are several alternative pricing models available. Wilson (1993) presents

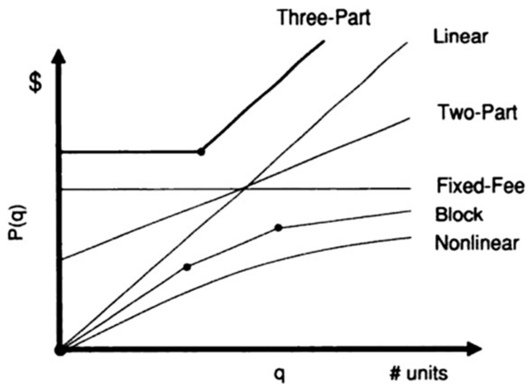

Figure 16.2: Pricing models or tariffs. *Source*: Wilson (1993, Fig. 1.1).

alternative pricing models used in regulated industries, both transport industries and other industries such as electricity provision, telephone companies, airlines, car rental and newspaper advertising. He points out that the way quantity affects prices paid differs among the different industries. Thus, in addition to two-part tariff pricing models, he finds several other forms of non-linear pricing models or tariffs that deviate from proportionality to the quantify bought. Figure 16.2 depicts different pricing models' schedules. The simplest is a fixed fee or a flat rate tariff irrespective of the number of units bought.

To capture customers' differing valuation of service qualities, transport firms may use non-linear pricing where, for example, increments in quality are priced above standard service. The above example from Tollpost Globe illustrates higher price for express delivery. Transport firms may offer several classes of fast delivery and charge prices above the standard delivery price that raises more for the speediest delivery than for an intermediate speed delivery. Hence, the non-linear pricing schedules allow the firm to offer a menu of differently priced services among which the customers may choose according to their willingness to pay. Thus, non-linear pricing models are especially suitable when the firms wish to cater to and capture extra revenue from heterogeneity among their customers.

Priority pricing opens up for differentiating delivery conditions according to priority or reliability of delivery. This is a case of quality pricing were reliability of delivery is the quality variable. Punctuality and speed are related quality variables, and so is quality of cargo handling that reduces the incidents of breakage and the following delays.

Express delivery at a higher cost is traditional in freight transport, parcel and post services. By offering different express services in addition to the standard service for customers to choose, the transport firm offers a menu of contracts to customers sensitive both to transport and time costs.

A two-part tariff may provide such priority pricing schemes by charging an access due and a priority surcharge, where the priority surcharge 'guarantees' reliability, punctuality and duration. Duration includes transport time but also handling and waiting. Maersk Line introduced such priority option against a charge in 2010. Named 'Priority Product', this contract offers priority loading against an extra charge. Strandenes and Marlow (2000) suggest priority pricing for port access, where all entering vessels pay an access fee. Against an extra charge ship owners obtain guarantees on the throughput time for the cargo and on the turnaround time for the vessel.

Wilson (1993) develops a pricing scheme for parcel deliveries for different classes or speed of delivery:

$$P(w|s) = A(s)w^b \tag{16.6}$$

where $P =$ price as a function of weight (w) and speed or class (s), $A(s) =$ customers higher valuation of higher speed and $w^b =$ the rise in the base freight tariff when parcel weight increases.

In freight terminals or ports where vehicles or vessels are queuing, giving priority to some induces longer waiting on others. Wilson (1993) also develops priority pricing in a service system for jobs that require time to be processed and where capacity is limited, so that other jobs must queue waiting for the processed job to finish. In this model low-priority jobs wait in queues while higher-priority jobs are handled. The different jobs also may require different service times. The aim is to find maximum total surplus, when processing times for the different jobs affect the efficient service scheduling.

For a set of assumptions not specified here, Wilson (1993) finds a pricing scheme that produces optimal priority schedules and submission rates:

$$P_k(t) = A_k t + \frac{1}{2} B t^2 \tag{16.7}$$

where $P_k(t) =$ the charge for a (k) priority job of duration (t), $A_k t =$ the delay costs imposed on other jobs with the same or lower priority compared to the one given priority, and $\frac{1}{2}Bt^2 =$ the waiting cost imposed on those new jobs that arrive while the on-going job is done. Hence, this pricing scheme caters both for differences in service time and differences in waiting costs for the jobs.

The above pricing scheme indicates that freight transport may cater for the external cost from delaying other freight by implementing a two-part tariff reflecting customer priorities from differing time costs.

16.5. Price Discrimination

Freight absorption by the producer is a kind of spatial price discrimination. Tirole (1988) illustrates multimarket or third-degree price discrimination by a spatial price discrimination example. He models a linear city with a monopolist producer where all customers buy the same amount at a given delivered price. All customers in the same location also pay the same delivered price. These simplifications allow us to focus on spatial differences.

The monopolist maximizes monopoly profit in the market formed by customers located in x:

$$(p - tx - c)D(p) \tag{16.8}$$

where $p(x)$ = the delivered price, x = distance to the customer, c = production cost per unit, tx = transport costs to location x and $(D(p) = q)$ = demand.

In this model, $(p(x) - tx)$ would be the FOB price. The transport costs have similar effects as an excise tax and represent a wedge between consumer or delivered price $(p(x))$ and producer or FOB price $(p(x)–tx)$.

A rise in the distance to the customer implies an upswing in the delivered price $(p(x))$ since the monopolist will be able to raise the price $(p(x))$ charged in order to cover both the higher transport costs (tx) and production cost (c). The rise in price comes at the cost of a reduction in the volume sold. When demand $D(p)$ is linear the rise in the optimizing monopolist's delivered price will be $(dp = t\ dx/2)$ reflecting this effect on the marginal revenue. Hence, the customer does not pay the full transport cost and the monopolist practices freight absorption. The monopolist discriminates against the closest customers. This discrimination furthermore is resistant to arbitrage between customers in different locations.

The freight absorption depends on demand conditions in the same way as does an excise tax. With constant-elasticity-of-demand instead of linear demand as assumed above, the effect on the delivered price for larger distance is $dp = tdx/(1–1/\varepsilon)$, and the monopolist instead discriminates against his farthest customers. In this case spatial arbitrage may counteract the discrimination since customers close to the plant will gain by buying the good and transporting it to customers further away, as long as the transport costs are not higher for them than for the monopolist. The third alternative is exponential demand. Then the effect of longer distance on the delivered price is $dp = t\ dx$, and there is no spatial discrimination.

Freight absorption is not uncommon in practice. Uniform delivered price and basing point pricing are examples. Absorption does not solely depend on the demand function, however. As Tirole (1988) points out demand conditions may differ among customers located close to the producer and those further away. The last group of customers may have other close suppliers and thus better substitution possibilities than the customers in the neighbourhood of the producer. So, if the demand elasticity increases with distance this opens up for spatial discrimination by freight absorption.

Yield management exploits price discrimination possibilities. Transport services fulfil other requirements for applying yield management techniques such as perishable service, limited capacity or finite selling horizon, price-sensitive customers and stochastic demand. Yield management is mainly relevant in industries with low variable relative to fixed costs, that is, for industries where a wide range of prices may cover variable costs.

Yield management techniques were initiated by the airline industry after deregulation of the industry in USA. Now yield management is used by most airlines and also by other transport or transport-related industries such as passenger railways, car rental and the cruise industry. These are passenger industries, but now some freight transport firms also employ yield management techniques, for example, container transport and air cargo or parcel distribution.

Passenger airlines offer high frequency of flights to attract passengers with a high value on reducing total travel time. For most airline markets such high frequency of flights produces excess capacity. To fill all the flights airlines offer several different fare classes on each flight. They introduced yield management techniques to find the most profitable allocations of seats to the different fare classes, and restrictions on low fare tickets to stop passengers with high willingness to pay from switching to low fare tickets. In the airline industry such restrictions typically reduce the flexibility for the passenger by introducing advance purchase, non-refund ability and other elements that reduce the quality of the service.

Air cargo firms also apply yield management. Yield management for cargo transport increases the complexity compared to passenger transport yield management, however. As Pak and Dekker (2004) point out, capacity limits for cargo are two-dimensional, that is, both weight and volume. In addition, different cargoes generate different packing and handling cost, whereas for passenger such differences are insignificant. Air cargo providers face additional challenges from not knowing until just before departure how much space is available, since passenger aircraft also carry air cargo and they do not know the volume of passenger baggage beforehand. Even so yield management now is well established in air cargo, while container shipping lags behind (Zurheide and Fischer (2012). They point at the changes in the container market following EU's decision to end exemption from the competition rules for liner conferences' from October 2008, and argue that this induces major changes to pricing in the container industry. They draw a parallel to deregulation in the US airline industry in the years after 1978, which induced the incumbent American airlines to implement the yield management systems.

Zurheide and Fischer (2012) simulate the effects of a container booking limit strategy where some slots are kept open for priority containers. Priority containers are charged a fixed extra fee above the normal tariff. Hence, this models a simple yield management system. They show simulation results on a fictional liner shipping network with booking limits for ordinary containers to keep space for priority containers that may be booked later. Compared to the ordinary first-come-first-served booking scheme, introducing such booking limits may raise the profit when capacity is limited. When capacity is ample, the ordinary first-come-first-served scheme is better. Maersk Line's 'priority product upgrade' mentioned above has a yield management element since the extra charge shippers face varies with demand and the available vessel space on the specific route (Maersk, 2010).

16.6. Concluding Remarks

In this chapter, we have discussed freight transport pricing models. Most freight transport firms apply pricing proportional to weight and distance. In trade and transport models the freight rates also are modelled fairly simply either as iceberg transport costs where the value of the good falls in proportion to distance moved or in proportion to distance and weight.

We discussed non-linear pricing models and how they can be used for priority pricing in freight transport. Spatial price discrimination takes place in freight transport pricing. Air cargo has developed methods for yield management and after EU prohibited liner conferences this deregulation probably will induce a wider use of yield management methods in container shipping. Yield management for cargo is more complicated than for passenger transport, since cargoes differs more than passengers both in volume and size.

References

Behrens, K., & Picard, P. M. (2011). Transportation, freight rates and economic geography. *Journal of International Economics*, *85*, 280–291.

Blauwens, G., De Baere, P., & Van de Voorde, E. (2012). *Transport economics* (5th ed.). Uitgeverij De Boeck NV: Antwerpen.

Button, K. (2010). *Transport economics* (3rd ed.). Cheltenham: Edward Elgar.

De Borger, B., & Proost, S. (2001). Transport problems a: The economic diagnosis. In B. De Borger & S. Proost (Eds.), *Reforming transport pricing in the European Union*. Cheltenham: Edward Elgar.

Francois, J. F., & Wooton, I. (2001). Trade in international transport services: The role of competition. *Review of International Economics*, *9*(2), 249–261.

Francois, J. F., & Wooton, I. (2008). Market structure and distribution services, goods trade, and the effects of liberalisation, in OECD/ITF benefiting from globalisation, transport sector contribution and policy challenges, 17th ITF symposium 25–27 October 2006, Paris, OECD, 99–130.

Fries, N., et al. (2008). Modal spilt functions for a Swiss national freight transport model. European Transport Conference, Leyden.

Hummels, D. (2008). Global trends in –trade and transportation, in OECD/ITF Benefiting from globalisation, transport sector contribution and policy challenges, 17th ITF symposium 25–27 October 2006, Paris, OECD, 17–35.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, *99*, 483–499.

Lamberts, D. M., et al. (1998). *Fundamentals of logistics management*. Boston: McGraw-Hill International Editions.

Maersk Line. (2010). The answer is "YES", Maersk Line News 1st June. Retrieved from http://www.maerskline.com/print/?node=2866385&page=news&path=/news/story_page/10/yes&lang=en_US. Accessed 20121205.

Ottaviano, G., Tabuchi, T., & Thisse, J-F. (2002). Agglomeration and trade revisited. *International Economic Review*, *43*(2), 409–435.

Pak, K., & Dekker, R. (2004). Cargo revenue management: Bid prices for a 0-1 multi knapsack problem. Research Paper ERS-2004-055-LIS, Rotterdam, Erasmus Research Institute of Management (ERIM).

Samuelson, P. (1954). The transfer problem and transport costs, II: Analysis of effects of trade impediments. *The Economic Journal*, *64*(254), 264–289.

Strandenes, S. P., & Marlow, P. B. (2000). Port pricing and competitiveness in short sea shipping. *International Journal of Transport Economics*, *XXVII*(3), 315–334.

Tirole, J. (1988). *The theory of industrial organization*. Boston, MA: MIT Press.

Wigan, M. R., & Southworth, F. (2006). What's wrong with freight models? and what should we do about it? Paper presented at the TRB 85th Annual Meeting 2006, Washington DC.

Wilson, R. B. (1993). *Nonlinear pricing*. New York, NY: Oxford University Press.

Winston, C. (1985). Conceptual developments in the economics of transportation: An interpretive survey. *Journal of Economic Literature*, *23*(1), 57–94.

Zurheide, S., & Fischer, K. (2012). Revenue management slot allocation model. *Maritime Economics and Logistics*, *14*(3), 334–361.

LOCAL (URBAN)

Chapter 17

Urban Freight Tour Models: State of the Art and Practice

José Holguín-Veras, Ellen Thorson, Qian Wang, Ning Xu, Carlos González-Calderón, Iván Sánchez-Díaz and John Mitchell

Abstract

The main objective of this chapter is to provide a comprehensive overview of empirical findings and models that focus on urban freight tours. In doing so, the chapter provides background information; reviews the research literature; classifies the tour models in simulation, hybrid, and analytical models; discusses strengths and weaknesses of a sample of key models; and concludes with a statement of chief findings.

Keywords: Freight demand model; tour model; urban freight

17.1. Introduction and Background

The phenomena that create freight demand are very different, and more complex, than the ones observed in passenger demand. To start with, freight demand is the result of the economic interactions of the multiple agents that participate in supply chains, including producers, shippers, carriers, freight forwarders, warehouses, receivers, as well as the public agencies that regulate various aspects of the system. In the context of this chapter, however, the role of producers and shippers is subsumed into the *supplier*, which is assumed to undertake the functions of producing and shipping the goods. The multiplicity of agents leads to a situation in which proper modeling of the freight demand process requires consideration of how their interactions shape freight demand. However, since considering all these interactions is a practical impossibility due to the mathematical intractability of the resulting

Freight Transport Modelling

ISBN: 978-1-78190-285-1

models, and lack of data about the interactions, simplifications are warranted. For that reason, the bulk of the research focuses on producers, shippers, carriers, and receivers as they play the key roles. Adding even more complexity to the mix, these interactions are governed by the tactical and strategic business considerations required to survive and thrive in an environment of market competition. The fact that producers, carriers, and receivers compete with others of their kind in a market, influences their behavior because they have to take into account what the competition would do (as opposed to optimizing their operations in isolation of others). Disregarding the effects of market competition is likely to lead to results that do not reflect the realities of urban freight markets.

The type of commodity being transported is an influential determinant of behavior. This is because the commodity type captures the effects of attributes such as degree of perishability, the handling technology required to transport the cargo, opportunity costs of the shipment, and others. In essence, the commodity type is a proxy for the industry segment in which the company operates. Not surprisingly, freight behavior research has identified the commodity type as a key explanatory variable (Holguín-Veras, 2002; Holguín-Veras, Silas, Polimeni, & Cruz, 2007; Holguín-Veras & Wang, 2011; McFadden, Winston, & Boersch-Supan, 1986).

The nature of the relationship among the economic agents involved in a supply chain is also a factor. The joint behavior of the agents that belong to the same company is different than the one in which they were independent companies. In the case of the integrated operations, all associated benefits and costs are internalized by the parent company. Thus, managers can decide on what is the overall best strategy and implement it. However, in independent operations, each agent is attempting to maximize their own returns. In doing so, they must account for both market competition, and the power relations that link it to the other participants in the supply chain. For instance, a powerful receiver could extract concessions from carriers and producers that a weaker one could not. This is important because, as a result of the deregulation of trucking, carriers have become the weakest element of the supply chain and must operate within the constraints imposed by both producers and receivers. In fact, the producers' decision concerning shipment size is so important that is "mode determining" (Holguín-Veras, 2002; Samuelson, 1977). In essence, contrary to what most individuals believe, freight mode choice is implicitly made by the suppliers when they decide on shipment size, not by the carriers that only make a choice that depends on the suppliers'. Similarly, the receivers' decisions regarding delivery time determine carriers' time of travel and how they respond to freight road pricing (Holguín-Veras, 2008; Holguín-Veras, 2011; Holguín-Veras et al., 2007; Holguín-Veras, Silas, Polimeni, & Cruz, 2008b). The large number of possibilities, as shown in Figure 17.1, is one of the factors that explain the multitude of behaviors exhibited by the urban freight system.

In response to the outcomes of these interactions, a set of commodity flows from producer sites to receiver locations is generated. This, in turn, leads suppliers to undertake complex tours by which they pick-up and deliver supplies to their customers. Modeling these tours is a very complex undertaking. Although technically speaking, the term urban-freight-vehicle-tour is the best descriptor, for simplicity

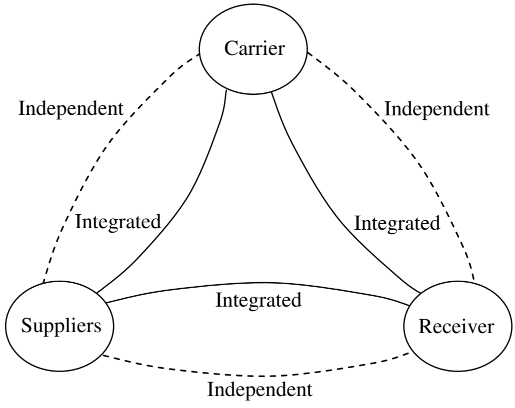

Figure 17.1: Potential links for a case with a supplier, a carrier, and a receiver.

sake the paper uses ''urban freight tour'' (UFT), ''freight tour,'' or simply ''tour'' to refer to the sequence of stops that freight carriers undertake to pick-up or deliver cargo and then return to the home bases.

The picture that emerges from this discussion is one of great complexity, in which multiple dimensions are at play: all participants are trying to maximize their profits (a financial metric), by trading commodities (the demand dimension), which leads them to generate freight tours (the logistical dimension) (Holguín-Veras & Patil, 2007; Holguín-Veras & Patil, 2008). All of this takes place in an environment of competition where not only individual agents compete with others of their kind, entire supply chains compete with other supply chains for market supremacy.

Modeling UFTs has to take into account, not only the complex dynamics that have just been described, but issues such as mathematical tractability, data constraints, and knowledge gaps regarding the dynamics that drive freight demand. In spite of these obstacles, significant progress has been made. The main objective of this chapter is to succinctly describe the state of the art and practice of UFT models. The discussion centers on system-wide freight demand models aimed at describing and forecasting UFTs. Operational and tactical models, for example, routing, are not discussed.

In addition to this introduction, the chapter contains five sections. Section 17.2 discusses basic concepts to provide a common base of understanding; Section 17.3 summarizes empirical results from the literature; Section 17.4 discusses key models; Section 17.5 focuses on models that could be considered representative of modeling typology; while Section 17.6 summarizes the work done.

17.2. Basic Concepts

To start with, it is important to establish some basic definitions to ensure the use of a uniform set of terms. Figure 17.2 shows the case of a supplier that sends shipments

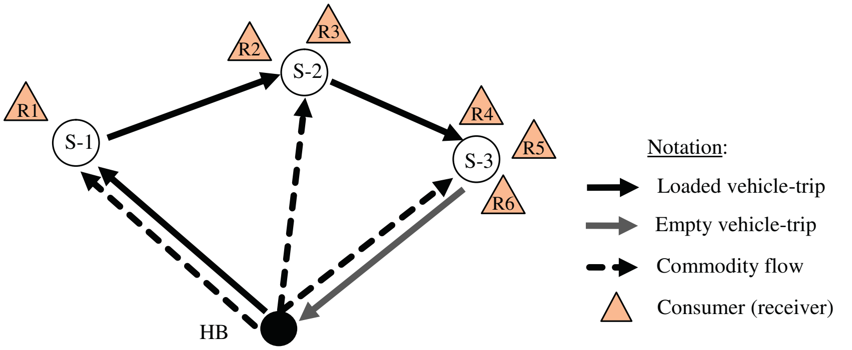

Figure 17.2: Vehicle trips, commodity flows, and delivery tours.

(commodity flows) from its home base (HB) to six receivers (R1, R2, R3, R4, R5, and R6) by means of a tour that is comprised of three individual trips (with stops at S1, S2, and S3) to deliver the cargo to the receivers (HB-S1, S1-S2, and S2-S3) and one empty return trip to the base (S3-HB). A number of important remarks can be made: (1) the UFT is comprised of multiple individual trip segments; (2) each one of these trip segments could be used to make deliveries to multiple receivers; (3) the tour generates both loaded and empty trips; and, (4) the commodity flows rarely run in the same direction as the vehicle-trips.

These features reflect the fact that the UFT is the logistical response to the need to transport cargo from supplier to multiple consumer locations. In essence, the underlying supply-demand relation is one of production and consumption linking the producer at HB to receivers. In contrast, the trip origins and destinations do not necessarily reflect production–consumption relations, as, in most cases, they are the result of a decision concerning vehicle routing. The Holy Grail in urban freight tour modeling is to capture both the underlying economic drivers of production and consumption and realistic delivery tours. The former represent the demand, while the latter is the physical expression of the vehicular supply.

The trade patterns that arise from these economic interactions depend on the industry segment in question. The few data available seem to indicate that the tour lengths (measured in number of stops) exhibit a wide range, where larger and more congested cities tend to have longer tours than smaller and less congested ones. The data also suggest that the type of operations and commodities transported influence tour behavior. The results from the literature are summarized in Section 17.3.

The research conducted on UFT can be classified depending on its focus on either characterization or modeling. The former focuses on describing the key features of UFTs. The latter intends to create computer systems and mathematical models that could be used to describe and forecast UFTs in cities and metropolitan areas. The remainder of the chapter discusses the work conducted in both areas and introduces selected models to help the reader understand the differences among the model classes.

17.3. Characterization of Urban Freight Tours

The research that focuses on characterizing UFTs is very important because it provides the foundation for modeling efforts by both describing real-life features and gaining insight into the inter-relations and influencing factors to be incorporated in UFT models. This section provides a summary.

Vleugel and Janic (2004) collected data from four different shopping areas in Amsterdam, Alphen aan den Rijn, Apeldoorn, and Rotterdam (the Netherlands). The purpose of the study was to shed light on route choice made by truck drivers and to study how urban policy impacts that choice. They found that trucks made in average 3.90 stops/tour in Alphen, 6.20 stops/tour in Amsterdam, 4.70 stops/tour in Apeldoorn, and 1.80 stops/tour in Schiedam. Their research revealed that, in many cases, some local policy measures have a negative impact on the efficiency of deliveries.

Holguín-Veras and Patil (2005) conducted a comprehensive analysis of trip diary data collected from a sample of commercial vehicles in Denver, Colorado. They found that the average number of stops per tour for single and combination trucks are 6.5 stops/tour and 7.0 stops/tour, respectively (some reported up to 19 stops). They analyzed the interconnections between tour length and trip purpose, and estimated conditional probabilities of travel to a next stop. They found that tour length decreases with the number of UFTs undertaken by the vehicle in a day, that is, the larger the number of tours the shorter the tour. The average tour lengths of a single truck were 7.24 stops/tour (one tour per day), 4.50 stops/tour (two tours per day), and 2.83 stops/tour (three tours per day). In the case of combination trucks, the numbers were 7.65 stops/tour, 3.65 stops/tour, and 3.33 stops/tour, respectively.

Holguín-Veras et al. (2006) analyzed the UFT patterns exhibited by the regular users of the Port Authority of New York and New Jersey facilities that connect the states of New Jersey and New York across the Hudson River. The analyses revealed there are significant differences in UFT patterns across company types and locations. The data show that common carriers undertake significantly longer tours than private carriers (15.70 stops/tour vs. 7.10 stops/tour), indicating a more intense utilization of assets. Location was found to be a factor as New Jersey carriers had longer tours than those located in New York (13.70 stops/tour vs. 6.00 stops/tour), which may be because carriers located farther away try to serve as many customers as possible to compensate for the longer travel to customer locations.

Figliozzi (2007) studied the generation of vehicle kilometers traveled. He considered the influence of the number of stops per tour, truck capacity, service frequency, tour length and time window length. He found that for equal payloads, multi-stop tours generate more vehicle kilometers than direct deliveries. He also showed that the efficiency of the tours with respect to vehicle kilometers traveled has no correlation with the percentage of empty trips. Finally, the author showed that the average trip length distribution is dependent on tour characteristics.

Wang and Holguín-Veras (2008) estimated discrete choice models to identify systematic patterns of tour construction behavior using synthetic data from Holguín-Veras, Thorson, and Ozbay (2004). The model assumes that the tour construction process could be decomposed into two decisions concerning: the next stop and tour termination. The models revealed that the selection of the next stop has an inverse relation with the distance to the next stop from the current one, and a direct relation with the amount of cargo to be picked up and delivered. The tour termination decision has an inverse relation with distance to the base and a direct one with the total amount of cargo that has been transported.

Figliozzi (2010) analyzed the impact of congestion on commercial vehicle tours and performed a sensitivity analysis on tours characteristics based on real data. The analysis revealed that a reduction of travel speed has a significant effect on vehicle kilometers and hours traveled, where the impact of congestion on them is higher when customers have longer service times. The author found that when congestion increases, the distance per tour decreases. Conversely, the number of tours needed, the percentage of time driving, and the average distance per customer increases.

Ruan, Lin, and Kawamura (2012) defined five different UFT types (i.e., single direct, single peddling, multiple direct, multiple peddling, and mixed) and estimated multinomial models to capture the decision process. The results imply that average trip distance between intermediate stops, average dwell time per stop, and the type of commodity transported are some of the variables that determine the selection of the UFT type.

To provide a fuller picture of UFT patterns, the authors pooled two different data sets in the New York City Metropolitan Area (Holguín-Veras, Silas, & Polimeni, 2008a; Holguín-Veras et al., 2008b; Holguín-Veras et al., 2006) and conducted statistical analyses. The data come from 485 carriers both private (47.26%) and for-hire carriers (52.74%). The results show that 12.60% of the UFT only have one stop, though 8.70% have more than 20 stops. The average is 7.98 stops/tour, though 54.91% make less than six stops per tour. As shown in Figure 17.3, the average

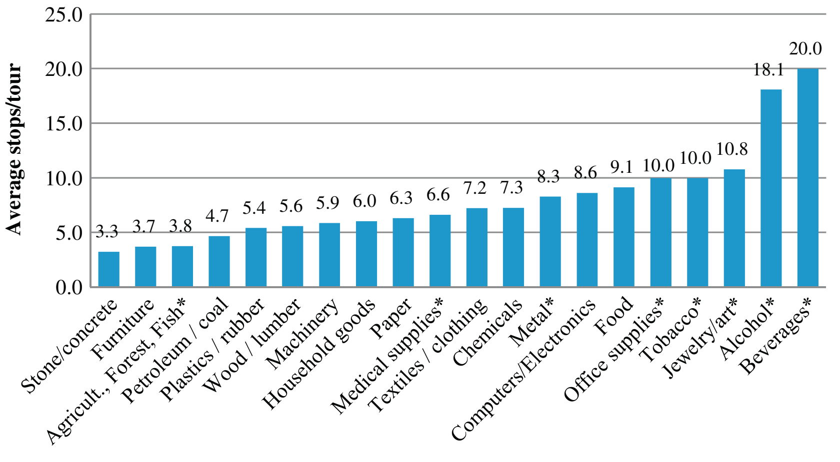

Figure 17.3: Average number of stops per tour (New York City metropolitan area). *Notes: (*) The number of observations is less than 10.*

number of stops/tour depends on the commodity that is transported. The commodities with the longer tours are beverages (20.00 stops/tour) and alcohol (18.09 stops/tour); while stone and concrete (3.25 stops/tour), and furniture (3.71 stops/tour) are the ones with the shortest tours.

17.4. Urban Freight Tour Models

The UFT models described in the literature were subdivided into simulation, hybrid, and analytical models. Simulation models are those that attempt to create the needed isomorphic relationship between model and reality by imitating the observed behaviors in a computer program. Analytical models, on the other hand, attend to achieve isomorphism using formal mathematic representations based on behavioral, economic, or statistical axioms. Hybrid formulations incorporate features of both simulation and analytical models (e.g., using a gravity model to estimate commodity flows, and a simulation model to estimate the UFTs). The models that are described in detail in Section 17.5 (Holguín-Veras, Thorson, & Mitchell, 2012a; Hunt & Stefan, 2007; Stefan, McMillan, & Hunt, 2005; Thorson, 2005; Wang, 2008; Wang & Holguín-Veras, 2009) are not discussed in this section.

17.4.1. Simulation-Based Models

Boerkamps and van Binsbergen (1999) proposed an urban model (GoodTrip) to estimate goods flows, urban freight traffic and its impacts based on logistical chains. The model computes the volume of goods attracted based on consumer demand. The commodity flows are then determined by the spatial distribution of activities and the market shares for each freight-related activity. As part of the simulation, a vehicle loading algorithm converts commodity flows into vehicle flows. To assign the commodity flows to vehicle tours, the authors determine the transport mode, vehicle capacity, maximum loading factor and maximum number of stops per tour based on the origin's activity type; while the minimal delivery frequency is determined based on the type of freight-related activity at the destination. This model was calibrated and validated for the food retail sector in Groningen, in the Netherlands.

FRETURB was developed by the Transportation Economic Laboratory at Lyon (France) to model urban goods flows. The FRETURB is an urban tour-based model composed of four sequential modules (Ambrosini, Routhier, & Toilier, 2004). The first one deals with freight trip generation using statistical techniques to estimate the number of pick-ups and deliveries at each transportation analysis zone (TAZ). The second module estimates road occupancy based on the tours followed by freight vehicles. These tours are simulated based on the distance traveled between stops, the number of stops per tour, the operating mode, and the density of activities for each TAZ. The third module estimates road occupancy by stationary vehicles and the fourth module road occupancy at any instant. The model estimates the impacts of the activity type and location on freight flows fairly well according to the authors.

17.4.2. Hybrid Models

van Duin, Tavasszy, and Taniguchi (2007) analyze the effects of uncertainty in demand on distribution channels. The model considers two types of contracts: long term and spot contracts. Long term contracts between shippers and carriers are repetitive over a period of time where volumes, origins, and destinations are known beforehand. The resulting loads for this type of contract are randomly assigned to a carrier and a truck. The second type of contract deals with the spot market. The paper proposes an auctioning protocol by which services of various carriers can be matched with shipper's demands. A simulation model is used to determine which carrier wins each additional load. The additional costs are estimated using a vehicle routing problem with time windows for each carrier and each additional load. The paper shows that the combination of the loads from the two types of contracts defines the tour followed by each carrier.

Wisetjindawat, Sano, Matsumoto, and Raothanachonkun (2007) estimated a micro-simulation model for urban freight movements incorporating freight agents' behavior interacting in the supply chain. The model considers commodity production and consumption, commodity distribution, and conversion of commodity flows to truck flows. Once the commodity flows are estimated, they are allocated to vehicle tours and modes, and ultimately are assigned to the network. Vehicle routing simulation is used to estimate the sequence of delivery stops. The results from their models show that the model estimates produce results similar to the observed number of truck trips.

Liedtke (2006) and Liedtke (2009) developed the micro-simulation model INTERLOG to simulate transport contracts using an agent-based approach to assess the effects of behavior-oriented transport policy measures. The model is comprised of modules describing company generation, supplier choice, shipment-size choice, carrier choice and tour construction. The results seem to indicate that the model can reproduce the effects of changes in shipment sizes, transport service providers, truck types, tour construction and route choice.

17.4.3. Analytical Models

Holguín-Veras (2000) introduced a modeling framework to estimate freight origin-destination matrices using UFTs. The process establishes the mathematical conditions for freight market equilibrium, under the assumption of a Cournot–Nash equilibrium.

Thorson (2005) developed a solution approach for the partial equilibrium version with freight productions and attractions that are exogenously determined. The solution approach entails a sequential process by which the competing suppliers: (1) deploy their transportation assets as delivery tours to pick-up and deliver cargoes; (2) determine the prices they would charge to each customer; and (3) readjust tours, pick-up and delivery actions, and prices in response to the market competition until price equilibrium is reached. The author used heuristics to find the delivery strategy that maximizes suppliers' profits in a competitive market. The performance of the model was assessed with the assistance of a numerical example inspired by the New York City metropolitan area. The model reproduced key parameters such as empty trips, load factors, and trip length distributions. In the analyses of the resulting market prices, it was found that, as expected, the more transparent the market, the lower the prices.

Xu (2008) and Xu and Holguín-Veras (2008) developed an analytical formulation to estimate UFT flows for the case of integrated shipper-carrier operations that explicitly considers UFTs. The formulation estimates the equilibrium production-routing patterns, and commodity flows that arise when the suppliers attempt to maximize profits in an environment of market competition. The problem is formulated as a Nash equilibrium in which the suppliers adjust production level, profit margin, and delivery routes, in response to competition until optimal strategies are found. The authors found that in small problems, the equilibrium was reached relatively quickly; in contrast, in larger problems equilibrium was not always found. They also found that market proximity is important.

17.5. Illustrative Models

This section discusses illustrative instances of the key model types to provide the reader with an idea about UFT modeling. Although for reasons of space, the reviews cannot be comprehensive, they serve the purpose of providing an introduction to the field.

17.5.1. Simulation Model: Calgary Model

Stefan et al. (2005) and Hunt and Stefan (2007) describes an agent-based micro-simulation framework that uses tours to model commercial vehicle movements in Calgary, Canada. The framework is based on three modules: a tour-based model, a fleet-allocator model and an external-internal flows model. The tour-based micro-simulation is based on a "rubber-banding" approach, in which a primary destination is identified first and then intermediary stops are added. At every stop after the first, a next-stop purpose decision is made until this decision corresponds to the termination of the tour and return to the base. The paper identifies five different industry categories. For each category, a number of tours are generated using regression models. For these tours, a tour purpose and a vehicle choice are assigned based on a Monte Carlo process in which selection probabilities are determined by logit models with land use, accessibility attributes, and establishment location as independent variables. The next component is the next-stop-purpose-model where trip purpose is selected for each stop using Monte Carlo simulation. Two additional components estimate the tour start time and the stop duration model. The model was found to produce robust estimates for Calgary (including service vehicles).

17.5.2. Hybrid Model: The Oregon Model

Donnelly (2007) describes a hybrid microsimulation model to estimate freight flows in Oregon, U.S. The model proposes the use of microsimulation to fuse data from different sources that otherwise would be incompatible. The data sources include businesses databases, the Commodity Flow Survey, the Vehicle Inventory and Use Survey, and truck intercept surveys. The model uses tour-based and activity-based modeling, sample enumeration, and dynamic microsimulation techniques. The process starts with the translation of aggregate production-consumption flows to commodity flows, then it determines trans-shipment points, and discretizes the flows into shipments. The simulated shipments are assigned to shipper-receiver pairs; these shipments are then assigned to a type of carrier (for-hire or private) and a vehicle type for each shipper-receiver pair using Monte Carlo simulations. For each shipper, multiple deliveries are consolidated by truck into daily itineraries, and the itineraries are optimized to reduce total network cost. The resulting itineraries are converted into origin-destination trip matrices and routed using a traditional traffic assignment model. The developers report that the model works fairly well.

17.5.3. Analytical Models

This section discusses two types of models with fundamentally different roots. The first type is represented by the models introduced by in Wang (2008) and Wang and Holguín-Veras (2009), which are aggregate models based on entropy maximization. The second formulation is that of Holguín-Veras, Xu, and Mitchell (2012b) that is based on spatial equilibrium principles and is disaggregated in nature.

17.5.3.1. Entropy maximizing tour flow model The approach used in Wang (2008) and Wang and Holguín-Veras (2009) is inspired by the derivation of gravity models using entropy maximization theory (Wilson, 1969; Wilson, 1970a; Wilson, 1970b). It computes the most likely solution to the system constraints. In this formulation, the problem is decomposed into two processes: a tour choice process, and the estimation of the tour flows that would traverse a given tour (see During calibration, a randomly generated set of tours is provided as an input to the tour flow calibration process. Among the many possible tour flow distribution solutions, the procedure computes the most likely set of tour flows that meet the system's constraints, such as trip generation of each zone, and the total travel cost of the network (Figure 17.4).

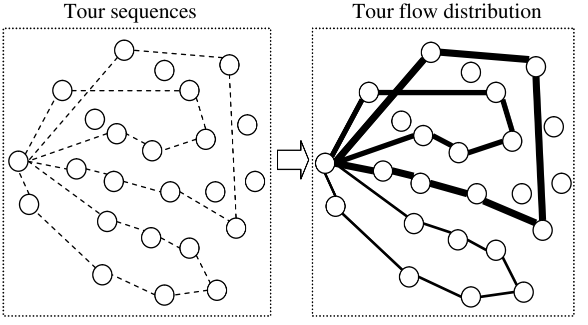

Figure 17.4: Tour sequence and tour flow distribution.

As shown, the former is nothing more than an ordered listing of the nodes visited, while the latter is the total number of vehicles that travels through it (which is represented in the right pane of the figure by the widths of the flows). During calibration, a randomly generated set of tours is provided as an input to the tour flow calibration process. Among the many possible tour flow distribution solutions, the procedure computes the most likely set of tour flows that meet the system's constraints, such as trip generation of each zone, and the total travel cost of the network.

The entropy function used to obtain the tour flow model is:

$$\text{Max}\, Z = \sum_{m=1}^{M} (t_m \ln t_m - t_m) \tag{17.1}$$

Subject to:

$$\sum_{m=1}^{M} a_{im} t_m = O_i, \quad \forall i \in \{1, 2, ..., N\} \quad (\lambda_i) \tag{17.2}$$

$$\sum_{m=1}^{M} c_{Tm} t_m = C_T \qquad (\beta_1) \tag{17.3}$$

$$\sum_{m=1}^{M} c_{Hm} t_m = C_H \qquad (\beta_2) \tag{17.4}$$

$$t_m \geq 0, \quad \forall m \in \{1, 2, \ldots, M\} \tag{17.5}$$

where

t_m = The number of commercial vehicle journeys (tour flows) following tour m;
O_i = Total number of trips produced by node i (trip production);
a_{im} = a binary variable indicating if area i is visited by tour m or not
C_T = Total travel cost in the area
C_H = Total handling cost in the area

c_{Tm} = Tour travel impedance on tour m;
c_{Tm} = Tour handling impedance on tour m;
λ_i = Lagrange multiplier associated with the i^{th} trip production constraint;
β_1, β_2 = Lagrange multipliers associated with the total travel impedance constraint and the total handling impedance constraint, respectively.

The resulting model is the UFT version of a traditional gravity model. As shown in Eq. (17.6), the amount of truck flows distributed to a given tour sequence (t^*_m) is a function of: 1) the marginal effect of trip production or attraction of each area visited along a tour, denoted by λ^*_i; and 2) the tour impedance variables such as tour travel time (c_{Tm}) and tour handling time (c_{Hm}).

$$t^*_m = \exp\left(\sum_{i=1}^{N} \lambda^*_i a_{im} + \beta^*_1 c_{Tm} + \beta^*_2 c_{Hm}\right) \tag{17.6}$$

The model was tested in the Denver metropolitan area (Wang & Holguín-Veras, 2009). The results clearly showed the model's accuracy and efficiency. The estimated tour flows closely match the observed values with the mean absolute percentage error of 6.6%. Equally important is that the parameters of the model and the resulting trip length distributions are conceptually valid. The test results indicate the model's strong potential as the travel demand forecasting model for urban freight, particularly for the situation in which only aggregate demand information is available as the input.

17.5.3.2. A spatial price equilibrium urban freight tour model Holguín-Veras et al. (2012b) develops a general spatial price equilibrium (SPE) model of urban freight demand that explicitly considers delivery tours in integrated shipper-carrier operations. The model estimates the commodity flows and vehicle trips expected to arise under competitive market equilibrium. The model also provides a comprehensive depiction of urban freight activity that considers delivery tours, vehicle trips, and the commodity flows. The authors point out that using SPE has many conceptual advantages accounting for delivery tours: it provides a coherent framework to model the joint formation of commodity flows and vehicle trips (Holguín-Veras, 2000). The consideration of delivery tours provides a closer depiction of delivery patterns that reflect the production-consumption linkages, previously illustrated in Figure 17.2.

The model builds on the seminal work of Samuelson (1952), as it seeks to maximize the economic welfare associated with the consumption and transportation of the cargo, taking into account the formation of UFTs. This is a more general version of the problem solved by Thorson (2005), Xu (2008) and Xu and Holguín-Veras (2008). The optimization problem is:

$$\text{MAX } NSP = \sum_{i=1}^{SN} S_i(E_i) - \sum_{u} C^u_V \qquad \text{(Social Welfare)} \tag{17.7}$$

Subject to:

$$S_i(E_i) = -\int_0^{E_i} s(x)dx \quad \forall i \qquad \text{(Area under excess supply function)} \tag{17.8}$$

$$E_i = \sum_u \sum_j e_{ij}^u \quad \forall i \qquad \text{(Excess supply)} \tag{17.9}$$

$$we^u_{i,p^u_l=j} - Q\vartheta^u_{ij} \le 0 \quad \forall i,j,u \qquad \text{(Linking flows to vehicle-trips)} \tag{17.10}$$

$$\left(t_{i,p_1^u} + \sum_{l=1}^{L^u-1} t_{p_l^u,p_{l+1}^u} + t_{L^u 0} \right) \le T \qquad \text{(Tour length constraint)} \tag{17.11}$$

$$\sum_{i=1}^{L^u-1} \sum_{j=i+1}^{L^u} we^u_{p_i^u,p_j^u} \le Q \quad \forall u \qquad \text{(Capacity constraint)} \tag{17.12}$$

$$\sum_{j,u} \vartheta^u_{ij} \le 1 \quad \forall j \qquad \text{(Conservation of flow)} \tag{17.13}$$

$$\vartheta^u_{ij} \in (0,1) \quad \forall i,j,u \qquad \text{(Integrality)} \tag{17.14}$$

$$e^u_{ij} \ge 0 \quad \forall i,j,u \qquad \text{(Non-negativity)} \tag{17.15}$$

$$c_D, c_T, c_H \ge 0 \quad \text{(Unit delivery cost)} \quad \text{(Non-negativity)} \tag{17.16}$$

$$d_{i,j}, t_{i,j} \ge 0 \qquad \text{(Distance, travel time)(Non-negativity)} \tag{17.17}$$

where

NSP = Net Social Payoff
SN = Set of supply nodes
E_i = Total supply exported by i
$S_i(E_i)$ = Area under the excess supply function s(x)
u = A delivery route (or vehicle)
L^u = Number of delivery stops in u
Q = Vehicle capacity in tons
w = Unit weight of commodity
p_l^u = Ordered index that stores the customer number visited in position l of tour u
C_V^u = Total variable cost
c_D = Unit distance cost ($/mile)
c_T = Unit time cost ($/hour)
c_H = Unit handling cost ($/ton) including loading and unloading
$t_{p_{l-1}^u,p_l^u}$ = Time traversed between stops l-1 and l in delivery tour u
T = Tour length constraint in hours
e_{ij}^u = Amount of cargo shipped from supply node i to demand node j in delivery tour u

ϑ_{ij}^{u} = 1, if vehicle tour departs from supply node i to demand node j (otherwise = 0)
p_{l}^{u} = Node visited in position l of tour u, l = 1, 2, 3 … L^{u}

This formulation is a significant contribution to the spatial price equilibrium literature as it extends this important field so that it explicitly considers delivery tours. Unfortunately, the use of this model presents significant computational challenges as it is both combinatorial and nonlinear. For that reason, its use will have to wait until efficient algorithms are developed for this class of problems. However, the fact that simplified instances have already been solved using heuristics by Thorson (2005), Xu (2008), and Xu and Holguín-Veras (2008) clearly indicates that there is an excellent chance that these economically consistent formulations will be ready for use in real life in the immediate future.

17.6. Concluding Remarks

This chapter presented a comprehensive overview of empirical findings and models dealing with urban freight tours. After providing background information and a literature reviewing pertaining to urban freight tours, the state of the art and practice of UFT models was described. UFT research was divided into two categories: research that characterizes the key features of UFTs and research that involves the development of models to describe and forecast UFTs in cities and metropolitan areas. The first category of research provides the foundation for the development of UFT models by describing real-life features and providing insight into the inter-relations and influencing factors to be incorporated into those models. Three categories of UFT models were identified: simulation, hybrid, and analytical. Simulation models involve the development of computer programs that imitate the observed behaviors of the agents involved in UFTs. Analytical models use formal mathematic representations based on behavioral, economic, or statistical axioms. Hybrid models combine features of both simulation and analytical models. To provide the reader with an introduction to these different types of models, examples for each category were briefly described and an illustrative example from each type was presented in more detail.

The most salient observation from this review is that there are reasons for optimism. This is because, for the first time in history, freight demand modeling researchers and practitioners are cognizant of the need to explicitly model UFTs, and are collecting data and developing models that account for UFT. There are some obvious limitations. The data collected are small and, in most cases, are unable to provide a comprehensive view of UFTs; while the models developed are still in need of significant improvements. Some of the most pragmatic approaches, such as simulations and hybrid models, are likely to require enhancements in their behavioral foundations as they rely on assumptions that are not always validated. On the other hand, the most theoretically appealing models – such as those based on spatial price equilibrium – still need computational improvements to make them ready for real-life use. The authors' expectation is that with continuing progress in the understanding of the behavioral foundations of UFT, and advances on computational algorithms, the various modeling strands will improve. It is also possible that, over time, unification of these distinct modeling approaches would take place. As is always the case, time will tell if this conjecture is correct.

References

Ambrosini, C., Routhier, J. L., & Toilier, F. (2004). *How do urban policies work on the urban goods transport flows?* Istanbul: WCTR.

Boerkamps, J., & van Binsbergen, A. (1999). *Goodtrip - A new approach for modelling and evaluation of urban goods distribution*. City Logistics I, Cairns, Australia.

Donnelly, R. (2007). A hybrid microsimulation model of freight flows. In E. Taniguchi & R. G. Thompson (Eds.), *Proceedings of the 4th international conference on city logistics, July 11–13, 2007* (pp. 235–246). Crete, Greece: Institute for City Logistics.

Figliozzi, M. A. (2007). Analysis of the efficiency of urban commercial vehicle tours: Data collection, methodology, and policy implications. *Transportation Research Part B: Methodological*, *41*(9), 1014–1032. doi: 10.1016/j.trb.2007.04.006

Figliozzi, M. A. (2010). The impacts of congestion on commercial vehicle tour characteristics and costs. *Transportation Research Part E: Logistics and Transportation Review*, *46*(4), 496–506. doi: 10.1016/j.tre.2009.04.005

Holguín-Veras, J. (2000). *A framework for an integrative freight market simulation*. IEEE 3rd annual intelligent transportation systems conference ITSC-2000. Dearborn, MI: IEEE.

Holguín-Veras, J. (2002). Revealed preference analysis of commercial vehicle choice process. *Journal of Transportation Engineering*, *128*(4), 336.

Holguín-Veras, J. (2008). Necessary conditions for off-hour deliveries and the effectiveness of urban freight road pricing and alternative financial policies. *Transportation Research Part A: Policy and Practice*, *42A*(2), 392–413.

Holguín-Veras, J. (2011). Urban delivery industry response to cordon pricing, time-distance pricing, and carrier-receiver policies. *Transportation Research Part A*, *45*, 802–824.

Holguín-Veras, J., & Patil, G. (2005). Observed trip chain behavior of commercial vehicles. *Transportation Research Record*, *1906*, 74–80.

Holguín-Veras, J., & Patil, G. (2007). Integrated origin-destination synthesis model for freight with commodity-based and empty trip models. *Transportation Research Record*, *2008*, 60–66.

Holguín-Veras, J., & Patil, G. (2008). A multicommodity integrated freight origin–destination synthesis model. *Networks and Spatial Economics*, *8*(2), 309–326. Retrieved from http://dx.doi.org/10.1007/s11067-007-9053-4

Holguín-Veras, J., Silas, M. A., Polimeni, J., & Cruz, B. (2007). An investigation on the effectiveness of joint receiver-carrier policies to increase truck traffic in the off-peak hours: Part I: The behaviors of receivers. *Networks and Spatial Economics*, *7*(3), 277–295. doi: 10.1007/311067-006-9002-7

Holguín-Veras, J., Silas, M., & Polimeni, J. (2008a). An investigation on the attitudinal factors determining participation in cooperative multi-carrier delivery systems. In E. Taniguchi & R. Thomson (Eds.), *Innovations in city logistics IV* (pp. 55–68). Hauppauge, NY: Nova Science Publishers.

Holguín-Veras, J., Silas, M. A., Polimeni, J., & Cruz, B. (2008b). An investigation on the effectiveness of joint receiver-carrier policies to increase truck traffic in the off-peak hours: Part II: The behaviors of carriers. *Networks and Spatial Economics*, *8*(4), 327–354. doi: 10.1007/S11067-006-9011-6

Holguín-Veras, J., Thorson, E., & Mitchell, J. (2012a). Spatial price equilibrium model of independent shipper-carrier operations with explicit consideration of trip chains. *Transportation Research Part B* (in review).

Holguín-Veras, J., Thorson, E., & Ozbay, K. (2004). Preliminary results of experimental economics application to urban goods modeling research. *Transportation Research Record: Journal of the Transportation Research Board*, *1873*(2), 9–16. doi: 10.3141/1873-02

Holguín-Veras, J., & Wang, Q. (2011). Behavioral investigation on the factors that determine adoption of an electronic toll collection system: Freight carriers. *Transportation Research Part C*, *19*(4), 593–605.

Holguín-Veras, J., Wang, Q., Xu, N., Ozbay, K., Cetin, M., & Polimeni, J. (2006). Impacts of time of day pricing on the behavior of freight carriers in a congested urban area: Implications to road pricing. *Transportation Research Part A: Policy and Practice*, *40*(9), 744–766.

Holguín-Veras, J., Xu, N., & Mitchell, J. (2012b). A dynamic spatial price equilibrium model of integrated production-transportation operations considering freight tours. (*in review*).

Hunt, J. D., & Stefan, K. J. (2007). Tour-based microsimulation of urban commercial movements. *Transportation Research Part B: Methodological*, *41*(9), 981–1013.

Liedtke, G. (2006). *An actor-based approach to commodity transport modelling*. Ph.D. thesis, Karlsruhe Universität, Karlsruhe, Germany.

Liedtke, G. (2009). Principles of micro-behavior commodity transport modeling. *Transportation Research Part E: Logistics and Transportation Review*, *45*(5), 795–809. doi: 10.1016/j.tre.2008.07.002

McFadden, D., Winston, C., & Boersch-Supan, A. (1986). Joint estimation of freight transportation decisions under non-random sampling. *Discussion Paper*, John F. Kennedy School of Government, Harvard University, Cambridge, MA.

Ruan, M., Lin, J., & Kawamura, K. (2012). Modeling urban commercial vehicle daily tour chaining. *Transportation Research Part E: Logistics and Transportation Review*, *48*(6), 1169–1184. doi: 10.1016/j.tre.2012.06.003

Samuelson, P. A. (1952). Spatial price equilibrium and linear programming. *American Economic Review*, *42*(3), 283–303.

Samuelson, R. D. (1977). *Modeling the freight rate structure*. Cambridge, MA: Center for Transportation Studies, Massachusetts Institute of Technology.

Stefan, K., McMillan, J., & Hunt, J. (2005). Urban commercial vehicle movement model for calgary, Alberta, Canada. *Transportation Research Record: Journal of the Transportation Research Board*, *1921*, 1–10.

Thorson, E. (2005). *The integrative freight market simulation: An application of experimental economics and algorithmic solutions*. Ph.D. thesis, Rensselaer Polytechnic Institute, Troy, NY.

van Duin, J. H. R., Tavasszy, L. A., & Taniguchi, E. (2007). Real time simulation of auctioning and re-scheduling processes in hybrid freight markets. *Transportation Research Part B: Methodological*, *41*(9), 1050–1066. doi: 10.1016/j.trb.2007.04.007

Vleugel, J., & Janic, M. (2004). Route choice and the impact of 'logistic routes'. In E. Taniguchi & R. Thompson (Eds.), *Logistics systems for sustainable cities* (pp. 221–234). Amsterdam: Elsevier.

Wang, Q. (2008). *Tour based urban freight travel demand models*. Ph.D. thesis, Rensselaer Polytechnic Institute, Troy, NY.

Wang, Q., & Holguín-Veras, J. (2008). Investigation of attributes determining trip chaining behavior in hybrid microsimulation urban freight models. *Transportation Research Record: Journal of the Transportation Research Board*, *2066*, 1–8.

Wang, Q., & Holguín-Veras, J. (2009). Tour-based entropy maximization formulations of urban commercial vehicle movements. *2009 Annual Meeting of the Transportation Research Board*. CDROM.

Wilson, A. G. (1969). The use of entropy maximising models in the theory of trip distribution, mode split. *Journal of Transport Economics and Policy*, *3*(1), 108–126.

Wilson, A. G. (1970a). *Entropy in urban and regional modelling*. London: Pion.

Wilson, A. G. (1970b). The use of the concept of entropy in system modelling. *Operational Research Quarterly*, *21*(2), 247–265.

Wisetjindawat, W., Sano, K., Matsumoto, S., & Raothanachonkun, P. (2007). *Micro-simulation model for modeling freight agents interactions in urban freight movement*. 86th Annual Meeting of the Transportation Research Board, Washington DC.

Xu, N. (2008). *A dynamic model of integrated production-transportation operations*. Ph.D. thesis. Rensselaer Polytechnic Institute, Troy, NY.

Xu, N., & Holguín-Veras, J. (2008). *A dynamic model of integrated production-transportation operations*. Annual Meeting of the Transportation Research Record, Washington DC.

Chapter 18

Modeling Behavioral Aspects of Urban Freight Movements

Francesco Russo

Abstract

The chapter presents the main behavioral aspects of urban goods movements and the approach to modeling them. The movements in question are generated by the restocking and purchasing decisions of retailers and end-consumers in the urban context.

The literature on recent shopping trends, such as e-commerce, is analyzed within the consolidated framework of a topological–behavioral model.

In analyzing push-and-pull movements in the delivery phase generated by the end-consumer and retailer, it is shown that the delivery phase can have other decision makers than those in the transaction phase.

A general model system to structure and interconnect all the choices, with relative decisions, is presented in the form of a wide-meshed grid in which researchers and practitioners may introduce specific models presented elsewhere, or calibrated ad hoc in a real urban context. The system considers the four main vehicle movements: push and pull, generated by end-consumers, and retailers. It also considers a commodity model where the quantity choice model is discussed with its relative implications in the main decisional level for each decision maker.

In the chapter, the advancements relative to the alternative set model and the choice within the set are proposed, referring to push and pull movements, giving possibility for future researches, considering implications on new discrete choice model, dynamic choice and presence of ITS, and dispersion of taste.

Freight Transport Modelling

ISBN: 978-1-78190-285-1

Some other implications regard the increase in organized delivery, for both decision makers, that may constitute an opportunity to reduce the number of vehicles with the relative practical and social implications.

The chapter is designed to form a bridge between the transport paradigm and the shopping-restocking process, and supply planners with an integrated updated vision.

Keywords: Urban goods movements; end-consumer and retailer choice; e-commerce; behavioral aspects; city logistics

18.1. Introduction

In 2008, over half of the world's population lived in urban areas, increasing from 25% in 1950. It is estimated that by the year 2050 75% of the world's urban population will be living in towns and cities. If we consider absolute values (rather than just percentages) and the growth rate of the world's population from 1950 to 2050, the figure becomes even more significant.

Population growth in urban areas and their transformation into megalopolises pose formidable problems in all countries. Urban conglomerations are becoming ever more common on a global basis, and there has been a sizeable increase in the number of cities, both those above one million inhabitants and those between 500,000 and 1,000,000. Processes of urban reorganization have often entailed a division of functions, with the separation of residential from commercial areas. The emergence of out-of-town shopping malls has emphasized this separation.

Alongside the growth of cities, the social structure is changing due to longer life expectancy: there has been an increase in the number of residents who do not undertake systematic trips. The overall scenario envisaged for urban area development makes the pursuit of sustainable mobility increasingly necessary. *Sustainability* is intended in the economic, social and environmental sense. The global objective is to ensure that transport systems meet society's economic, social and environmental needs, at the same time minimizing negative repercussions on the economy, society and environment.

Economic sustainability, focusing particularly on efficiency and effectiveness, may be expressed in specific objectives ranging from the reduction of costs incurred by the user and of service production costs, with an increase in production efficiency and product effectiveness, to improvement in the quality of (transport and logistics) services and performance, and to liberalization and regulatory processes that may become an instrument of such sustainability. It is necessary to adopt measures to improve economic and environmental performance for all transportation modes, while in urban areas logistical processes need to be redefined.

Social sustainability is structured into different specific objectives. The most important comprise aspects of safety and security, with problems connected to the various areas of risk (including food).

Environmental sustainability lies at the core of the debate on an international scale. The main issues are those in the *road map* of the Bali Conference and the Cancun Agreement, which decide what route to follow after the Kyoto protocol. Development of a *green economy* is beginning to emerge, from the historical Earth Summit of Rio in 1992 to the UN Conference of Rio in 2012. Many steps have been taken in general awareness, even if political responses are more difficult to obtain. The issues concern: an increase in energy efficiency and ecologically compatible fuel; a reduction in pollution produced by the transport system with particular reference to air quality in urban and metropolitan areas; a reduction in visual intrusion and noise levels in urban and nonurban areas.

Such population concentrations and changes in the social structure, combined with the need for sustainable mobility, heighten problems connected to the purchasing of goods by the end-user and the restocking of retail outlets. Trips for purchases are, in percentage terms, ever closer to those of homework, exceeding in urban areas systematic homework trips. In the United States in 1983 the average annual person trips per household amounted to 537 to/from work and 474 for shopping. By 2009 this had changed to 541 to/from work and 725 for shopping (NHTS, 2009). The traditional structure of restocking of neighborhood shops, and hence of buying consumer goods in such shops, is being increasingly modified by new forms of purchase and distribution, resulting from the Internet. In the United States, e-commerce sales in 2010 increased by 14% against 2009 (Census, 2011). End-users have modified their behavior and purchasing decisions over time, gradually switching from traditional distribution channels.

Hence the ever-increasing importance of knowing not only urban freight movements, but also focusing on behavioral aspects, and on the chance to model them. The general objective of this chapter is to model the behavioral aspects of urban freight movements, mapping the behavior of the end-consumer and the retailer that generates freight movements in an urban context.

Starting from an analysis of existing studies on shopping processes, and the definition of the actors and their relevant actions, insofar as they influence urban freight movements, the chapter has the following specific objectives:

- to categorize the main classes of urban freight movements made by end-consumers and retailers;
- to analyze the main research areas for the different decision makers involved, presenting an integrated modeling system that allows a linkage between final consumer choices and restocking choices made by retailers;
- to present the problems relative to the specification of quantities for movement, and introduce some guidelines that are likely to obtain good advancements in little time.

The chapter is structured as follows: in Section 18.2 we give an overview of general definitions; in Section 18.3 we present the pattern of movements with the behavioral components; in Section 18.4 the main elements for modeling are illustrated.

18.2. General Definitions

In this chapter, we focus on freight trips due to consumer behavior and not on the overall behavior of consumers as regards shopping (for purchases). This initial restriction is required in order to differentiate the contents of our contribution from studies concerning purchasing decisions that constitute a broad field of research for all those dealing with marketing and related sciences.

First let us recall the main aspects of the structure of the shopping process in order to highlight the phases that entail movements of freight and users. Below we only deal with phases that are directly correlated with freight movement. In the sector concerned, much has been written (Peterson, Balasubramanian, & Bronnenberg, 1997; Salomon & Koppelman, 1988) about the decisional sequences in the shopping process. Here we refer to the literature revisited by Mokhtarian (2004), since its breakdown allows the phases connected with freight movement to be identified. The overall phases are: Desire; Information gathering and receiving; Trial/experience; Evaluation; Selection; Transaction; Delivery/possession; Display/use; Return.

In the study by Couclelis (2001, revisited by Mokhtarian, 2004) shopping can be segmented into three phases: before, purchase, and after, each of which must be carried out in the store or by e-commerce, obtaining $2^3 = 8$ possible patterns. The eight combinations obtained define as many types of users: from the traditional shopper who performs the three phases inside the shop (*store*|*store*|*store*), to the innovative consumer or cybernaut who conducts all the phases by Internet (*e-commerce*|*e-commerce*|*e-commerce*), to new types like the careful, well-documented user (*e-commerce*|*store*|*e-commerce*), and to the bargain hunter or free rider (*store*|*e-commerce*|*store*). An even more complex segmentation is proposed by Cao (2012) who considers 16 alternatives to compare e-shopping and store shopping, starting from the basic four indicated by Mokhtarian: Information, Trial, Transaction, and Delivery.

These segmentations are particularly important because they allow consumer behavior to be first fragmented and then reconstructed in respect of the increasingly imposing use of purchase channels connected with the Internet. They also highlight the current and potential impact of the Internet upon all the phases of the shopping process.

The above segmentations, and the others found in the literature, do not allow direct specification of trips that occur in relation to the consumption of goods. They may be used as a starting point, but it is necessary to set further initial definitions and introduce specific hypotheses on the patterns of movements and the behavior that generates them.

Initial definitions concern:

- spatial definitions, which restrict the analyst's field of attention;
- definition of end-consumer, which identifies the individual making the choices;
- definition of freight type, which identifies what is chosen.

By their very nature, the three definitions are interwoven and a strict hierarchy cannot be specified. Thus the necessary presentation in series should be flanked by a parallel interpretation.

The structure of the overall model referred to in this chapter follows the topological behavioral paradigm established since the 1990s. In terms of topological aspects, the main point of reference is Beckman, McGuire, and Winsten (1956) while behavioral aspects derive from the works of McFadden (1974), and Ben-Akiva and Lerman (1985). The general theory of transport systems obtained is the one which represents real systems in abstract form (Cascetta, 2007; Manheim, 1979; Sheffi, 1985).

18.2.1. Spatial Definitions

First of all two places (or rather, two meta-places) have to be defined:

the place in which the generic good is consumed;
the place in which the generic good is purchased.

The term *meta-place* is useful since the physical sites of final consumption and purchase for final consumption may be very disparate, and are continually evolving alongside changes in the user's daily behavior.

Below we refer to the main-albeit not exhaustive-behaviors that unfold in a reference urban area. An attempt is made to identify those behavioral aspects that, though not prevalent, could nonetheless become important and significant with the growth of the use of the Internet for purchases; that is, with the increase in de-localization. What emerges is that the components in user behavior change in relation to the type of urban area defined both on the strictly physical plane and in economic and social terms, apart from in relation to the end-user's own characteristics.

For the specification of the basic topological structures, reference should be made to Figure 18.1. The urban area, containing the transport system to be analyzed, may be expressed by means of a graph $G = (N, L)$, where N is the set of internal and external nodes and L the set of directly linked pairs of nodes belonging to N, called links.

It is useful to divide the geographical area into traffic zones and approximate all points at the beginning and end of trips in each zone with one point (centroid). The internal centroid set is $\{c\} = C \subseteq N$; the external centroid set is $\{z\} = Z \subset N$ and $\{c\} \cap \{z\} = \emptyset$.

Among internal and external centroids four different sets are defined:

$$\{o\}, \{d\}, \{w\}, \{z\}$$

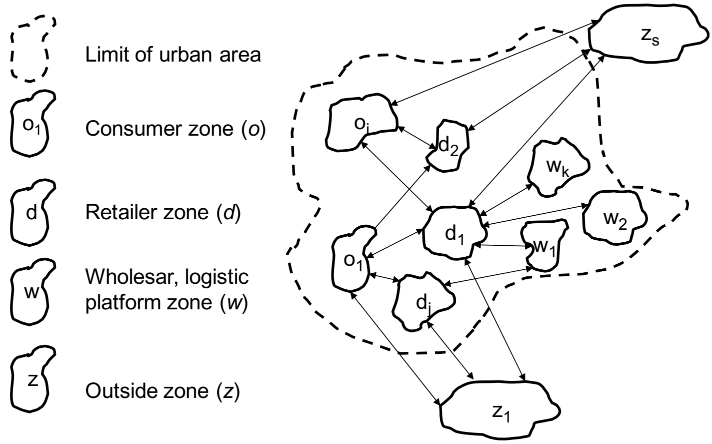

Figure 18.1: Elements for spatial definitions.

where
$\{o\} = O \subseteq C$ is the set of *internal* zone centroids in which the end-consumer consumes (uses) the goods and the residences and services (offices) are located; in these zones the goods are consumed, or their use starts;
$\{d\} = D \subseteq C$ is the set of *internal* zone centroids in which the end-consumer can purchase the goods that the retailer sells; the shops are located in these zones;
$\{w\} = W \subseteq C$ is the set of *internal* zone centroids to which the retailer can bring goods sold in his/her shops;
$\{z\} = Z$ is the set of zone centroids *outside* the study area where the end-consumer can purchase the goods that an out-retailer (or other) sells, or the (internal) retailer can bring the complementary goods sold in his/her shops.

Many zones are thus at the same time type o and type d; others may be type d and type w. Hence in general:

$$\{o\} \cap \{d\} \cap \{w\} \neq \emptyset$$

Note that with the set definitions, no constraint was introduced concerning the trip from o to d: in particular, the decision maker for the trip was not defined. Likewise, no constraint was introduced for trips from o to z, and for those from d to w and from d to z.

In many urban contexts the presence of large retailers (Great Organized Distribution) in central areas offers end-users the possibility of gaining access to sales outlets traditionally available only to small retailers. In this sense all the zones with outlets accessible to end-consumers are indicated by d. The areas with specific warehouses of retailers or producers for wholesalers, or with logistics platforms of shippers to supply retailers are indicated by w. This is in line with United Nations

Statistics, according to which "wholesale" is the release (sale without transformation) of freight to retailers, industrial, commercial, institutional, or other professional business users, or to other wholesalers and related subordinated services; in general, it is the sale to anyone other than a final consumer. Everything sold outside the study area is indicated by z, whether by a retailer, wholesaler or a logistic platform, or by a producer.

18.2.2. Definition of Final Consumer in Relation to Transaction and Delivery Phases

The definition of end-consumer is not straightforward, and often in the analyses in the literature, and in the available models, it is somewhat ambiguous. The ambiguity arises from the angle with which the generic consumer is observed. Hence, for example, a retailer who has to change the roll in the cash till to issue receipts is, in terms of the paper roll, an end-consumer, while for his/her main activity he is not an end-consumer, but a retailer.

This is why it is useful to define two types of end-consumers. A first type consists of family groupings in the broad sense, of which the family or the single person is the most widespread type, which represents the end-users of goods across the board. However, also in this case there might be forms of ambiguity connected to the spread of work via the web, which might be carried out at home, thereby introducing promiscuity in the use of specific goods. A second type is that consisting of producers/sellers of services and goods, who purchase goods not earmarked for sale, or who contribute to the production and sale of services and goods only indirectly related to that activity. Consider, for example, the consumption of tea or coffee in a bank office. Applying the general formulation of the i/o model, this is the purchase of goods that would fall within the functional form of the good produced, in the sense that their values are much lower than the price of services or goods on sale by producers/sellers considered.

Such specifications may resolve ambiguities, especially for the second type. The first type may be easily indicated as consumers and they are linked to transactions of type B2C (business to consumer). In the second type, the subject is a business but, as stated above, the purchase of the good in question does not increase the firm's core business. This is why, also in this case, we are dealing with a transaction similar to B2C, rather than B2B (business to business). Coming under the second type are supplies to public authorities, which are usually considered B2G (business to government). Note that the abbreviations B2C, B2B, and B2G were introduced into the literature for electronic transactions related to e-commerce, while they may also refer, without introducing any ambiguity, to traditional transactions.

For each different kind of *transaction* or *purchase* (Couclelis, 2001; Mokhtarian, 2004), there are various movements of goods possible:

- the end-consumer buys from a retailer or from a wholesaler or from a producer within the study area ($d \in D$);

- the end-consumer buys directly from the producer or from a retailer/wholesaler in a zone outside the study area ($z \in Z$)
- the retailer buys from a wholesaler or a producer within the study area ($w \in W$);
- the retailer buys directly from the producer or from a wholesaler in a zone outside the study area ($z \in Z$).

With this definition, and recalling the *delivery* phase, we can specify two main classes of urban freight movements:

- the consumer trip, to which the first and second of the above-defined freight movements belong, in which the transaction decision maker is the consumer;
- the logistic trip, which comprises the third and fourth of the above-defined freight movements, in which the transaction decision maker is the retailer.

Our attention in this chapter chiefly concerns freight movements connected to the end-consumer. The movements connected to retailers are only recalled for similarity's sake where similar hypotheses may be extended. Hypotheses of similar behavioral structures are valid when the retailer has a traditional outlet, not belonging to any national or international chain, nor to franchise groups.

Identifying the decision makers for the transactions, hence for the movements resulting from the transactions, does not yet allow a complete specification to be obtained. Identification of decision makers for the next step in the shopping process, *delivery*, is necessary because the freight movement is a trip chain (from producers to consumers) and in each trip there may be various decision makers. This holds both for consumers and for retailers. Indeed, consumers may introduce the delivery phase within their daily activities or may have the goods delivered in the reference zone o. By the same token, retailers may assume different behavior whether they transport the good themselves, or entrust delivery to others.

18.2.3. Definition of Freight Type

To define the subject of the trip, we must refer to the generic consumption good indicated by i. It may be generally assumed that i belongs to the class of goods identified by I ($i \in I$). In the literature various classifications are available in relation to the use to be made of the classes. Classifications that more closely address the economic and financial components of the freight of the NACE type (Nomenclature générale des activités dans les Communautés Européennes is the General Industrial Classification of Economic Activities in the European Communities) are used in national censuses while classifications more oriented to the vehicle type required for transportation, of the NST type (standard goods Nomenclature for Transport Statistics, used by Commission of the European Communities), are used in statistics relative to transport infrastructures, roads, railways, ports and airports.

For the treatment of urban freight movements, it is worth using a more aggregate classification than those introduced previously, which nevertheless have the benefit of being able to be obtained, through aggregation, from each of the two classifications described above, retaining the minimum complexity useful for specifying behavior. The first level of classification is split into durable and nondurable goods, in relation to their life-time for the purposes of the consumer in question. The first class contains goods that are continuously and repeatedly used by consumers, ranging from clothing to furnishings and household appliances.

In the context of nondurables, two classes may be identified, namely food and nonfood. Belonging to the class of nondurables and nonfood are perishable goods like daily newspapers and weekly magazines. The class of food may in turn be broken down into two subclasses: perishables and nonperishables. The first subclass includes commodities like bread, fresh milk, fruit and vegetables, meat, and fish. The second subclass comprises beer, wine, salt, sugar, pasta, UHT milk, etc.

The above definitions together with the segmentation concerned (Mokhtarian, 2004) are a necessary prelude to identifying the patterns of movements and the behaviors that generate them. The combination of actors that decide on the transaction and the subject of the transaction with specific class characteristics allows us to introduce aggregative hypotheses on the trip types that constitute the basis for formulating the models. At the same time, the combination of actor and subject does not allow a unique definition of decisions with regard to the delivery phase.

18.3. Commodity-Purchase Trips

On the basis of the definitions in the previous section we may introduce some hypotheses that allow description of behavior in delivery, both in the case of the consumer and the retailer. To differentiate the two purchasing processes, *attraction* is used for the process of the end-consumer, and *acquisition* for the retailer process, by referring to the definitions in the literature (Russo & Comi, 2010).

Identifying the decision maker of delivery is crucial; since it is used to determine the whole decision-making process that, in the final analysis, allows network flows to be simulated on the basis of the above-described topological and behavioral model. The problem that is posed lies in defining who decides how the freight has to move between the two extremes (the two meta-places). At this stage of the analysis "how" concerns the aggregate channel through which the freight may arrive. The decision concerns the alternatives that respond to the basic demand: are the purchased goods transported by the purchase decision maker or not? These alternatives postulate both that the transaction was carried out by the end-consumer (who purchases in the shopping process), and that it was performed by the retailer (who purchases in the restocking process). To this end, it is worth separating the respective processes. Only in later specifications, and on the very basis of identifying the decision maker of the delivery, will it be possible to analyze, besides the location where the good is purchased, the mode, service, path, schedule, and so forth.

18.3.1. End-Consumer, Decision Maker in the Transaction

The type of behavior in purchasing products induces movement of the products from the generic point *d* (or *z*) to the place of consumption *o*. It is hypothesized that the decision maker decides, as he/she is at *o*, where to purchase the commodity, and how to make it reach *o*. Alternatively, the decision maker decides where to purchase the good and how to make it reach *o*, without him/her being in *o* at the moment of the decision.

Although there are many alternatives, the macrobehavior may be summarized into two classes (Figure 18.2):

- pull-type behavior: the end-consumer arrives at place *d*, performs the transaction and purchases the commodity; the user transports the good to the consumption site *o*. Both in going from *o* to *d* and from *d* to *o*, the user may make other stops.
- push-type behavior: the end-consumer may or may not go to place *d*, perform the transaction and purchase the good; the commodity is transported to site *o* of consumption by actors other than the user.

Traditional pull-type behavior chiefly prevailed until the arrival of the Internet, or at any rate until the first forms of e-commerce (Golob & Regan, 2001). Such behavior may be broken down into different classes in relation to:

- spatial path: choice of the purchase place *d*, and trip from *o* to *d* (1) by loop or (2) chain;
- temporal frequency: systematic trip during the week, or nonsystematic (identical at the weekend or different);
- bundle of goods: trip for purchase of a single good q_{ii}, or for purchase of different goods q_{i1}, q_{i2}, q_{i3}; and in this case: (1) the different goods are all of class *I* or (2) there are also goods of class *J* of type q_{j1}, q_{j2}.

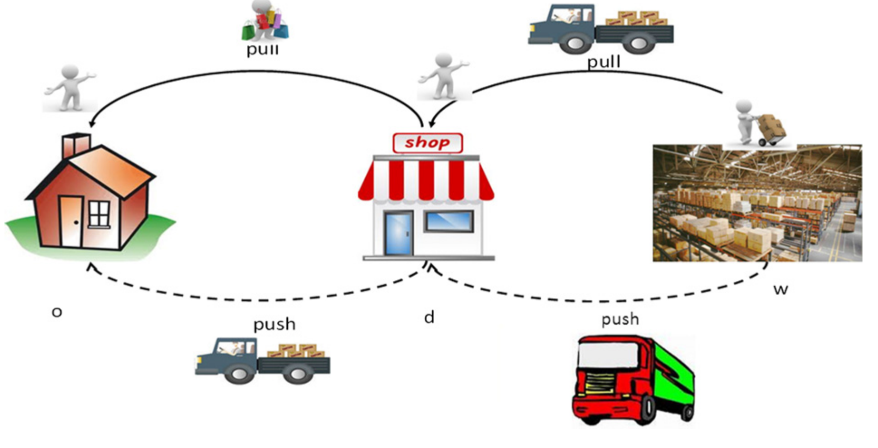

Figure 18.2: Aggregated alternatives for the end-consumer and retailer.

So-called push behavior and movements have a different history: their first forms were door-to-door selling or the traveling sales of goods, especially fruit and vegetables, or the classic vacuum cleaner. Door-to-door has not only retained its market share, it has increased it. Today, push freight movements include all the trips generated by e-commerce also due to the vast use of call centers in the first phase of contact with the potential customer. In this case, except for specific situations in which trial phases are particularly important (consider the case of the purchase of second-hand cars), movements are of a push type.

Push behavior may be broken down into different classes in relation to:

- spatial path: choice or otherwise of the purchase place *d*; alternatives available: (1) decision on the part of the user of product q_{ii} and purchase place *d*, with web shopping or telephone purchase; (2) decision on part of the user of product q_{ii}, but not of the place of origin *d*, for web purchase; (3) direct sale in place o of product q_{ii}, with traditional forms of sale;
- temporal frequency: systematic purchase during the week (fresh milk, news-papers,..), or nonsystematic;
- bundle of goods: purchase of the individual good q_{ii}, or of different goods q_{i1}, q_{i2}, q_{i3}; and in this case: (1) the different goods are all of class *I* or (2) there are also class *J* goods of type q_{j1}, q_{j2}.

18.3.2. Retailer, Decision Maker in the Transaction

The behavioral mode in product acquisition leads to products being transported from the generic point *w* (or *z*) to the sales outlet *d*. The general hypothesis is that the decision maker decides, when in *d*, where to purchase the good, and how to make it arrive at *d*. A complementary hypothesis is that the decision maker decides where to purchase the goods and how to make it arrive at *d*, without being at *o* at the moment of the decision. Also in this case, as previously for the end-consumer, although there are many alternatives, the macrobehavior may be summarized in two classes (Russo & Comi, 2010) (Figure 18.2):

- pull behavior: the retailer goes to place *w*, or *z*, purchases (acquires) the goods; the retailer transports the goods to the retail outlet *d*; along the path the retailer may undertake other stops.
- push behavior: the retailer may or may not go to place *w*, or *z*, purchases (acquires) the goods; the goods are transported to sales outlet *d* by actors other than the retailer.

Pull and push behavior is variously interwoven, and both may be found in the same decision maker in the restocking phase. Hence in the same shop and with the same decision maker there may be class *I* goods that are purchased with pull behavior and hence movement, and class *J* goods that are purchased with push movements.

The prevalence of one type over another is historically connected with the organization of national logistic systems. Hence in countries such as Italy where ex-factory sales prevail there is a greater quantity of *pull* trips. In countries such as the United Kingdom where the producer pays for delivery there is a strong *push* component. As in the case of the end-consumer, also for the retailer the spread of the Internet has further increased *push* movements. By contrast, unlike what has occurred for the end-consumer, for perishable food commodities, with particular reference to fruit and vegetables, pull-type trips are very widespread; push trips are developing in relation to the availability of new forms of packaging suited to such goods.

Pull behavior may, in turn, be broken down in relation to:

- spatial path: choice of the purchase place w (z), and loop trip or chain trip;
- temporal frequency: systematic trip during the week, or nonsystematic.
- bundle of goods: trip to purchase the individual good q_{ii}, or for different goods q_{i1}, q_{i2}, q_{i3}; and in this case: (1) the different goods are all class I or (2) there are also class J goods of type q_{j1}, q_{j2}.

Push behavior may be broken down into different classes in relation to:

- spatial path: choice or otherwise of the purchase place $w(z)$, available alternatives: (1) decision by the retailer of product q_{ii} and place of purchase; (2) decision by the retailer of product q_{ii}, but not of the place of purchase; (3) attempted sale by the producer at the retailer's in place d of product q_{ii}, or sale by target in franchising, etc.;
- temporal frequency: systematic purchase during the week (milk and dairy products, newspapers,..), or nonsystematic;
- bundle of goods: purchase of the individual good q_{ii}, or purchases of different goods q_{i1}, q_{i2}, q_{i3}; and in this case: (1) the different goods are all of class I or (2) there are also class J goods of type q_{j1}, q_{j2}.

Note that, given the same transported goods, with push-type trips the number of vehicles used is considerably reduced, even if the utility of the user making the transaction (in purchase) cannot be assumed to increase. The increase in consignments organized by the same subject (if working in a push way) affords economies of scale, such as better use of vehicles, of time windows, loading capacity, inventory, drivers and so on. At the same time, not all users will necessarily increase their utility: for instance, the shop that has the goods may be the last in the vehicle tour. We can thus consider push movements as following the normative (system optimum) approach and pull movements as following the descriptive (user optimum) approach, as first stated by Wardrop (1952) in his seminal work.

It is thus shown that, in shifting from pull trips to push trips the overall efficiency of the system may improve on average, especially on the urban scale. Knowledge of push or pull types in delivery, at all levels, is decisive for any policy that is intended to be implemented on an urban scale.

In Figure 18.2 the aggregated alternatives (push and pull) in delivery are represented. The arrows indicate the direction of goods, while the line type refers to the decision maker of the delivery.

18.4. Models for Trip Simulation

This section contains the main references for modeling the behavior concerning the trips in question. It should be noted that the paradigm offering a general framework is topological–behavioral.

The different trips presented in disaggregate form in the previous section could be simulated with only one general model containing the alternatives consisting of different combinations. The difficulty both in analytical specification and in subsequent calibration and then in validation makes such an avenue impracticable. It is thus necessary to make a segmentation in simulating decisions, even when the latter are taken by the decision maker him/herself. In segmenting the decisions, on the part of the analyst, specific interpretations are introduced that must be tested and have to reach minimum thresholds of statistical acceptability.

In this section, we introduce some of the main macroelements present in the literature. To do so, given that much has been written on global taxonomy (see the recent papers by Chow, Yang, & Regan, 2010 and De Jong, Vierth, Tavasszy, & Ben-Akiva, 2012), we use a structured framework. In this framework we propose the different elements not at the same level of importance, but in a tree, in which the macroelements are selected and constitute the main branch. For each macroelement specific twigs are presented, and so on to the final leaves. For each macroelement some reference papers are considered, but the challenge is to give to the reader-analyst just a grid, an outline from specification to calibration, in which models that differ extensively from that proposed may be introduced. With this approach the focus is on the problems and the literature is only a tool, where only one paper is cited for each leaf in question.

Drawing up an overall system to frame and interconnect all the decisions is thus fundamental. The first macroelement concerns the overall system that allows the use of many of the findings from the literature, obtained by different researchers, especially in recent years in which attention to such trips has increased. In this way, to introduce multistep modeling we consider one reference work (Russo & Comi, 2010), but there are other works that can be used. Only the main elements are introduced, deferring analytical formulation to the original papers recalled.

The second macroelement that has to be investigated involves determining the quantity of freight transported in each shipment. This is a classic freight transport problem, which fundamentally differentiates freight transport from passenger transport, together with that treated above of the decision maker who in passenger transport (almost) always coincides with the transported individual, while in freight transport the decision maker is always (clearly) different.

The third macroelement concerns the modeling structure to be used in specific choice models at individual levels, considering recent advances made in urban freight models.

In this proposed structural framework other macroelements may be introduced. Two are specifically omitted: a branch "before" and a branch "after."

The *branch before* refers to the relationship between the hypothesized pattern, with the mathematical formulation, and actual city life. This relation is represented by the data collection with respect to the various decisional levels for each decision maker. For each choice presented and for each type of model proposed, the specific difficulties that arise must be investigated. The absolute privacy concerning information on purchasing habits – user profiles, propensity to purchase certain goods, decisions on past purchases – prevents their use in constructing models. This characteristic means that specific research is required in the potential offered by Information and Communications Technology (ICT) in different forms to construct suitable data sets that don't open the privacy: from intelligent transportation systems (ITS)-based information to that based on electronic transactions, that based on relations among the various goods that can be obtained from new Internet technologies (Internet of thinks). Major studies are welcome in this new field to solve the data problem. A general approach to the use of traffic counts, time measurements and other data to calibrate demand models is presented in Russo and Vitetta (2011).

The *branch after* involves the policy implications of urban freight movements. This other macroelement can be developed by first considering the time horizon of reference (operative, tactical or strategic), and then, for each of them, the different policy implications for decision makers in different positions. For each policy the different "what to" solutions that can be supplied must be investigated, in terms of specific measures and expected results.

In this very flourishing branch we consider the behavioral aspects of urban freight movements that require tactical decisions. In this branch a large number of papers have been proposed (Rizer, Cornelis, Browne, & Leonardi, 2010), including that involving the author (Russo & Comi, 2011a, 2011b) that includes the main research strands and a synopsis of the literature.

Interaction with urban development involves different players other than the end-user and retailer. The formal main player is the mayor as an official representative of decision making regarding urban strategy and measures for society. However, external conditions have to be considered, such as the development of economic interaction in terms of complexity and new ITC. The relationship between ITC and transport at the urban level is intensifying on a daily basis and the consolidated pattern is changing.

In this chapter following the path cleared by Mokhtarian (2004), Circella and Mokhtarian (2010) and Mokhtarian and Tang (2013) – we underline the new role of push movements with the possibility in a very near future to endogenize the main part, as proposed here in approximate terms. For a general overview on advanced operative and strategic aspects, refer to the updated literature presented in Ben-Akiva, Meersman, and Van de Voorde (2013).

The circular relationship between after (the mayor's policy and measures) and before (user choices) inserted in the dynamic evolution of the urban context makes each city specific, along with the characteristics of urban freight movements. The urban context defined in terms of social, economic and environmental elements modifies the nature of the behavioral model from the relative weight of the different acquisition and attraction models to the parameters of retailer and end-user path choice.

Moving from a high quality and innovative city such as Seattle, to a high-quality and historical location such Florence, to a new town such as Milton Keynes in the United Kingdom, specific modeling can differ according to different quantities of infrastructures, private vehicles, congestion, transit, metro and so on. In the following sections, we discuss the three macroelements always present in modeling urban movement.

The direct relationship between urban context and transportation is developed in works that treat strategic aspects involving land use (see, e.g., Russo & Musolino, 2012). Other issues regard matters concerning the behavioral aspects of freight movements over distances greater than those in urban networks, namely from the metropolitan to the regional. As stated above, we again refer to the specific book by Ben-Akiva et al. (2013) for a review of the recent literature.

18.4.1. The General System of Models

The various behavioral aspects relative to urban freight trips have been extensively covered elsewhere. Using a highly reductive scheme, it is possible to refer to two large areas of analysis. The first considers the user, hence the end-consumer, as the main point of reference, with his/her variously articulated behavioral choices. The reference work is that of Ben-Akiva and Lerman (1985), which presents the formulations of the theories of individual choice behavior, hence multinomial choice, and the system of models within which are specific models of shop trip frequency. For the second area the firm is the chief point of reference: it has to minimize costs from the production to the consumption of its own products. In this respect, importance should be attached to all the studies that fall within the discipline traditionally and generically termed logistics, in the field of operations research. In this framework, mention should be made of the work by Daganzo (1991). The monograph represents an attempt to examine logistics systems in an integrated way. After all, besides transportation, a logistic system usually includes other activities such as inventory control, handling, and sorting, which must be carefully coordinated if cost effectiveness is to be achieved.

Although much has been written about the two areas in question, there have been few studies treating the global problem of urban freight transport simulation. Existing models mainly simulate some aspects of the consumer in the choice process, and others of the retailer in the restocking process. The latter focus on movements between firms (producers) and distribution centers on a wide scale. In particular, they do not consider the possibility of combining freight and passenger flows, hence of representing the interacting behavior of commodity consumers and commodity suppliers/shippers/retailers.

In the context of works on urban freight transport, we should mention the seminal paper by Ogden (1992). Problems of city logistics, with the possibility of connecting the two research areas described above, have been widely investigated in the last ten years. Differing approaches have been adopted to the various problems of urban freight transport. These approaches differ in: modeling structure; reference unit for transport; distribution channels; aggregation level; basic assumption; calibration and published results.

Of the main models for urban freight movements, mention should be made of some classes of problems: ITS and city logistics (Taniguchi, Thompson, Yamada, & Van Duin, 2001); retailer push and pull vehicle routing (Holguín Veras, 2002; Polimeni, Russo, & Vitetta, 2010); freight transportation and congestion (Crainic, Ricciardi, & Storchi, 2004); peak-hour movements (Munuzuri, Cortes, Onieva, & Guadix, 2010); and end-consumer movements (Gonzales-Feliu, Toilier, & Routhier, 2010). A specific group of models is the one that considers the relationship between e-shopping and store shopping. This includes already cited seminal works (Golob & Regan, 2001; Mokhtarian, 2004) but also Farag, Schwanen, Dijst, and Faber (2007) and Cao (2012) who report complete calibrations.

A modeling system to simulate goods movements at an urban scale may be found in the work by Russo and Comi (2010). The modeling system is developed on two levels and concerns a medium-size city, applying a disaggregated approach for each decisional level:

- *commodity level* (first level) concerning estimation of quantity flows; at this level the models concern calculation of demand by freight type, by *o*-*d* consumption pair (attraction macromodel), and *d*-*w* or *d*-*z* restocking pair (acquisition macromodel);
- *vehicle level* (second level) that allows quantity flows to be converted to vehicle flows; at this level the models concern determination of the service, vehicle and time used (service macromodel) as well as the path chosen for restocking sales outlets (path macromodel).

The structure of the models in question (Russo & Comi, 2010) may be updated or developed with other models from the literature, referring to the elements underlined in this chapter. First, the structure may be developed by explicitly introducing the vehicle level component for the end-consumer. In this way the second level concerns not only the vehicles used in the restocking process, but also those used in shopping processes. For vehicles involved in shopping processes, the macromodels are those of mode and path choice with assignment. The complete structure could endogenize the push/pull dichotomy.

Given the behavior, and the relative movements, presented in Section 18.3, it is necessary to model push and pull behavior both for shopping and for restocking. We thus reach the general framework of the model system reported in Figure 18.3, in which it is simple to identify the individual macromodels concerned. In the general model, it is worth noting that the retailer intervenes twice, at the vehicle level, with extremely different characteristics: first when he/she operates as a decision maker in restocking processes with pull behavior; a second time as a decision maker in shopping processes with the push behavior of the end-consumer. Clearly, on a

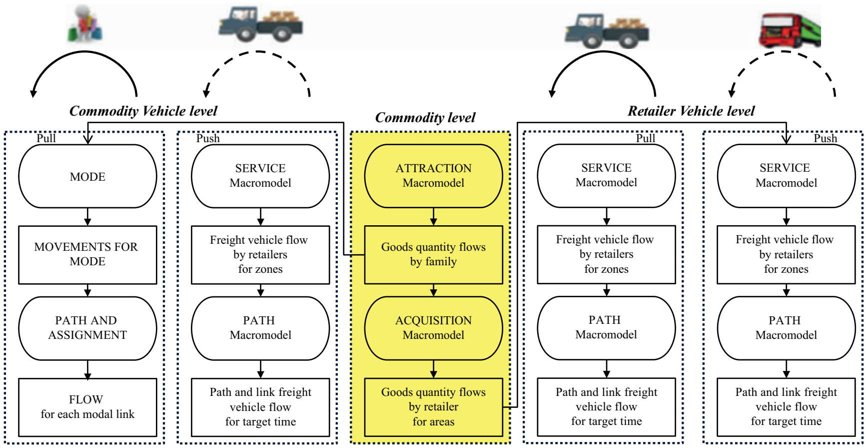

Figure 18.3: General modelling system structure.

theoretical basis, the retailer's two chains could in turn overlap on the final path on the network.

The copious literature in question can contribute to specify the different proposed macromodels, to fill in the grid and make it implementable in its complete form. In the same way, a single block (i.e., a macroelement) may be used to solve a specific freight transport problem at the urban scale. The key question that permeates all of the proposed models is the relationship between ITS and transportation that must be investigated and introduced in each of them.

18.4.2. Quantity Choice Model

One of the main problems in freight models, of whatever area scale, is to determine the quantities that affect behavior. Here we wish only to point out that much of the formulation stems from determination of the quantities of reference. This quantity, in each model, and the relevant problem of measurement is one of the major barriers to understanding freight mobility phenomena. For this problem mention should be made of the work by Bayliss (1988), which deals precisely with the measurement of supply and demand in freight transport; the problem of measurement was posed by Ben-Akiva and Lerman (1985) as one of the main directions for future research.

In simplified form it can be assumed that the reference quantities for the end-consumer may be:

- q_i the quantity of goods that the end-consumer wishes to purchase in his/her generic shopping;
- Q_i the overall quantity of goods of type i required for a reference period T by the family group, or the final business user; q_i being $<=Q_i$.

Similar assumptions in terms of q_i and Q_i may be made for the retailer, and for all the other decision makers involved in the logistics chain from the producer to the end-consumer. As regards the reference quantity for the behavioral choice model, two approaches emerge from the literature: the first, called the consignment approach, considers the quantity per single trip q_i as the fixed reference for the choice; the second, called logistic, considers the global quantity Q_i, for the reference period T, as the fixed reference and the quantity q_i as a dependent variable chosen by the decision maker. Within the logistic approach it is possible to recall the preliminary models set up at the MIT (Modenese Vieira, 1992) up to the advanced model on Aggregate-Disaggregate-Aggregate freight model system (Ben-Akiva & de Jong, 2008). Within the consignment approach a large body of literature has been developed. The cited works mainly deal with movements at national levels. For specific references, see Russo (2005).

The considerations made in this section should be interpreted in light of what was stated in Section 18.4.1 above, hence bearing in mind the need that (consignment or logistic) models may participate in an overall system of simulation. Further, specifications on quantities must be consistent with the definitions in Section 18.2 and with the hypotheses on behavior introduced in Section 18.3. In this sense the quantities of pull and push trips, and the quantities for an individual class of goods or for a bundle of goods of different classes, change in relation to the characteristics of the urban area of reference (spatial definition), to user characteristics (final consumer definition), and in relation to freight characteristics (definition of freight).

18.4.3. The General Definition of the Model: Alternative Set Model and Choice within the Set

In the context of formulating the generic choice model, belonging to the overall system, attention should be paid to the problem of defining the set of alternatives, and to the choice within the set, in terms of modeling structures to be used. The alternative set is always assumed as given, yet this is not so, starting from the problems connected with the specification of the path choice model.

In the overall structure introduced above, the problem is clearly posed both for the vehicle level referring to the end-consumer and for the same level referring to the retailer. In both cases the problem is posed both for *push* movements and for those of a *pull* type, with four different models having to be specified. The problem of the alternative set is also posed in specific parts of the attraction macromodel and the acquisition macromodel.

Defining the set of alternatives in problems related to shopping is especially important since the hypothesis that the user considers all the path alternatives loopless with a fixed choice set appears somewhat forced. The hypothesis that the generic user considers that only paths without loops can be accepted, in the sense that the user does not consider paths along which an intersection is crossed two or more times. In the literature a fixed choice set is assumed and all loopless routes are considered with the same probability of being perceived, since high-performance algorithms have been developed to calculate all loopless routes for each pair of centroids. However, in order to take account of the perception process that influences user route choice behavior, considerable researcher effort has been spent in recent years. Choice set generation is not so easy due to the large number of alternatives existing between each origin-destination pair and is a critical phase to investigate the perception process that cannot be skipped, assuming that the same choice set comprises all topological alternatives.

Manski (1977) defines the choice probability $p_n(k)$ of a generic alternative k for user n, considering two models: generation of the choice set and choice of the alternatives. Given the set F_n of all the feasible alternatives that, for route choice, are all the loopless routes for user n, the probability $p_n(k/S_n \neq \emptyset)$ can be calculated as

$$p_n(k/S_n \neq \emptyset) = \sum\nolimits_{\forall I_n \in S_n} p_n(I_n/S_n \neq \emptyset) p_n(k/I_n)$$

where

- S_n is the set containing all the subsets that can be extracted from F_n, considered by the users;
- I_n is the generic subset perceived by user n (I_n belongs to S_n);
- $p_n(I_n/S_n \neq \emptyset)$ is the probability of user n choosing subset I_n conditional that S_n is not empty;
- $p_n(k/I_n)$ is the probability of user n choosing alternative k conditional on choosing subset I_n.

For explicit route choice modelling, starting from this theory and the consideration that only a subset of all the possible routes (choice set) is actually perceived by the users ($S_n = \{I_n\}$), two conceptual steps are considered in the literature:

1. Set choice ($p_n(I_n/S_n \neq \emptyset)$) that consists in the construction of sets perceived by users (formation level) and in obtaining the evaluation probability for each choice set perceived (extraction level) (Antonisse, Daly, & Ben-Akiva, 1985; Ben-Akiva, Bergman, Daly, & Ramaswamy, 1984; Morikawa, 1996; Russo and Vitetta, 1995; Russo, Vitetta, & Quattrone, 2007);
2. Route choice $p_n(k/I_n)$ that consists in the route choice evaluation of each alternative contained in the extracted sets (Ben-Akiva et al., 1984; Ben-Akiva & Bierlaire, 1999; Cascetta, Nuzzolo, Russo, & Vitetta, 1996; Russo & Vitetta, 2003; Ortùzar & Willumsen, 1994).

Following the possibility to use the Manski (1977) model, the explicit construction of perceived sets can be introduced straightaway into the four path choice models (pull and push; end-user and retailer). Our aim is to extend, also to attraction and acquisition macromodels, the two steps: generation and choice. The two probabilities, $p_n(I_n/S_n \neq \emptyset)$ and $p_n(k/I_n)$, thus need to be determined.

As regards the choice within a set, three problems are currently proposed and debated by analysts, representing three key aspects for modeling the behavior in question. The first concerns the need to consider, for the specification of the individual discrete choice model, different approaches other than consolidated approaches exemplified by random utility models (RUM), recalled in Section 18.4.1 above. Of the non-RUM approaches, of interest are those that consider possibility instead of probability, derived from fuzzy-type structures. For a comparison between random and fuzzy models for road route choice for national freight transport, see Quattrone and Vitetta (2011).

The second aspect concerns the temporal nature of the alternatives, hypothesizing that problems connected to spatial correlations have been implicitly cited. There is a need to be able to use models that take explicit account of user choice dynamics, especially for end-consumers. Indeed, the more commonly used models tend to be formulated first and then used, with reference to systematic trips historically prevalent in urban areas, with regard to work, school, and to systematic purchases. It is necessary to introduce elements that take explicit account of time variations, starting, for example, from models developed in psychology in the field of sequential analysis (Bakeman & Gottman, 1997; Gottman & Roy, 1990), considering the subsequent increase in explicit information, both from personal experience and from the presence of intelligent transportation systems (ITS).

The third problem concerns the dispersion of taste. The demographic changes in cities, both in number and size as recalled in Section 18.1, have altered the paradigm of systematic trips that has accompanied the development of analysis of urban transport systems. The advent of new forms of labor, and new forms of shopping and restocking, requires the use of the new generation of models. There are now models with particularly accurate analytical forms that allow the presence of various heterogeneities in preferences to be taken into account, hence allowing for the presence of latent attributes for variously defined classes or for continuous distributions on the whole population in question. In relation to more refined formulations, equally valid software is available for calibration (Greene & Hensher, 2010; Train, 2003). Such functional forms are required to ensure better modeling of the decision maker's behavior, allowing for subsequent variations previously introduced.

Acknowledgments

The author wishes to thank the anonymous referees for their suggestions, which were most useful in revising the chapter.

References

Antonisse, R. W., Daly, A. J., & Ben-Akiva, M. E. (1985). Highway assignment method based on behavioural models of car driver's route choice. *Transportation Research Record*, *1220*, 1–11.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction. An introduction to sequential analysis*. New York: Cambridge University Press.

Bayliss, B. (1988). *The measurement of supply and demand in freight transport*. Aldershot, England: Avebury.

Beckman, M., McGuire, M. C., & Winsten, C. B. (1956). *Studies in the economics of transportation*. New Haven, CT: Yale University Press.

Ben-Akiva, M. E., Bergman, M. J., Daly, A. J., & Ramaswamy, R. (1984). Modelling interurban route choice behaviour. *Proceedings of the 9th international symposium on transportation and traffic theory* (pp. 299–330). Utrecht, Netherlands: VNU Science Press.

Ben-Akiva, M. E. & Bierlaire, M. (1999). *Discrete choice methods and their applications to short term travel decisions*. Chapter for the Transportation Science Handbook, Preliminary Draft.

Ben-Akiva, M. E., & de Jong, G. (2008). The aggregate-disaggregate-aggregate (ADA) freight model system. In M. E. Ben-Akiva, H. Meersman & E. Van de Voorde (Eds.), *Recent development in transport modelling: Lessons for the freight sector*. Bingley, UK: Emerald Group Publishing Limited.

Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA: The MIT Press.

Ben-Akiva, M. E., Meersman, H., & Van de Voorde, E. (2013). *Freight transport modeling*. Bingley, UK: Emerald Group Publishing Limited.

Cao, X. (2012). The relationship between e-shopping and store shopping in the shopping process of search goods. *Transportation Research Part A*, *46*, 993–1002.

Cascetta, E. (2007). *Transportation system analysis: models and applications*. Berlin: Springer.

Cascetta, E., Nuzzolo, A., Russo, F., & Vitetta, A. (1996). A new route choice logit model overcoming IIA problems: Specification and some calibration results for interurban networks. In Jean-Baptiste Lesort (Ed.), *Proceedings of the 13th international symposium on transportation traffic theory*. Pergamon Press.

Census. (2011). Estimated quarterly US Retail Sales: Total and e-commerce (Table 1). Washington, DC.

Chow, J. Y. J., Yang, C. H., & Regan, A. C. (2010). State-of-the art of freight forecast modeling: Lessons learned and the road ahead. *Transportation*, *37*, 1011–1030.

Circella, G., & Mokhtarian, P. L. (2010). *Complementarity or substitution of online and in-store shopping: An empirical analysis from Northern California*. Paper no. 10-0894 presented at the 89th annual meeting of the Transportation Research Board (conference DVD), Washington, DC, January 2010.

Couclelis, H. (2001). Pizza over the internet: E-commerce, the fragmentation of activity, and the tyranny of the region. *Paper presented at the workshop on entrepreneurship, ICT and the region*, Amsterdam, June 7–8.

Crainic, T. G., Ricciardi, N., & Storchi, G. (2004). Advanced freight transportation systems for congested urban areas. *Transportation Research Part C*, *12*, 119–137.

Daganzo, C. (1991). *Logistics systems analysis*. Berlin: Springer-Verlag.

De Jong, G., Vierth, I., Tavasszy, L., & Ben-Akiva, M. E. (2012). Recent developments in national and international freight transport models within Europe. *Transportation*. doi: 10.1007/s11116-012-9422-9 published online: 13 June 2012.

Farag, S., Schwanen, T., Dijst, M., & Faber, J. (2007). Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping. *Transportation Research Part A*, *41*, 125–141.

Golob, T., & Regan, A. C. (2001). Impacts of information technology on personal travel and commercial vehicle operations: research challenges and opportunities. *Transportation Research Part C*, *9*, 87–121.

Gonzales-Feliu, J., Toilier, F., & Routhier, J. L. (2010). End consumer goods movement generation in French medium urban areas. In Taniguchi, E., & Thompson, R. G. (Eds.), *Procedia Social and Behavioral Sciences*, 2. Science Direct, 6189–6204.

Gottman, J. M., & Roy, K. A. (1990). *Sequential analysis*. Cambridge: Cambridge University Press.

Greene, W. H., & Hensher, D. A. (2010). *Modeling ordered choice*. Cambridge: Cambridge University Press.

Holguín Veras, J. (2002). Revealed preference analysis of the commercial vehicle choice process. *Journal of Transportation Engineering*, *128*(4), 336–346. doi:10.1061/(ASCE)0733-947X(2002)128:4(336)

Manheim, M. (1979). *Fundamentals of transportation system analysis*. Cambridge, MA.

Manski, C. (1977). The structure of random utility models. *Theory and Decision*, *8*, 229–254.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). New York, NY: Academic Press.

Modenese Vieira, L. F. (1992). *The value of service in freight transportation*. PhD Dissertation, Massachusetts Institute of Technology, Department of Civil Engineering, Cambridge, MA.

Mokhtarian, P. L. (2004). A conceptual analysis of the transportation impacts of B2C e-commerce. *Transportation*, *31*, 257–284.

Mokhtarian, P. L. & Tang, W. (Laura). (2013). Trivariate probit models of pre-purchase/purchase shopping channel choice: Clothing purchases in Northern California. In S. Hess & A. Daly (Eds.), *Choice modelling: The state of the art and the state of practice – Selected papers from the second international choice modelling conference*. Edward Elgar: Cheltenham, UK: Edward Elgar.

Morikawa, T. (1996). A hybrid probabilistic choice set model with compensatory and non compensatory rules. *Proceedings of the 7th WCTR*, Sydney, Australia.

Munuzuri, J., Cortes, P., Onieva, L., & Guadix, J. (2010). Modelling peak-hour urban freight movements with limited data availability. *Computers & Industrial Engineering*, *59*, 34–44.

NHTS. (2009). National Household Travel Survey Average Annual PMT (Person Miles of Travel), Person Trips and Trip Length by Trip Purpose (Table 5). U.S. Department of Transportation, Federal Highway Administration.

Ogden, K. W. (1992). *Urban goods movements*. Hants, England: Ashgate.

Ortùzar, J. de D., & Willumsen, L. G. (1994). *Modelling transport* (3rd ed.). Wiley.

Peterson, R. A., Balasubramanian, S., & Bronnenberg, B. J. (1997). Exploring the implications of the Internet for consumer marketing. *Journal of the academy of marketing science*, *25*(4), 329–346.

Polimeni, A., Russo, F., & Vitetta, A. (2010). Demand and routing models for urban goods movement simulation. *European Transport-Trasporti Europei*, *46*, 3–23.

Quattrone, A., & Vitetta, A. (2011). Random and fuzzy utility models for road route choice. *Transportation Research Part E: Logistics and Transportation Review*, *47*(6), 1126–1139.

Rizer, C., Cornelis, E., Browne, M., & Leonardi, J. (2010). GHG emissions of supply chains from different retail systems in Europe. In Taniguchi, E., & Thompson, R.G. (Eds.), *Procedia social and behavioral sciences*, 2. Science Direct, 6154–6164.

Russo, F. (2005). *Sistemi di trasporto merci: Approcci quantitativi per il supporto alle decisioni di pianificazione strategica tattica ed operativa a scala nazionale*. Milan, Italy: Franco Angeli.

Russo, F., & Comi, A. (2010). A modelling system to simulate goods movements at an urban scale. *Transportation*, *37*(6), 987–1009.

Russo, F., & Comi, A. (2011a). Measures for sustainable freight transportation at urban scale: Expected goals and tested results in Europe. *Journal of Urban Planning and Development*, *137*(2), 142–152.

Russo, F., & Comi, A. (2011b). A model system for the ex-ante assessment of city logistics measures. *Research in Transportation Economics*, *31*(1), 81–87.

Russo, F., & Musolino, G. (2012). A unifying modelling framework to simulate the Spatial Economic Transport Interaction process at urban and national scales. *Journal of Transport Geography*, *24*, 189–197.

Russo, F., & Vitetta, A. (1995). Networks and assignment models for the Italian national transportation system. *Proceedings of the 7th WCTR*, Sydney, Australia.

Russo, F., & Vitetta, A. (2003). An assignment model with modified logit, which obviates enumeration and overlapping problems. *Transportation*, *30*(2), 117–201.

Russo, F., & Vitetta, A. (2011). Reverse assignment: Calibrating link cost functions and updating demand from traffic counts and time measurements. *Inverse Problems in Science and Engineering*, *19*(7), 921–950.

Russo, F., Vitetta, A., & Quattrone, A. (2007). Route choice modelling for freight transport at national level. *Proceedings of the European transport conference*, Strasbourg, France.

Salomon, I., & Koppelman, F. S. (1988). A framework for studying teleshopping versus store shopping. *Transportation Research A*, *22*(4), 247–255.

Sheffi, Y. (1985). *Urban transportation networks: Equilibrium analysis with mathematical programming methods*. Englewood Cliffs, NJ: Prentice-Hall.

Taniguchi, E., Thompson, R. G., Yamada, T., & Van Duin, R. (2001). *City logistics – Network modelling and intelligent transport systems*. Oxford, UK: Pergamon.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press.

Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*. Part II, Vol. 1, 325–378. London, UK.

Chapter 19

Urban Freight Distribution: Urban Supply Chains and Transportation Policies

Romeo Danielis, Elena Maggi, Lucia Rotaris and Eva Valeri

Abstract

The chapter moves from the belief that adopting a supply chain approach is crucial to understand how urban freight distribution works and what will be the impact of the various potential urban transport policies on supply chain performance. According to a recent strand of the literature, the chapter aims at (a) characterizing the urban supply chains; (b) discussing how an urban supply chain can be modelled, which role do actors play and how the coordination issue can be handled; (c) showing how transport decisions, in particular whether to use own-account or third-party transport operators, are dealt within each urban supply chain and by each actor; and (d) analysing how urban supply chains are affected by the many proposed freight transport policies. Although much progress has been made in the field, both with regard to modelling and to the empirical analysis, it is concluded that important progress needs to be made with regard to both the *ex-ante* and the *ex-post* evaluation of the private and social efficiency of the different urban supply chains and on how they are impacted by local authorities' transport policies.

Keywords: Urban freight transport; supply chain approach; models; policy impact evaluation

19.1. Urban Freight Distribution as an Urban Supply Chain Issue

In order to analyse the impact of public transport policies on urban freight distribution, in line with some recent literature, in this chapter we underline the

Freight Transport Modelling

ISBN: 978-1-78190-285-1

importance to adopt a supply chain approach. The urban transport policy, in fact, might influence the level of efficiency and effectiveness not only of the urban distribution chain, but also of the whole supply chain.

For a firm, the *logistics* challenge consists of planning, controlling, implementing and monitoring all internal and network-wide material, part and product flows, including the necessary information flows, for the purpose of satisfying customer needs and realizing a profit. The main fields of the logistics activity can be broken into procurement logistics, production logistics, distribution logistics, sales logistics and disposal logistics. The more complex the firm structure (in terms of number of manufacturing plants and warehouses), the greater the difficulty of efficiently managing its logistics, particularly when part of the chain is operated by third-party firms.

When many firms work together, the term *supply chain* is used to identify a system of activities, people, technologies, information and resources targeted at transferring a product or a service along the entire chain: from the provider of raw materials to the end-customer. *Supply chain management*[1] is the term used to identify the set of techniques used in analysing and managing a supply chain.

When the end-consumers are located in urban areas, a supply chain does include a distribution system set in that complex environment that is a town or a city. In this chapter, we use the term *urban supply chain* (hereafter, USC) to identify the part of a supply chain in charge of delivering the goods to an urban area.

Since supply chain decisions are typically taken in order to achieve commercial efficiency disregarding wider environmental objectives, city authorities try to reconcile the conflict between the private objectives and the social ones. The branch of science dealing with these issues is called *city logistics* (see e.g. Taniguchi & Thompson, 1999).

The efficient and effective performance of the USCs is crucial for the attractiveness and competitiveness of a city. However, because of its complex nature – consisting of intertwined actors (producers, wholesalers, transport operators, retailers and consumers) with possibly conflicting objectives and multiple constraints – and because of the interaction with other urban functions – such as the residential, touristic, commercial, recreational, educational function – achieving a good performance is not an easy task.

The need for investigating urban freight distribution from a USC point of view is recognized by several authors. Since in an urban setting the relevant infrastructures and agent behaviours interact in especially complex and important ways, Friesz, Ilsoo, and Cheng-Chang (2011) model such interaction as a dynamic not cooperative Nash game and highlight that the performance of any aspect of the USC depends

1. Gibson, Mentzer and Cook (2005, p. 23) state: 'Supply Chain Management encompasses the planning and management of all activities involved in sourcing and procurement, conversion, and all logistics management activities. Importantly, it also includes coordination and collaboration with channel partners, which can be suppliers, intermediaries, third-party service providers, and customers. In essence, supply chain management integrates supply and demand management within and across companies'.

on the performance of the rest of the supply chain. Arunotayanun and Polak (2009, p. 2) write: 'Existing approaches to the empirical analysis of freight demand have, almost without exception, ignored the influence of supply chain and logistics concepts and have instead relied to conceptual and methodological approaches developed in the passenger sector. Little attention has been paid either to characterizing the actual behaviours of freight agents in this evolving structure of supply chains or to how to most effectively deploy modern demand modelling techniques to accommodate these complexities. Hence, little progress has currently been made in understanding and modelling freight agents' behaviours embraced with the supply chain concepts, leaving a large gap in this area of research'. Such remarks, that Arunotayanun and Polak (2009) formulate having in mind intercity, business-to-business freight modelling, is also true or, we would dare to say, it is even truer, for urban freight distribution because of the reasons that will be illustrated below. Aggregate models do have a value, but a deeper understanding of the supply chain dimension in which urban freight distribution takes place is essential to understand and forecast how a given set of policy or regulations affect goods distribution.

Similar conclusions are drawn by Browne and Gomez (2011) in a paper that investigate and quantify the impact of delivery restrictions on costs and environmental performance for a distribution operation. They state that their research 'highlighted the importance of adopting a supply chain approach to the removal of restrictions and the need for public/private sector cooperation'.

This chapter contributes to the existing literature by (a) characterizing the USCs; (b) discussing how a USC can be modelled; (c) showing how transport decisions are dealt within each USC; and (d) analysing how USCs are affected by the many proposed freight transport policies.

19.2. Characterizing Urban Supply Chains

As the USC has a very complex nature and can assume different profiles according to the characteristics of the city and of the other urban activities, a standardized characterization is not possible, but a USC perspective is needed. Nevertheless, according to the literature and to the results of some European projects, such as CityPorts (Regione Emilia Romagna, 2005), Bestfus (www.bestufs.net) or City Freight (http://www.transport-research.info), it is possible to classify the USC by three dimension: (a) product characteristics; (b) structure and organization of the whole supply chain; (c) structure and organization of the last mile, that is, the retail systems.

An initial classification of the USCs can be based on the product type (called, hence, product USC). The most common product types distributed in an urban area are

- fresh fruit and vegetables;
- milk and dairy products;

- meat and fish;
- bread and pastry;
- dry food: including long-life packaged products and beverages;
- frozen food: including ice-cream products;
- clothing and footwear;
- pharmaceutical products;
- flowers;
- newspapers;
- home furnishing and electronics;
- stationary and tobacco.

Important attributes of a USC relate to consignment frequency, consignment time, consignment size and the possibility of consolidation with other product supply chains. As illustrated in Table 19.1, compiled drawing from the literature and from

Table 19.1: Product USC and their logistics characteristics

	Consignment frequency	**Consignment time**	**Consignment size**	**Possibility of consolidation with other product supply chains**
Fresh fruit and vegetables	Daily	Fixed	Large	No
Milk and dairy products	Daily	Fixed	Small	No
Meat	Weekly	Flexible	Large	No
Fish	Daily	Fixed	Large	No
Bread and pastry	Daily	Fixed	Small	No
Dry food and beverages	Weekly	Flexible	Large	No
Frozen food	Weekly	Flexible	Small	No
Clothing and footwear	Seasonally	Flexible	Large	Yes
Pharmaceutical products	More than once a day	Fixed	Small	No
Home furnishing and electronics	Monthly	Flexible	Large	Yes
Stationary and tobacco	Weekly	Flexible	Small	Yes
Newspapers	Daily	Fixed	Small	No
Flower	Daily	Flexible	Small	No

the authors' own field research (Danielis, Rotaris, & Marcucci, 2010; Marcucci & Danielis, 2008), such attributes are quite differentiated among product USC.

Starting with consignment frequency, Table 19.1 indicates that the USC requiring the higher frequency, often more than once a day, is that distributing pharmaceutical products. In Italy the 17,000 pharmacies existing in 2006 were supplied almost exclusively by 138 wholesalers via 150 distribution centres (Dallari, 2006). The pharmacies share an electronic database with real-time listings of the quantity of products required by each pharmacy, and the information is transmitted to the wholesalers via electronic data interchange and the Internet. The up to 4-times-a-day consignments are performed by third parties hired and paid directly by the wholesalers, using refrigerated trucks and multi-drop routing.

Fresh fruit and vegetables, milk and dairy products, fish, bread and pastry and newspapers are supplied on a daily basis. The common feature of these USCs is that they need to guarantee that their products are fresh and new every day: this implies highly effective procurement systems. In the case of fresh fruit, vegetable and fish, the products might have also very distant origins. They arrive at the local wholesaler markets and, at least in Italy, they are directly bought by the retailers and carried to their shops using their own vehicles. Milk and dairy products, bread and pastry and newspapers, are distributed by specialized wholesalers or local producers.

Meat, dry food and beverages, frozen food and stationary and tobacco do not require daily supply but tend to be supplied on a (bi-)weekly basis. These products are typically distributed by specialized wholesalers serving more than one city with their own vehicles.

Home furnishing and electronics and clothing and footwear tend to be consigned once a month or even once a season. In case of special needs, the consignments are arranged via couriers.

Consignment time varies also among USCs. The ones that require daily consignments are often characterized by fixed time requirements because the product has to be available on the shelf usually early in the morning. The remaining USCs are more flexible: they can be consigned during the day or even during off-shopping hours.

As to the consignment size, this usually depends on the product type. Fresh fruit and vegetables, meat, fish, clothing and footwear, dry food and beverages and home furnishing and electronics usually have large consignment sizes, where the remaining USCs have typically small consignment sizes.

With regard to the possibility of consolidation between different product supply chains, this is limited, to some extent, to clothing and footwear, home furnishing and electronics and stationary and tobacco. These products are usually packaged and transported by third-party transport operators.

A second classification of the USCs can be based on the structure and organization of USC. A USC can be short, or direct (as illustrated by Kotler & Armstrong, 1999). Examples of direct USCs can be found in the delivery of fresh fruit and vegetable, bread and pastry, clothing and footwear, pharmaceuticals and flowers. In such cases the producers sell their goods directly to retailers. This type of organization avoids intermediary costs, but requires an effort by the producer or by the retailer in order to organize, via own-account or third-party, the transportation of the goods. This form of USC is typically used either for perishable and/or locally produced goods, or when the consumers are willing to pay a premium price for goods that come directly from a specific producer.

If the production of the good is fragmented among many small firms, the producers can enhance their bargaining power and increase their logistics efficiency by joining in associations (or cooperatives) aimed at collecting and selling the product to the retailers. Similarly, if the retail system is fragmented in many small firms, the retailers may join in associations or cooperatives.

The vast majority of the products are, however, intermediated by wholesalers, including general- (cash and carry), single-brand and specialty-line wholesalers. They are independent intermediaries that buy goods from manufacturers and sell them to retailers. Since they take title to the goods, they assume certain risks and can suffer losses if products are damaged, become out-of-date or obsolete, are stolen, or are not sold. At the same time, because they own the products, they are free to develop their own marketing strategies including setting prices. Merchant wholesalers include full-service wholesalers and limited-service merchant wholesalers. Limited-service wholesalers include cash-and-carry wholesalers, truck jobbers, drop shippers, mail-order wholesalers and rack jobbers. Wholesalers play often a central role in organizing a USC and on the delivery of the goods to the retailers' stores.

Alternatively, the intermediary function of a USC might be performed by agents and brokers.

Indeed, the complexity of the channels increases as the number of actors interacting along the distribution chain raises, implying, if channel members adopt mark-up strategies, higher management costs for producers and higher prices for consumers. However, having a large number of intermediaries allows the producers to reach many target markets and the consumers to choose among broader assortments of products.

Based on the literature and on the authors' field research, the five USCs types are associated with the product UCS as represented in Table 19.2.

It can be seen that some product types can be distributed via multiple structure and organizational types of USC. The most notable example is fresh fruit and vegetables which can be distributed by three type of USC (direct, wholesalers and producers cooperatives).

A third classification of USCs focuses on the retail systems that supply the final consumer.

Two features differentiate the retail systems that distributed the goods in an urban area: dimension and property.

Originally, all goods where distributed via small, independent stores and few bigger regional stores. They were mostly family business with the owner working him\herself in the shop. In the last decades, such distribution systems have been complemented by stores organized and owned by retail groups (often, stock companies or cooperatives) offering one-stop shopping for a wide variety of products in stores of different dimensions and complexity: convenience stores (less than 3,000 square feet); supermarkets (3,000–25,000 square feet); superstores (25,000–60,000

Table 19.2: Structure and organizational types of UCSs and product types

Type of USCs	Product type
1 Direct supply	Fresh fruit and vegetables; bread and pastry; clothing and footwear; pharmaceutical products; flower
2 Producers' cooperatives	Fresh fruit and vegetable; milk and dairy products; fish
3 Retailers' cooperatives	Pharmaceutical products
4 Wholesalers	Fresh fruit and vegetables; meat; fish; dry food and beverages; frozen food; clothing and footwear; flower; newspapers
5 Brokers and agents	Stationary; dairy products; dry food and beverages; frozen food

square feet); hypermarkets (above 60,000 square feet). The latter retail systems are characterized by large economies of scale, economies of scope and enhanced purchasing power which allows them to negotiate lower prices from their suppliers.

The size and the scope of the retail system have a strong influence on the procurement systems used. The small independent stores have small or no warehousing room. The goods are often delivered to them using small trucks. Usually they do not have a vehicle and consignment take place during the shopping hours. On the contrary, a retail group has large warehousing space, vehicles and personal in charge of goods procurement (often procurement takes place at the group level not at the store level) and consignment size and time is chosen by the retailers on the basis of his\her own needs. Such features allow the retail groups to have higher consolidation levels, larger trucks, night time deliveries and, in general, a greater attention on the private efficiency of the procurement logistics.

Whereas family businesses specializes in specific products (all above product USC are possible), retail groups manage and offer many product USC at the same time.

To the above classification, two more retail systems should be added: (a) the home retail system which delivers the goods bought on the Internet directly to the consumer's house and (b) the Ho.Re.Ca. (Hotel, Restaurant and Catering) system, where the end consumers directly consume a specific variety of goods, not a home, but in a hotel, restaurant, bar or cafe. Since the Ho.Re.Ca. system is often located in the historical centre of a city, the goods' procurement to this retail system has highly specific features.

19.3. Modelling the Urban Supply Chain

Within a supply chain coordination is the key feature. Lack of coordination imply: inaccurate forecast, low capacity utilization, excessive inventory and inventory costs, inadequate customer service, multiple inventory turns, bad time to market, slow order fulfilment response, bad quality, insufficient customer focus and poor customer satisfaction.

Coordination may take three forms: (a) vertical integration; (b) long-term contracts aimed at increasing the total SC profit, reducing the overstock/understock costs and sharing the risks among the partners; (c) short-term contracts.

The simplest USC is the direct producer-retailer supply chain. A successful producer-retailer relationship is crucial for the efficiency of an urban distribution system. It has been shown that the optimal results, both from private and system point of view, are obtained when the strategies of the two actors are coordinated (Goyal, 1976). This requires that both the vendor (producer) and the buyer (retailer) make their data available to the other party to better plan the inventory management activities and that they agree on how to share the extra-profits gained via the inventory cost reduction. Full cooperation takes place in practice also when the vendor is authorized to manage the inventory levels of the buyer, taking decisions about the order quantity, the timing of the order, the reorder point and the replenishment of inventory. Such a case is termed the vendor-managed inventory system and allows the vendor to assume responsibility for maintaining inventory levels and determining order quantities for its customers. The vendor-managed inventory system has proven to be particularly beneficial for products having high demand variance and high outsourcing costs. In such a system, the operations of the vendor and buyer can be integrated through information sharing by using the information technologies such as electronic data interchange or internet-based protocols. The vendor can use this information to plan production, schedule deliveries and manage inventory levels at the buyer. Formally the problem can be modelled as a joint economic lot size model so that system costs are reduced and capacity utilization is increased. The benefits of the vendor-managed inventory (VMI) system have been widely recognized in different industries, especially in the retail industry, for example, Wal-Mart, Kmart, Dillard Department Stores and JCPenney are among the earlier adopters of VMI (Dong & Xu, 2002, p. 75).

Modelling the optimal solution of this inventory control problem requires making some assumptions about:

- the demand: whether it is deterministic (external demand at each designated activity are known in advance with certainty) or stochastic (demand is assumed known up to a given probability distribution) and whether unsatisfied demand can be backlogged or not (unsatisfied demand is not retained);
- the supply chain: whether it is a single-product or a multi-product one;
- the time dependency of the problem: whether the parameters of the problem are stationary or not over time;
- the inventory position of each actor: whether it can be periodically reviewed or not.

Various types of coordination mechanisms have been used in supply chain coordination literature such as quantity discount, credit option, buy back/return policies, quantity flexibility, commitment of purchase quantity, etc.

Coordinating USC partners is not an easy task, since it requires: sharing strategic information, defining a common data collection methodology and monitoring and measuring the management performance on the basis of the profits of the USC as a whole and not of the profits gained by each partner. A graphical illustration of the coordination issues in the supply chains is illustrated by Arshinder and Deshmukh (2008, p. 329).

The coordination task is even more difficult if all the actors involved in the delivery activities within the urban context are considered, since it requires to detect and include both all the actors belonging to the supply chain and the stakeholders affected by the SC activities within the urban context (residents, workers, students, consumers, tourists) and to take into account their specific highly heterogeneous needs, as illustrated by Schoemaker, Allen, Huschebeck, and Monigl (2006, p. 8) (Table 19.3).

In the real world, unless the parties involved represent components of the same company supply chain, cooperation might not prevail since there is often a superior/subordinate relationship inherent in the situation where the dominant party prefers her/his priorities to lead the solution. Moreover, in the reality the urban freight distribution activities are very often fragmented and de-structured, with lack of coordination between the different actors having conflicting goals. As a result, decentralized modelling of the problem is necessary. Two models can be envisaged: a vendor-driven decentralized model and a buyer-driven decentralized model.

In a vendor-driven decentralized model, the vendor has the greater channel power and makes decisions (e.g. supply, replenishment and manufacturing) independently to maximize its individual profit. Consequently, the buyer takes decisions subject to the vendor's optimal decisions. A vendor-driven decentralized model consists of two sub-problems: (1) the vendor maximizes its profits subject to his constraints and (2) the buyer maximizes its profits subject both to his constraints and to the vendor's optimal decisions.

The opposite takes place in the buyer-driven decentralized model.

Table 19.3: Interests of stakeholders involved in UCSs

Stakeholders	**Interests**
Residents	Product and services
	Negative environmental impact
Retailers	Competitiveness and profitability
Authorities and Public Service	Accessibility
	Governance and legislation
	Negative environmental impact
Suppliers	Market growth
	Profitability
Carriers	Congestion
	Cost-effectiveness

Source: Anand, Yang, van Duin, and Tavasszy (2012, p. 11945).

Without coordination between the vendor and buyer, a decentralized vendor–buyer system is profit inefficient because double marginalization takes place, since each party makes decisions considering only a portion of the total system profit. In fact, the decentralized model is used as a benchmark to evaluate the increase in efficiency deriving from the various coordination mechanisms.

In real markets, however, the issue is even more complicated because each buyer is typically supplied by many producers and each supplier generally serves many buyers. As a result, the joint economic lot size optimization problem requires to coordinate the deliveries carried out by the suppliers' network to the buyers' network and to take into account not only the inventory but also the transportation costs (Kim & Goyal, 2009).

The supply chains comprising the producers' cooperatives or the retailers' cooperatives represent a potential coordinating network to optimize inventory and transportation decisions.

The models discussed above mainly focus on the integration between two levels. But a supply chain can be made of more than two levels and its scope can be much larger. Technically, the models dealing with the optimization of the inventory and supply activity of more than two levels are called multi-echelon models. The most common notion of a multi-echelon inventory system is one involving a number of retail outlets (stores, facilities, installations' bases) to satisfy customer demands for goods and which, in turn, act as customers of higher-level wholesale activities (warehouses, depots, factories). The wholesale activities themselves may be customers of still higher-level wholesale activities or production facilities. It is important to note that the system might pertain to a single product such as a particular kind and brand of food, but that for the same grocery store a different product may have a completely different structure (there may be a different factory, or no regional warehouses, or a different mix of retailers, etc.). A multi-echelon inventory system can be portrayed as a directed network wherein the nodes represent the various activities or facilities in the system and the linkages represent flows of goods and information. If the network has at most one incoming link for each node and flows are acyclic (that is there are no loops in the network), it is called 'arborescence' or inverted tree structure. Although much more complex interconnected systems of facilities can exist (a retailer may obtain re-supply from more than one wholesaler, or a wholesaler may procure from more than one factory, or a retailer may sometimes supply another retailer), most of the work in multi-echelon inventory theory has been confined to arborescence structures (Clark, 1972).

As in the vendor-managed inventory system, the problem consists in managing the inventories throughout the entire supply chain so that for any echelon the stocking level should be planned taking into account not only its own inventory position but also the inventory of all the downstream echelons. The optimal planning of the inventory level for each echelon should not be determined on the basis of the demand information derived from the next downstream echelon, but on the demand from the end-consumers, reducing the demand variability and, consequently, the inventory costs. Yang and Wee (2001), for instance, show that the integrated approach results in a significant cost reduction compared to that of independent decision making by each individual entity of the supply chain.

Only few European projects focusing on city logistics use a supply chain approach to analyse freight movements within the urban context (Patier & Routhier, 2008). According to Bestufs I (www.bestufs.net), for example, only a few cities are using planning tools for urban freight transport aimed at modelling distribution chains, although, for evaluation of measures and forecasts, such models proved to be efficient tools. City Freight (http://www.transport-research.info), instead, examines some selected supply chains in order to detect their impacts on urban congestion. According to this research, urban freight is part of transport and logistics chains which often involve a larger area than only one city, therefore the problems, objectives, solutions, benefits and drawbacks beyond the city's boundaries should be considered. Furthermore, City Ports (Regione Emilia Romagna, 2005) provides a methodology to identify and analyse the supply chains mostly contributing to urban freight movements and to detect both the main actors involved in the organization of the delivery activities and the constrains characterizing each supply chain in terms of time windows, vehicle characteristics, personnel, storage facilities, etc. (Gentile & Vigo, 2007).

Finally, it is important to highlight that the efficiency and the effectiveness of the USCs are strongly dependent also on (a) the organization of the other urban activities, (b) the frequency and the type of conflicting interactions between passengers and freight flows, (c) the characteristics of the city in terms of geographical conformation and land use planning (urban shape), (d) the urban transport and infrastructure policies, concerning not only freight but also and primarily passengers flows. All this elements are exogenous factor that constrain the choices and the performance of the USC. Despite the importance of this issues, the editorial constraints doesn't allow us to explicit consider also them.

Having identified from a theoretical point of view how a USC can be modelled, we turn in the next section to the description of some real-world example, concentrating mainly on the transportation features.

19.4. Urban Supply Chains and the Organization and Performance of the Transportation Activities

The structure of the USC influences the transport organization. In the following subsection it will be described which transport decisions are taken by which actor of the supply chain with regard to whether to be involved in the delivery\procurement of goods or to leave it to other actors of the supply chain or to specialized third-party transport operators. When an actor decides to be part of goods delivery\procurement, s\he needs to buy a vehicle or use her\his own car. The question is then which vehicles are bought in terms of size and type of engine (diesel, electric, alternative fuels). Next, it is interesting to know how efficiently the vehicle is used in terms of load factor.

In a research project carried out at the University of Trieste, various product supply chains have been studied and a field survey has been made, mostly in Rome and in Trieste, to answer these research questions.

Generally, three actors of the supply chain were interviewed: the producer, the wholesaler and the retailer. For the details about each chain refer to Danielis (2013). For example, we report in Figure 19.1 the structure of the food supply chain.

The food supply chain is one of the most complicated. The producers buy inputs from other suppliers or wholesalers. The delivery is performed by transport operators or own-account by the wholesaler. Producers deliver their goods with own vehicles or using third-party transport operators to the wholesaler, to the retailers and to end-consumers. Some producers deliver also to supermarkets, usually with own vehicles.

The wholesalers buy the goods from the producers. The procurement is performed with own vehicles, by third-party transport operators and with the producers' vehicles. The wholesalers deliver the goods to the retailers, to supermarkets and to other wholesalers. The wholesalers normally make use of their own vehicles, only sometime they make use of third-party transport operators.

The food retailers buy directly from the producers, transporting the goods with their own vehicles or with the producers' vehicles. Alternatively, they buy from the wholesaler with deliveries made by the wholesalers' vehicles or by third-party transport operators. Furthermore, they buy from the supermarkets or other retailers, with own vehicles or third-party transport operators' vehicles.

In our sample, two out of ten retailers consign their products also to the end-consumer using own vehicles. When a third-party transport operator is hired, the sender is in charge of organizing the delivery. Note that own-account transport tend to prevail in the food USC relative to third-party transport.

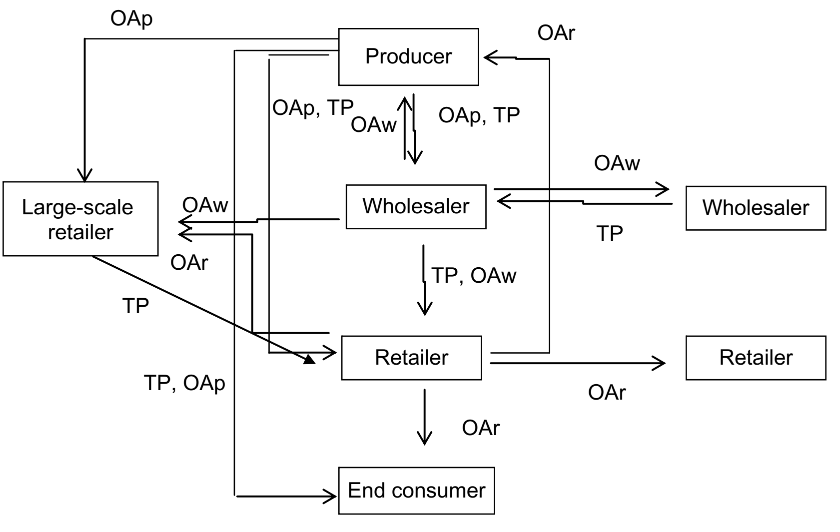

Figure 19.1: Actors and transport relationships in the food supply chain. *Legend*: TP = third-party transport; OAp, OAw, OAr = respectively, producers' own account, wholesalers' own account and retailers' own account.

19.4.1. The Degree of Transport Outsourcing

The survey allowed us to estimate the degree of transport outsourcing across urban product supply chains, using as an indicator the number of weekly deliveries made via own-account or third-party. As reported in Table 19.4, the results are quite differentiated among USCs.

In the interviewed sample, USCs in which the use of third-party transport operators prevails are clothing and footwear, frozen food, building material and publishing. The USCs in which own account prevails are bread and sweets, frozen foods, dry foods, wines and liquors, stationery, flowers and plants, drugs, tobacco and newspapers and the fish cool. The deliveries are made in similar percentages by own-account or third-party transport operators in these USCs: fresh food, soft drinks and minerals, furniture and electrical and computer science.

A specific section of the survey asked for reasons why own-account or third-party transport are chosen or not chosen. The results are reported in Tables 19.5 and 19.6.

Table 19.5 shows that on average maintaining a direct relationship with the customer and controlling over the service quality are the most important factors in the decision to perform own-account transport instead of using third-party transport operators. Performing joint services, managing jointly warehousing and transport and purely economic reasons are of secondary importance. Even less relevant are the possibility of carrying advertisements on the vehicles, fiscal advantages, the locked-in situation of having already made the investments and the possibility of performing

Table 19.4: Own account and third party: percentage of weekly deliveries

Urban supply chains	Own account (%)	Third party (%)
Clothing and shoes	12.5	87.5
Food: sweets and diet food	71.4	28.6
Fresh food	46.2	53.8
Ice cream	0.7	99.3
Frozen food	78.4	21.6
Dry food	79.6	20.4
Fresh fish	71.4	28.6
Drink and beverages	47.9	52.1
Alcoholic beverages	87.9	12.1
Stationary	99.3	0.7
Flowers and plants	76.3	23.7
Construction material	6.6	93.4
Furniture	50.0	50.0
Printing material	0.0	100.0
Drugs and beauty products	81.1	18.9
Electric and computer products	49.7	50.3
Newspapers and tobacco	100.0	0.0

Table 19.5: Advantages of performing own-account transport (54 interviews)

Motivation	Food	Furniture	Beverages	Statio-nary	Pharma-ceuticals	Flow-ers	Construction material	Electric and computer	Average
Control over service quality	21	27	27	44	17	22	22	8	22
Direct relationship with customers	29	22	32	22	17	31	17	11	25
Joint services during the consignment	6	17	9	11	11	9	17	33	13
Advertisement	7	8	1	0	0	2	0	3	4
Cheaper than TP	12	8	6	6	0	18	28	15	11
Previous investments	5	2	12	0	50	0	11	6	7
Joint warehousing and transport management	15	8	6	0	0	16	6	11	11
Fiscal advantages	3	5	1	11	0	2	0	9	3
Checks on quality and prices	3	3	6	6	6	0	0	3	4
Column total	100	100	100	100	100	100	100	100	100

Table 19.6: Disadvantages of performing own-account transport (58 interviews)

Motivation	Cloth and shoes	Food	Furniture	Beverages	Stationary	Printing	Pharmaceuticals	Flowers	Construction material	Electric and computer	Average
Need to have a person in charge	0	7	6	8	0	14	10	16	14	31	11
Cost of drivers	29	20	15	25	0	0	36	29	25	5	20
More expensive than TP	29	8	10	21	0	43	12	0	13	3	11
Costs of vehicles	43	40	56	26	57	0	33	50	34	39	38
Lower service quality	0	2	1	0	0	0	0	0	7	0	1
Unsuitable for long distances	0	21	11	18	0	43	7	6	7	12	15
Inventory management not feasible	0	2	1	2	43	0	2	0	0	10	3
Column total	100	100	100	100	100	100	100	100	100	100	100

checks on quality and prices at the customer store. The answers, however, vary among supply chains. The pharmaceutical wholesalers deem as an important motivation the fact that previous investments have already been made. In the electronics and computers supply chain the possibility of performing joint services when delivering the goods is seen as relevant. For the construction materials supply chain the economic advantages are very large. In the flower and in the beverages supply chain the direct relationship with the customers is considered extremely important.

Turning to the disadvantages of performing own-account transport, buying and maintaining a vehicle and the need for a driver are considered the most relevant (Table 19.6). Also a disadvantage is its unsuitability for long distances and the need to have a person in charge of organizing deliveries. It may also turn out disadvantageous from an economic point of view. The effect on service quality and the increased difficulty of running a warehouse are not seen as disadvantages. Also for the disadvantages, there is large variation among product supply chains. The need of having a person in charge of managing the deliveries is seen as a disadvantage especially in the electronic and computer supply chain. The need of hiring drivers is felt as a problem primarily in the pharmaceutical supply chain. The printing material supply chain deems that the own-account is more expensive and unsuitable for long distances, while the negative effect on inventory management is felt in the stationary supply chain.

19.4.2. The Degree of Efficiency in the Use of the Own-Account Vehicles

The general question of measuring the private and social efficiency of a USC will be deal with in the last section. At this point we focus specifically on some aspects of efficiency in the use of vehicles. In the sample the average number of vehicles is equal to two for manufacturers, four for the wholesalers and one for retailers. The producers have vehicles registered on average in 2004 (7 years old, since the survey took place in 2010), the wholesalers in 2005 and the retailers in 2003. Regarding the fuel type, diesel vehicles are mostly used by all stakeholders. The loading rate has been estimated both for procurement and delivery (Tables 19.7 and 19.8).

It results that in procurement the average loading rate for own account is equal to 61% for retailers, 55% for producers and 40% for wholesalers. In particular, the retail chains of flowers and plants and wines and liquors have a higher loading rate whereas it is significantly lower than the dry food. The producers have high loading rates in the procurement of dry food and furniture. The wholesalers, instead, have a rate high enough only in the case of pharmaceuticals, equal to 60%, while in other chains the value is lower.

Given that the use of their own account with respect to the third party responds to a variety of needs-such as the characteristics of the product and the requirements of speed, the length of the paths, etc.., – retailers and manufacturers are in procurement more efficient from the standpoint of the loading rate than wholesalers.

Table 19.7: Loading rate for own-account vehicles in procurement

Urban supply chains	**Producer (%)**	**Wholesaler (%)**	**Retailer (%)**
Clothing and shoes			50
Food: sweets and diet food	15		
Fresh food		23	62
Frozen food	50	50	
Dry food	90		33
Fish			65
Drink and beverages		50	75
Flowers and plants	50		80
Furniture	72		
Drugs and beauty products		60	
Electric and computer products		17	
Average value	55	40	61

Table 19.8: Loading rate for own-account vehicles in distribution

Urban supply chains	**Producer (%)**	**Wholesaler (%)**	**Retailer (%)**
Food: sweets and diet food	70		
Fresh food		80	
Frozen food	50	80	
Ice cream	75		
Dry food		40	100
Drink and beverages	100	78	
Alcoholic beverages		78	43
Flowers and plants	50	100	10
Construction material		100	
Furniture	85	100	
Drugs and beauty products		50	
Electric and computer products		60	
Average value	72	77	51

In distribution, the picture is rather different. First of all, note that the wholesalers use own account much more often than producers and retailers. It is therefore very interesting and meaningful that the loading rate of own-account vehicles is the highest for wholesalers, compared to producers and retailers rate, which is 77%. Namely, the phase of the distribution is handled by the wholesaler also with attention to the efficiency in achieving a higher degree of vehicles filling on their own account. This is especially true in some sectors, such as furniture, building material and flowers and plants, where the loading rate is 100%.

19.5. Impact of Urban Freight Transport Policies on Urban Supply Chains

In this section the impacts on USC and consequently on the whole single supply chain of five specific urban freight transport policies, more frequently implemented by the local authorities, will be examined: goods vehicle time-access regulation, vehicle type restrictions, loading\unloading (L\U) policies, fiscal policies and the promotion of urban transhipment and consolidation centres (UTCC). Other relevant policies such as public–private partnerships, information and communication technology, intelligent transportation systems and land use planning will not be discussed in this chapter not because they are not relevant but in order to limit the scope of the chapter.

The research question on how these public policies affect the different USCs will be discussed from a conceptual point of view and drawing from the so far limited existing literature.[2]

Access times restrictions, very frequently used in practice, aim at limiting the use of road space to trucks in favour of passenger traffic and the liveability of the city. Since this policy implies a reduction of the time available for delivery, it imposes an additional constraint in the search of a solution to the carriers' routing optimization problem.

The impact of access time restrictions on own-account carriers is likely to be marginal since they perform a single origin-destination trip (e.g. from the general fruit and vegetables market to their shop) often outside the restricted time windows. On the contrary, third-party carriers have to solve a more difficult routing optimization problem, often with multiple origin-destination points, spread all along the working time and on geographically distant shops. Hence, an additional constraint is likely to impose large adjustment costs, generate suboptimal routing, decrease load factors and require larger fleets and a larger number of lorry drivers (Holguín-Veras, Silas, Polimeni, & Cruz, 2007; 2008). Large retail organizations are an intermediate case between own-account and third-party operators. Such speculation is confirmed by Quak and de Koster (2009) who study the impact of governmental time windows, vehicle restrictions and different retailers' logistical concepts on the financial and environmental performance of retailers. They use a case study with two cases that differ in their drop sizes as input for an experiment. They show that the cost impact of time windows is the largest for retailers who combine many deliveries in one vehicle round-trip. The cost increase due to vehicle restrictions is the largest for retailers whose round-trip lengths are restricted by vehicle capacity. Vehicle restrictions and time windows together do not increase a retailer's cost more than individually. They also find that variations in delivery volume and store dispersion hardly influence the impact of urban policy and the retailer's logistical concept decisions.

2. Danielis et al. (2010) and Maggi (2004, 2007) have previously discussed this aspect at some length.

Quak and de Koster (2007) discuss which dimensions related to a retailer's network structure and logistical planning will determine its sensitivity to time-windows. They focus on various retailer characteristics (product characteristics, network structure, logistical planning) and distribution performance (operational, financial, environmental). Since they use 14 cases, the sample is too small to consider any statistical analysis. They conclude that the impact of increasing time-window pressure varies among different retailers. The retailers that supply more stores during the time-window hours – thanks to the short distance, short unloading time and larger drop size – are affected the least. On the contrary, the retailers which use their vehicles most during a 24-h period in the current situation are the most affected by time-windows.

Turning to the impacts on the product USCs, one observes that pharmaceutical products although distributed by third-party carriers using multi-drop routing are, in practice, exempted from this restriction, hence, they suffer no impact. With regard to fresh food, some products (especially meat and milk) are distributed by third-party carriers using multi-drop routing practices, whereas others (such as non-industrial fish, fruit and vegetables) are procured via own-account transport. If the former cannot adjust to the time window regulation by developing an early schedule in non-restricted times because of self-implied time-window constraints, the impact is likely to be relevant. Some segments of Ho.Re.Ca. are impacted when supplied by third-party carriers or wholesalers. Lastly, clothing and footwear will bear only marginal effects since deliveries are less frequent and mostly single-drop. Both Ho.Re.Ca. and clothing and footwear might have self-implied time windows because of staff constraints.

Vehicle restrictions, either by weight, engine type or load factor, aim at containing the environmental and congestion externalities. They cause an increase in fleet size, and an accelerated fleet renewal rate. Larger companies are likely to be able to cope with this policy better than small firms or own-account carriers. The most affected distribution channels are those which require large quantities and frequent deliveries. Pharmaceutical products are delivered frequently by small vehicles since the size of the parcels is generally small. Moreover, they are typically exempted from this regulation, hence they are not affected. On the contrary, fresh food, being characterized both by large quantities and frequent deliveries are highly affected, especially when distributed via large retail organization, because of the weight restrictions, and via own account, because of the engine type restrictions. Ho.Re.Ca. requires less frequent deliveries and small to medium quantities. Hence, it should suffer relatively less the impact of this policy, although procurement takes often place by own-account transport. Clothing and footwear involves infrequent deliveries, seasonal large quantities and small replenishment deliveries. Hence, it is only marginally impacted. A potential side effect is on the size of the consignments, probably reduced in order to be transported with smaller vehicles.

Urban goods distribution needs proper spaces for L\U activities. Unless a shop has its own private, internal L\U bay, it relies on off-street or on- street parking spaces. This generates a conflict with the parking needs for passenger cars, highly requested in central areas. The aim of *L\U policies* is, hence, both to regulate the loading bay use among truck users and to prevent private cars to use the L\U bays. L\u policies affect substantially consignment costs and times. The actual practice of irregular on-street parking contains the carrier costs and times but generates high congestion costs. The provision of a larger number of L\U bays and their effective enforcement would probably leave unchanged or slightly increase the private costs and times, but certainly reduce social costs. Pharmaceutical products are frequently delivered but require short L\U times since the parcels are typically small, hence, the impact of a L\U regulation is likely to be modest. On the contrary, fresh food is highly impacted since it is frequently delivered, and requires medium\large L\U times. Most shops located in city centres do not have internal bays for L\U activities, with the exception of large retail organizations who generally have personnel dedicated to stock management activities. Ho.Re.Ca. products are less frequently supplied with medium L\U times; hence, the impact of a L\U regulation is likely to be small. Clothing and footwear stores are occasionally supplied in large quantities for seasonal orders and require large L\U times. More frequently, though, they are supplied in small quantities with short L\U times.

Fiscal policies, either in the form of congestion toll or, more frequently, area licensing fees, are used in some cities to regulate access to central areas. They aim to internalize congestion costs and to achieve an optimal congestion level. The literature on fiscal policies' impacts has concentrated its efforts on the effects on passenger's flows or even on generic road users, rarely distinguishing the freight users from the passenger users. The objective of this literature is to develop useful model capable of capturing traffic dynamics, in terms of path choices modification and the resulting network flow patterns, and heterogeneous users' responses to toll charges for the design and evaluation of time-dependent pricing schemes (Lu & Mahmassani, 2011; Lu, Mahmassani, & Zhou, 2008). The reduction of travel time and the increase of punctuality in the deliveries improve the efficiency of freight urban transport. A study of 2001 in Stockholm has estimated that 'the total gain in travel matched about 2–3% of the total time production, and was valued to about € 50 million per year' (Eliasson & Lundberg, 2002). It is possible to suppose that in certain USCs, where the time is a key competitiveness factor, the value of the reduced travel time is higher than the charge to pay. Also the study of Greater London Authority (2006) confirms that the London congestion charging helps to free road space and improve journey times for service and delivery vehicles.

Fiscal policies should also induce an increase in load factors and multi-drop deliveries. As a result social efficiency of the overall transport system should be achieved.

Congestion tolls affect negatively consignment costs, unless the value of the time saved is higher than the fare. But this happens only when the value of the goods transported is high. A survey of Transport for London (2006) on the retail, hotels & restaurants, wholesale, manufacturing, education and construction, highlights that just under half of businesses in the charging zone and over half of businesses in the boundary zone consider transport costs to be negligible. Of those businesses in the charging zone, 13% of businesses suggest that transport costs account for over 10% of total business costs. Thus, for the majority of sectors, there are no patterns that indicate a possible congestion charging effect on the business performance. The City of Stockholm (2006) founds that, in most cases, the road pricing had a marginal impact on the companies' overall transport costs, but generally transport costs account for a very small proportion of the final price of goods and services. In particular, the cost to companies of congestion taxes over a 12-month period equated to less than 0.5% of the value of total production of goods and services in the area.

As area licensing fees are not proportional to the number of daily access entries, the higher is the daily consignment frequency the lower will be the per-trip cost. As a consequence, the delivery costs of own-account carries are more impacted by fees than those of third-party carriers since their load factors are lower. Finally, according to the literature (McKinnon, 2006), the cost of the fiscal policy can be more easily shifted from carriers to retailers when consignments are less-than-truck load.

Danielis et al. (2010) speculate that the impact of fiscal policies might differ among distribution channels. It is observed that pharmaceutical products are generally exempted. With reference to fresh food, when an area licensing fee is applied, the impact on fresh food delivery costs is likely to be small since they are characterized by frequent deliveries of large quantities. On the contrary, the impact on the Ho.Re.Ca. and clothing and footwear delivery costs are likely to be relatively higher, since they have less frequent consignments made mostly via own-account. However, the final effect on goods prices paid by final consumers is uncertain, as well as the effect on land rents since they depend on the characteristics of the specific markets. As McKinnon stated (2006), fiscal policies produce most likely small effects, at least in the short run, on strategical and commercial decisions, while they might have some effect on the tactical and operational ones.

The empirical evidence on the impact of the pricing policies on the USCs is still scant.[3]

Some evidence relates to the impact of the congestion charging on the overall economy and in some cases on retail. Despite the difficulty of discerning the impact on business performance from the range of factors impacting the economy (Transport for London, 2006), congestion charging seems to have a neutral or very modest impact on the economy, depending on charge level (Hensher & Puckett, 2005; Holguín-Veras, 2006; Whitehead, 2005; for London see Ernest & Young, 2006; Transport for London, 2006, 2008; Department for Transport (DfT), 2005; for Stockholm case see City of Stockholm, 2006; Mattsson, 2003; Eliasson, Hultkrantz, & Rosqvist, 2009; and Eliasson, 2008; for Trondheim case see Tretvik, 2003).

Winsor-Cundell (2003), however, illustrating the results of a London Chamber of Commerce and Industry investigation on the impact of London congestion charging based on a direct survey on 1,430 retailers, states that 76% of the respondents report reduced revenues and a fall of productivity, 33% consider relocation and 28% consider closing their business due to the charge. Food shops appear to be more

3. As Beria and Boggio (2012) state the literature on acceptability (both ex-ante and ex-post) is much more detailed than the literature on the actual effects of the implemented measures.

impacted than luxury goods. Moreover, 'even though just-in-time delivery is common in many sectors, the short shelf-life of food products means that food distribution has particularly limited flexibility in responding to new charging regimes' (Steedman, 2006). Very frequently, in the food supply chain any additional costs are passed down the chain to the last ring.

Similar conclusions are reported in London First (2006). The retail and leisure sectors, small businesses and those close to the boundary of the zone are considered to have been most adversely affected. As concerns the size of firms, a survey on retailers within the London charging zone in 2003 developed by the Commission for Integrated Transport (2003) founds that small retailers of convenience and food products, with respect to the large ones, appear to be less able to take advantage of the reduced congestion, more concerned with the costs of the scheme than with the benefits. This is exacerbated by the widespread practice of couriers, express delivery and other transport organizations imposing surcharges for deliveries in the charging area or reducing service levels (e.g. three deliveries a week in place of daily deliveries). The suppliers increase the price of their services in order to make a quick profit, generally taking no account of increased efficiency as a result of reduced delivery time or increased reliability. Instead, the chains of convenience stores have the power to exert pressure on suppliers, avoiding the surcharges.

However, May, Koh, Blackledge, and Fioretto (2010) claim that the impact of congestion pricing on the economy has been much smaller than that predicted by the business community.

In a counterfactual study, Quddus, Carmel, and Bell (2007) focus on the John Lewis Oxford Street (JLOS) store, one of the biggest retail stores located within the charged zone. He also finds a drop in sales attributable to congestion charging of 5.5% by the time-series model and of 8.2% by the panel model over a period of about eleven months following the introduction of the charge, compared to the other five stores outside the area.

To opposite conclusions comes the study performed by Daunfeldt, Rudholm, and Ramme (2009) using revenue data from 14 shopping malls, 9 within the charged area of Stockholm and 5 outside the charged area. The data also include revenue data from a sample of retail stores located along the main shopping streets in Stockholm. The results show that the Stockholm road pricing trial did not negatively affect retail revenue, neither in shopping malls nor in the sample of retail stores. As possible explanations for this result the authors argue that (a) probably that most stores and shopping malls are open in evenings and on weekends in Sweden, making it easy to avoid the congestion fee by changing the time when the shopping is being done, (b) people to a larger extent uses public transport for shopping trips, (c) as parking fees are quite expensive in the Stockholm city, it is also likely that car-borne shoppers are high-income earners that are less sensitive to changing their shopping behaviour when congestion charges are introduced and (d) as habits that change slowly, it is possible that the retail businesses might be affected negatively by the introduction of congestion charges even though the results do not support so far this view.

A further interesting issue is the spatial impact of a charging scheme. Löchl (2006) performs a theoretical analysis and reviews the existing evidence. The case of

Trondheim has been particularly researched after the introduction of road pricing in 1991. Avant Management A/S (1992) found that 10% of the customers had changed their shopping behaviour by moving their shopping to other destinations or times after the introduction of the cordon pricing. Moreover, while business people located in the city centre had predicted major negative swings in trade prior to the cordon pricing, the Chamber of Commerce of Trondheim concluded from an own ex-post survey that there was hardly any effect on trade at all. Anyway, there was a long lasting general trend of growth in areas outside and decline in areas of the cordon. Tretvik (1999) even concludes a general trend line of modest but steady growth in retail sales in real terms inside the cordon since the introduction as he later reports (Tretvik, 2003, p. 89).

Finally, one other possible impact of road pricing is also the increase of home deliveries from shops to final consumer (Steedman, 2006). According to Winsor-Cundell (2003), since the introduction of the congestion charge in London there was an 11.8% increase in telephone orders. The small size of these deliveries and the widespread location of the consumers can cause problems of vehicle load factor maximization, negatively influencing the congestion.

The *development of a UTCC* aims at optimizing the consolidation and routing patterns of the existing distribution channels and at using less polluting vehicles. Since it introduces an extra node in the distribution channel which imposes extra logistics costs, own-account or third-party operators or the logistics coordinator of the distribution channel might not be willing to use it. The actual urban goods distribution regulation may obviously influence the decision favouring the use of UTCC vehicles against all other non-UTCC vehicles (see the case of Vicenza, Italy). Given the previous discussion we believe that pharmaceutical products and fresh food will not make use of a UTCC, since the specific characteristics of the goods distributed require dedicated and integrated channels and infrastructures, strong logistics coordination and fast and frequent deliveries. On the contrary, clothing and footwear and Ho.Re.Ca, especially when supplied via own-account, clothing and footwear for occasional replenishment orders and Ho.Re.Ca for goods other-than-fresh food might accept to use a UTCC. Gonzalez-Feliu and Morana (2010), analysing the case of Cityporto, the UTCC for the city of Padua (Italy), in operation since 2004, report that the main customers are third-party transport operators, which buy the last mile service from Cityporto, who is also granted the privilege of accessing the limited traffic zone of Padua, hence avoid the inconvenience of entering the city centre. More recently, also a soft drinks distribution company has signed a partnership with Cityporto for restaurant and bar deliveries.

19.6. Conclusions and Further Research Needs

This chapter moves from the belief that adopting a supply chain approach is crucial to understand how urban freight distribution works and what will be the impact of the various potential urban transport policies. The chapter has tried to contribute to the existing literature results on the following issues: (a) characterizing the USCs; (b) discussing how a USC can be modelled, which role do actors play and how the coordination issue can be handled; (c) showing how transport decisions, in particular whether to use own-account or third-party transport operators, are dealt within each USC and by each actor; and (d) analysing how USCs are affected by the freight transport policies more frequently proposed by the local authorities.

Although much progress has been made in the field, both with regard to modelling and empirical analysis, we think that important progress needs to be made with regard to both the *ex-ante* and the *ex-post* evaluation of the private and social efficiency of the different UCSs and on how they are impacted by transport policies.

Adopting at the urban scale the methodologies proposed for the supply chain evaluation can be a way forward and should be tested. As reviewing the various supply chain evaluation methodologies and the corresponding literature, which is outside the scope of this chapter, we just mention the two most important ones. One is the Supply-Chain Council's Operations Reference (SCOR) model that decomposes the processes within a supply chain and incorporates multiple performance indicators into one measurement system. It deals with (1) reliability measures (e.g. fill rate, perfect order fulfilment); (2) cost measures (e.g. cost of goods sold); (3) responsiveness measures (e.g. order fulfilment lead-time); and (4) asset measures (e.g. inventories). The second one is called Balanced Scorecard model and employs performance metrics from financial (e.g. cost of manufacturing and cost of warehousing), to customer (e.g. on-time delivery and order fill rate), business process (e.g. manufacturing adherence-to-plan), innovation and technology perspective (e.g. new-product development cycle time). Although it is aimed at measuring the producer–customer relationship, that is two subsequent agents within the same supply chain, it can easily be adapted to track the supply chain as a whole.

Adapting and applying these methodology to the different USCs, we believe, would greatly enhance our understanding on how to improve freight distribution in an urban area.

References

Anand, N., Yang, M., van Duin, J. H. R., & Tavasszy, L. (2012). GenCLOn: An ontology for city logistics. *Expert Systems with Applications*, *39*, 11944–11960.

Arshinder, A. K., & Deshmukh, S. G. (2008). Supply chain coordination: Perspectives, empirical studies and research directions. *International Journal of Production Economics*, *115*, 316–335.

Arunotayanun, K., & Polak, J. (2009). Accounting for supply chain structures in modelling freight mode choice behaviour. Paper presented at the European transport conference 2009. Leeuwenhorst Conference Centre, The Netherlands.

Avant Management A/S. (1992). *Handelsundersøkelsen 1992, Trondheim med omkringliggende kommuner*. [The trade inquiry of 1992, Trondheim with surrounding municipalities.] quoted by Ciari, F., Löchl, M., & Axhausen, K.W. (2008). Location decisions of retailers: an agent-based approach. 15th International conference on recent advances in retailing and services science, Zagreb, July 2008.

Beria, P., & Boggio, M. (2012). *The role of transport supply in the acceptability of pollution charging extension. The case of Milan*. XIV conference of the Italian Association of Transport Economics, Bari, Italy, 28th–30th June 2012.

Browne, M., & Gomez, M. (2011). The impact on urban distribution operations of upstream supply chain constraints. *International Journal of Physical Distribution & Logistics Management*, *41*(9), 896–912.

City of Stockholm. (2006). *Stockholmsforsoket: Facts and results from the Stockholm trials*, first version. Congestion Charge Secretariat, Stockholm.

Clark, A. (1972). An informal survey of multi-echelon inventory theory. Working Paper Serial T-259. The George Washington University School of Engineering and Applied Science Institute for Management Science and Engineering, University of Washington DC.

Commission for Integrated Transport. (2003). *The impact of congestion charging on specified economic sectors and workers*. Retrieved from http://webarchive.nationalarchives.gov.uk/20110304132839/http://cfit.independent.gov.uk/pubs/2003/index.htm

Dallari, F. (2006). *La filiera farmaceutica. Quali criticità e quali prospettive per la logistica dei farmaci in Italia?* [The pharmaceutical supply chain. Which are the key nodes and the perspectives of drug logistics in Italy?]. Centro di Ricerca sulla Logistica, Verona, Italy, 25 October 2006. Retrieved from http://www.liuc.it/ricerca/clog/dallari_pharma25ott06.pdf

Danielis, R. (Ed.), (2013). La distribuzione urbana delle merci. Analisi e politiche. [Freight Urban Distribution. Analyses and Policies] Trieste, Italy: EUT.

Danielis, R., Rotaris, L., & Marcucci, E. (2010). Urban freight policies and distribution channels: A discussion based on evidence from Italian cities. *European Transport/Trasporti Europei*, *46*, 114–146.

Daunfeldt, S.-O., Rudholm, N., & Ramme, U. (2009). Congestion charges and retail revenues: Results from the Stockholm road pricing trial. *Transportation Research Part A*, *43*, 306–309.

Department for Transport (DfT). (2005). *Transport and land use interaction in the context of road pricing*. Discussion paper. Department for Transport (UK). Retrieved from www.dft.gov.uk

Dong, Y., & Xu, K. (2002). A supply chain model of vendor managed inventory. *Transportation Research Part E*, *38*, 75–95.

Eliasson, J. (2008). Lessons from the Stockholm congestion charging trial. *Transport Policy*, *15*, 395–404.

Eliasson, J., & Lundberg, M. (2002). *Road pricing in urban areas*, Vägverket, Publication 136E Swedish National Road Administration and T&E. Retrieved from http://www.transport-pricing.net/download/swedishreport.pdf

Eliasson, J., Hultkrantz, L., & Rosqvist, L. (2009). Stockholm congestion charging trial. *Transportation Research Part A*, *43*, 237–310.

Ernest & Young. (2006), *Review of transport for London's assessment of business and economic impacts of the congestion charge*. Third annual report 2005. Retrieved from http://www.london.gov.uk/mayor/economic_unit/docs/impacts-of-cc.pdf

Friesz, T. L., Ilsoo, L., & Cheng-Chang, L. (2011). Competition and disruption in a dynamic urban supply chain. *Transportation Research Part B*, *45*(8), 1212–1231.

Gentile, G., & Vigo, D. (2007). *Movement generation and trip distribution for freight demand modelling applied to city logistics*, Technical report, Università di Bologna, Bologna, Italy. Retrieved from http://151.100.152.220/gentile/wp-content/uploads/2010/04/39-MovGen-TS2007.pdf

Gibson, B. J., Mentzer, J. T., & Cook, R. L. (2005). Supply chain management: The pursuit of a consensus definition. *Journal of Business Logistics*, *26*(2), 17–25.

Gonzalez-Feliu, J., & Morana, J. (2010). Are city logistics solutions sustainable? The Cityporto case. *TeMa. Journal of Land Use, Mobility and Environment*, *3*(2), 55–64.

Goyal, S. K. (1976). An integrated inventory model for a single supplier-single customer problem. *International Journal of Production Research*, *15*(1), 107–111.

Greater London Authority. (2006). *Retail in London,* Working Paper I, Service and delivery issues, June 2006. Retrieved from http://www.london.gov.uk/publication/economic-development-publications-2000-2008#retail-london

Hensher, D. A., & Puckett, S. M. (2005). Refocusing the modelling of freight distribution: Development of an economic-based framework to evaluate supply chain behaviour in response to congestion charging. *Transportation*, *32*, 573–602.

Holguín-Veras, J. (2006). *The truth, the myths, and the possible in freight road pricing in congested urban areas*. Book Chapter in City Logistics VI, pp. 615–630.

Holguín-Veras, J., Silas, M., Polimeni, J., & Cruz, B. (2007). An investigation on the effectiveness of joint receiver–carrier policies to increase truck traffic in the off-peak hours. Part I: The behavior of receivers. *Network Spatial Economics*, *7*(3), 277–295.

Holguín-Veras, J., Silas, M., Polimeni, J., & Cruz, B. (2008). An investigation on the effectiveness of joint receiver–carrier policies to increase truck traffic in the off-peak hours, Part II: The behavior of receivers. *Network Spatial Economics*, *8*(4), 327–354.

Kim, T., & Goyal, S. K. (2009). A consolidated delivery policy of multiple suppliers for a single buyer. *International Journal of Procurement Management*, *2*(3), 267–287.

Kotler, P., & Armstrong, G. (1999). *Principles of marketing* (8th ed.). Upper Saddle River, NJ: Prentice Hall.

Löchl, M. (2006). *Land use effects of road pricing – A literature review*. 6th Swiss transport research conference paper, Monte Verità/Ascona (CH).

London First. (2006). *Getting London to work, London*. Retrieved from http://www.london-first.co.uk/documents/042_Getting_London_to_Work_-_full_report.pdf

Lu, C.-C., & Mahmassani, H. S. (2011). Modeling heterogeneous network user route and departure time responses to dynamic pricing. *Transportation Research Part C*, *19*(2011), 320–337.

Lu, C.-C., Mahmassani, H. S., & Zhou, X. (2008). A bi-criterion dynamic user equilibrium traffic assignment model and solution algorithm for evaluating dynamic road pricing strategies. *Transportation Research Part C*, *16*(2008), 371–389.

Maggi, E. (2004). Il governo dei flussi logistici per uno sviluppo urbano sostenibile. In F. Boscacci & E. Maggi (Eds.), *La logistica urbana. Gli elementi conoscitivi per una governance del processo [Freight urban logistics. The key issues for a process governance]* (pp. 117–145). Milan, Italy: Polipress.

Maggi, E. (2007). *La logistica urbana delle merci. Aspetti economici e normativi [Freight Urban Logistics. Economics and Normative Issues]*. Milan, Italy: Polipress.

Marcucci, E., & Danielis, R. (2008). The potential demand for a urban freight consolidation centre. *Transportation*, *35*, 269–284.

Mattsson, L. G. (2003). *Modelling road pricing reform in Stockholm*. IMPRINT-EUROPE seminar: Implementing Pricing Policies in Transport: Phasin, Leuven, Belgium.

May, A. D., Koh, A., Blackledge, D., & Fioretto, M. (2010). Overcoming the barriers to implementing urban road user charging schemes'. *European Transport Research Review*, *2*(1), 53–68.

Mckinnon, A. C. (2006). A review of European truck tolling schemes and assessment of their possible impact on logistics systems. *International Journal of Logistics: Research and Applications*, *9*(3), 191–205.

Patier, D., & Routhier, J. (2008). How to improve the capture of urban goods movement data. Paper presented at the session "Urban good surveys", *8th international conference on survey method in transport*, Annecy, France, May 2008.

Quak, H. J., & de Koster, M. B. M. (2007). Exploring retailers' sensitivity to local sustainability policies. *Journal of Operations Management*, *25*, 1103–1122.

Quak, H. J., & de Koster, R. (2009). Delivering goods in urban areas: how to deal with urban policy restrictions and the environment. *Transportation Science*, *43*, 211–227.

Quddus, M. A., Carmel, A., & Bell, M. G. H. (2007). The impact of the congestion charge on retail: the London experience. *Transport Policy*, *14*, 433–444.

Regione Emilia Romagna. (2005). *City Ports Project*. Quaderni del Servizio Pianificazione dei Trasporti e Logistic, n° 7, Bologna, Italy. Retrieved from http://mobilita.regione.emilia-romagna.it/entra-in-regione/Pubblicazioni/i-quaderni-del-servizio-pianificazione-dei-trasporti-e-logistica

Schoemaker, J., Allen, J., Huschebeck, M., & Monigl, J. (2006). *Quantification of urban freight transport effects I, best urban freight solutions II*. Retrieved from www.bestufs.net

Steedman, P. (2006). *Road pricing. Could it promote sustainable food systems?* Report of an expert workshop held in association with the RSA, Food Ethics Council, December 2006. Retrieved from http://www.foodethicscouncil.org/system/files/Road_pricing_expert_workshop.pdf

Taniguchi, E., & Thompson, R.G. (Eds.), (1999). *City logistics I*. Institute for City Logistics, First International Conference on City Logistics, 12–14 July, Australia, Institute of Systems Science Research, Kyoto.

Transport for London-TfL. (2006). *Central London congestion charging. Impacts monitoring*. Fourth annual report, June 2006. Retrieved from http://www.tfl.gov.uk/microsites/freight/publications.aspx

Transport for London-TfL. (2008). *Central London congestion charging. Impacts monitoring*. Sixth annual report. Retrieved from http://www.tfl.gov.uk/microsites/freight/publications.aspx

Tretvik, T. (1999). The EUROPRICE project: The Trondheim toll ring and the effects on retailing, *SINTEF Memo N-12/99*, SINTEF, Trondheim.

Tretvik, T. (2003). Urban road pricing in Norway: Public acceptability and travel behaviour. In J. Schade & B. Schlag (Eds.), *Acceptability of transport pricing strategies* (pp. 77–92). Oxford: Elsevier.

Whitehead, T. (2005). Road user charging and business performance: Identifying the processes of economic change. *Transport Policy*, *9*, 221–240.

Winsor-Cundell, E. (2003). *The retail survey. The impact of the congestion charge on the retail sector*, The London Chamber of Commerce and Industry Campaigns Team, May 2003. Retrieved from http://www.londonchamber.co.uk/DocImages/395.pdf

Yang, P. C., & Wee, H. M. (2001). A quick response production strategy to market demand. *Production Planning & Control*, *12*(4), 326–334.

Chapter 20

Modeling for Public Policies Inducement of Urban Freight Business Development

Rosário Macário

Abstract

Urban logistics, being responsible for the delivery of goods inside the city, is a sector of activity that has huge economical importance. However, it also poses serious problems to the other users of the urban space – congestion, noise, obstruction of roads and sidewalks, infrastructure degradation, and so on. These problems have enormous costs for the society and hinder the overall urban sustainability.

In this chapter, we recognize that the diversity of problems associated with urban logistics prevents the organization of this sector with a single solution. Instead we recognize diversity of urban profiles and of business models that must be successfully matched to solve the complexity that involves the sector. It is expected that this rational serves as the basis for designing a future master plan for urban logistics, which might be used by local authorities to regulate their territory and to implement measures that enable inducement of better behavior by the economic agents who take part of the urban supply chain.

The research was developed following a "3-step approach" methodology. The first step comprises the definition of "logistic profiles"; the second step encompasses the evaluation of different logistic solutions based on different criteria (e.g., logistic urban platforms, distribution centers, and car-sharing freight solutions) in terms of their suitability for serving different logistic profiles. This leads to the third and final step, which consists in modeling the changes that will occur if the solutions are implemented. This "3-step approach" is validated by modeling its application and benefits to a first pilot

Freight Transport Modelling

ISBN: 978-1-78190-285-1

area in Lisbon – Alvalade neighborhood, and more recently transferred to downtown.

Keywords: Urban freight; urban logistics; business models; decision maker; policy

20.1. Introduction

Globalization and other new market trends have turn down many of the barriers that for a long time kept companies from freely developing their operations. As a consequence, companies have engaged in new productive and distribution solutions, which embrace multiple agents based in very different locations. The transportation sector was called to bridge the growing distances between these locations, and to ensure a continuous supply of goods between these locations OECD (2003).

In fact the forecast is that changes in world economics will consubstantiate in displacement of factories to cheap-labor countries (Asia, Eastern Europe, Central and South America, North Africa, etc.), fragmentation of production with products no longer manufactured in a single factory, and instead provided by different producers in multiple different locations. This new geometry causes great impact in the logistic and supply chain giving room to the emergence of new distribution concepts CEC (2007a, 2007b).

From several studies, we could compile TURBLOG (2011) some indicative dimensions for what a city generates in urban freight. These figures report to the years 2005–2007[1]:

- 0.1 delivery or pick-up per person per day;
- 1 delivery or pick-up per job per week;
- 300–400 truck trips per 1000 people per day; and
- 30–50 tons of goods per person per year.

And urban freight represents:

- 10–15% of vehicle equivalent miles travelled in city streets
- 2–5% of the employed urban workforce.
- 3–5% of urban land is devoted to freight transport and logistics.

1. The figures presented in this paragraph have been compiled by the author from different scattered sources, since no homogeneous figures exist in the literature. These sources were Rodrigue (2006), Stantchev and Whiteing (2006), START (2008), Visser van Binsbergen, and Nemoto (1999), BESTUFS (2008), BESTUFS (2002), ECOTEC et al. (2007), FHA (2009), TURBLOG (2011).

In addition, we could also find the evidence that a city does not only receive goods, but is often also a place of shipping:

- 20–25% of all truck-km in urban areas are outgoing freight,
- 40–50% is incoming freight and
- 25–40% is originated from and is delivered within the city

Furthermore, the increasing activity of the logistic activities within urban areas is jeopardizing the sustainability of such regions and international organizations estimate that 75–80% of population lives in urban regions (DfT, 2003). Interpretation of the evolution of economics and society in general gives no indication of reverting the growing trend of urbanization (OECD, 2003).

Logistic companies in turn have no incentive in engaging for sustainable solutions, because the costs they are responsible for are partly being supported by the whole society, as externalities. This raises the need for regulation, so that externalities are properly incorporated in the activities that are provoking them and behaviors of respective agents. However, public intervention in urban logistics has had so far a rather narrow scope, being commonly limited to traffic restrictions (e.g., access times or vehicles' dimensions). One of the reasons behind this seems to be the lack of knowledge on adequate tools to deploy effective measures regarding urban logistic problems (Taniguchi & van der Heiden, 2000). Moreover, from the interviews undertaken, there is a strong belief that an increase in costs in logistic activities would result in a decrease in economic competitiveness for the companies and, ultimately, of the whole regions (Hensher & Puckett, 2005).

The most relevant impacts emerging from urban freight can be divided into the following:

- Economic impacts: congestion, inefficiency, and resource waste, and the effects on urban competitiveness;
- Environmental impacts: pollutant emissions including the primary greenhouse gas carbon dioxide, the use of nonrenewable fossil fuel, land and aggregates, and waste products such as tires, oil, and other materials;
- Social impacts: the physical consequences of pollutant emissions on public health (death, illness, hazards, etc.), the injuries and death resulting from traffic accidents, noise, visual intrusion, and other quality of life issues (including the loss of Greenfield sites and open spaces in urban areas as a result of transport infrastructure developments).

Furthermore, according to project TURBLOG (2011), there are other private costs to consider, "goods vehicle operators and drivers face a range of difficulties when carrying out freight operations in urban areas." These include:

- Traffic flow/congestion issues caused by traffic levels, traffic incidents, inadequate road infrastructure, and poor driver behavior;

- Transport policy-related problems including, for example, vehicle access restrictions based on time and/or size/weight of vehicle and bus lanes;
- Parking and loading/unloading problems including loading/unloading regulations, fines, lack of unloading space, and handling problems;
- Customer/receiver-related problems including queuing to make deliveries and collections, difficulty in finding the receiver, collection and delivery times requested by customers and receivers.

In this chapter, we recognize that the diversity of problems associated with urban logistics prevents the organization of this sector from using a single solution. Instead we recognize diversity of urban profiles and corresponding business models (Osterwalder, 2004) that must be successfully matched to solve the complexity involving the sector. We aimed to develop this rational to provide the basis for designing future master plans for urban logistics, which might be used by local authorities to regulate their territory and to implement measures that enable inducement of better behaviors by the economic agents who take part of the urban supply chain. Interviews done to stakeholders in different countries reveal that successful urban logistic solutions are the ones that bring a positive value proposition to the supply chain and/or to the shop business.

The research was developed following a "3-step approach" methodology. The first step comprises the definition of "Logistic Profiles"; the second step encompasses the evaluation of different logistic solutions based on different criteria (e.g., logistic urban platforms, distribution centers, car-sharing freight solutions, and so on) in terms of their suitability for serving different logistic profiles. This leads to the third and final step, which consists in modeling the changes that will occur if the solutions envisaged are implemented. This construct is validated by modeling its application and benefits to a first pilot area in Lisbon –Alvalade neighborhood, and at the moment of writing this chapter the concept is transferred to downtown Lisbon.[2] Finally, the work undertaken paves the ground for the next stage of the research, the development of the architecture for modeling urban freight distribution to support policy decision making.

20.2. Stages of Development on Urban Logistics

The study of urban freight tasks is complex and heterogeneous, involves an interdisciplinary engagement as a consequence of the difficulty to identify the common features between the requirements of different users and vehicle operators. Furthermore, urban freight is strongly interrelated with many other aspects of the

2. An extended survey for the city of Lisbon with data collection to be undertaken during February 2012. This survey addresses not only origin destination but also the variables that characterize the different logistic profiles as described in this chapter.

urban system: urban passenger system, land use, regional development, socio-economic environment, employment, etc. When considering urban freight planning it is necessary to devote some effort towards understanding its integration within urban mobility planning. As pointed out by Macário and Caiado (2004), ''acting on urban logistics domains implies intervening in different aspects of urban mobility management, particularly institutional, regulatory, social, infrastructural and technological, therefore requiring the joint and coordinated action of the different stakeholders in the urban logistics arena.''

In order to define best policies for urban freight distribution it is important to distinguish the different relationships that exist between urban areas, not only in terms of public authorities but also considering the distribution of public space (e.g., land uses and transport infrastructures), once recommendations should vary according to the geographic approach. For example, considering metropolitan areas, urban freight must be considered in master planning and a roadmap should be developed with guidelines for the development of business and commerce. In the other hand, in a small town, access and parking regulations can be useful, but in city centers sometimes these restrictions may not be sufficient, and so innovative measures have to be taken to change peoples' behavior.

In the case of a small city or a part of a city (e.g., city center) innovative last mile solutions have also to be considered, and in this case logistic business concept and logistic profiles have to be considered, in order to access the most efficient and effective improvements for urban logistics.

The following criteria were chosen to distinguish the type of urban areas:

- dimension of the urban areas (area and population);
- governance, that is, which level of authority are we giving advice (local authorities, government, environmental agencies);
- relationships between the city and the metropolitan area (intermodal hubs, transport infrastructure, important economic sectors, traditions).

Each city has its own life cycle and its growing pace and the success for the implementation of a good solution (and/or a policy) has to consider that pace as an indicator of the city's capacity to change.

Regarding the types of measures adopted in urban logistics, it is worth noticing that the emerging economies, such as China, India, Mexico, Chile, and Brazil, seem to be at an early stage of development with regard to urban logistics practices, compared to more developed countries, such as France, the Netherlands, and Japan.

The more developed countries show a broader range of measures, varying from restrictions to incentives and often including market-oriented initiatives by companies. In contrast, other cities in, for example, Latin America, India, and China seem to focus more on restrictions and/or measures influencing freight transportation in general.

Some cities are already one step ahead in terms of urban logistics development and their policy is close to the one adopted in the mobility management approach, they are striving to change the behavior, while others are still trying to organize the public space and transport accessibility, mainly focusing on transport infrastructure and traffic restrictions.

We can find three stages of development in what concerns urban freight, being: the first stage, the **restriction stage**, characterized by urban areas that show no evidence (or very little) of urban logistics policies in policy documents or even sections of other documents not logistic oriented. Also no evidence (or very little) of urban logistics solutions adopted. This stage corresponds to cases that are still raising awareness of the problems that urban logistics might cause and choosing to face these problems with access restrictions and some reduced incentives normally associated to price of access. The second stage, **new practices stage**, corresponds to urban areas that already have some evidence of urban logistics policies in policy documents and have some restrictions/incentives already being applied. In these cities there is typically some evidence of urban logistics solutions that are being adopted by the local authority, logistic operators, etc. Finally, the third stage, **policy and practice stage**, corresponds to urban areas that have already a good sample of policies, solutions and restrictions/incentives. This stage corresponds to cases that have solutions applied with good impacts but have already understood that it is not enough, they are striving to change behaviors that requires a deeper intervention.

In order to tackle the problems presented above, rationalization of the distribution process (from the economic, spatial and temporal perspectives) is required. This means reducing the flow of goods yet keeping the adequate level of distribution to satisfy consumer's needs. As a way to achieve this, various solutions have been pointed out in several cities. These solutions are not only aimed at the transport activity, but to the organization of the whole logistic chain. On some cases, the implemented solutions were adopted as "stand alone" solutions, focused on a specific case or problem; on some other cases, combined solutions were applied, that is, several measures were implemented in a combined form, as part of a broad political strategy for urban logistics (DETR, 1999; Freight Transport Association, 1998).

Generally, the different authors try to group the adopted solutions in accordance to different criteria (Muñuzuri, Larrañeta, Onieve, e., & Cortés, 2005). In this work, we decided to group the measures identified according to their focus of application. This systematization can be found in Table 20.1, where they were placed according to the degree of intervention needed for their implementation (i.e., from "soft" to "hard" measures)

Some of the measures identified are very ambitious, while others lack the needed coherence and fail to achieve the goals for which they were designed. This happens for a number of reasons, like the relative novelty of the introduction of this subject on the urban governance agenda, and also the lack of knowledge about the implementation processes involved in urban logistics and hence on the appropriate way to tackle the collateral problems accruing from these processes.

Another very important issue relates with the difficulties associated to modeling urban freight movements, given its fragmented character largely caused by the fact that a significant amount of transport is done on private basis. It is worth noting that in all experiences reported the main difficulty in handling urban freight lies in

Table 20.1: Measures implemented to solve the urban logistics problems[a]

Stage of development	Type of measure	Examples
Restriction stage	Access restriction measures	Access restrictions according to vehicle characteristics (weight or volume), conditioning access to pedestrian areas, urban tolls, periodic restrictions
Restriction stage	Legislative and organizational measures	Cooperative logistic systems, encouraging night deliveries, public–private partnerships, intermediate delivery depots
New practice stage	Territorial management measures	Creation of loading and unloading areas, of load transfers, and mini logistic platforms
New practice stage	Technological measures	GPS, track and tracing systems, route planning software, intelligent transport systems, adoption of non polluting vehicles and vehicles adapted to urban characteristics (size and propulsion)
Policy and practice stage	Infrastructural measures	Construction of urban distribution centers, and peripheral storing facilities, use of urban rail for freight (freight trams), underground freight solutions
Policy and practice stage	Policy oriented	Development of master plans, subordinate implementation and location of activities to impacts caused on mobility systems, implementation of different solutions to respond to different market segments

[a]Collected and summarized in TURBLOG research project at www.turblog.eu under the coordination of the author.

understanding the associated origin/destination (O/D) matrix and identify the different conditions that result from the interaction between type of product, type of business, and type or urban space.

20.3. The Rational for City Logistics: Logistic Profile Concept

The logistic profile (LP) concept (Macário et al., 2007a) is based on the hypothesis that it is possible to identify, for some well-defined areas inside a city, reasonably homogenous groups of logistic needs, based on three key points: the urban characteristics of the area, the requirements of the logistic agents (i.e., the requirements

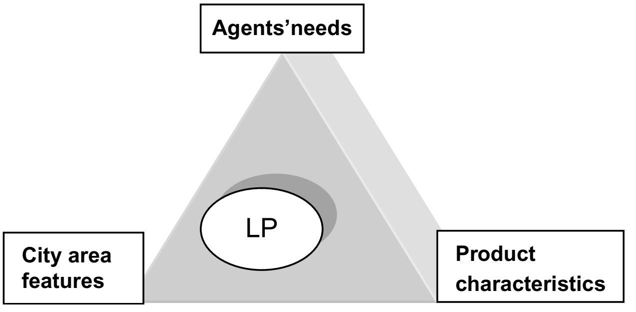

Figure 20.1: Definition of logistics profile.

concerning the type of delivery), and the characteristics of the products being transacted. The LP of a given urban area is thus defined by the interaction of these three key aspects (Figure 20.1).

Our hypothesis is that in the areas of the city in which LPs can be defined,[3] it will then be possible to adjust urban logistic services that will optimize the consumption of the involved public and private resources (space, vehicles, etc.), in function of the needs of the different market segments.

Several variables are used to qualify the logistic profiles; these are identified bellow (Table 20.2a, 2b, and 2c). These variables can be combined in an array of different LPs. However, more than getting a full comprehensive list of profiles, it is of higher importance to stress that it is possible to match different logistic services to specific logistic profiles, in order to optimize the delivery processes. The way this is done is explained bellow.

A key characteristic of the urban logistic system is its high number of stakeholders and the heterogeneity of their needs, which force a very thin segmentation of the proposed services, often going below the minimum scale for economic feasibility. The introduction of the LP concept has the main goal of introducing market segmentation in urban freight, giving the possibility to best fit different delivery solutions to the different needs of each profile. Following this rational, Table 20.3 matches the most relevant LPs found with the different delivery services identified above, and analyses their fitness for each of the former. This table does not mean to be exhaustive, but simply to show how the devised methodology works. It is also worth stressing that slight changes on each LP might result in different levels of adequacy of the different services. For instance, by simply changing the special needs of the products mentioned on the first LP on the Table 20.3, makes the solution "Hierarchic Terminals with bicycle or other eco-friendly, light delivery service" become unsuitable for the resulting profile.

3. It will only be possible to define a LP of an area if it has a significant group of homogenous "logistic needs," which is a minimum critical mass.

Table 20.2a: City area features. Variables used to determine the logistic profiles (TURBLOG, 2011)

Features	**Classification**		
1.1. Commercial density	*Low* <30% Commercial face to residencies/services/industry	*Medium* 30–70% Commercial face to residencies/services/ industry	*High* >70% Commercial face to residencies/services/industry
1.2. Homogeneity	*Low* Several types of services and products	*Medium* Mix of residential areas with offices and commercial stores	*High* Cluster of one type of service or similar products
1.3. Logistic accessibility	*Bad*	*Reasonable*	*Good*
1.3.1. Measures considering logistic needs	Bad level of access between the shop and the parking (e.g., no loading bays)	Some specific measures considering logistic needs (e.g., loading bays non-exclusive)	Transport network suited for the logistic needs (e.g., exclusive loading bays)
1.3.2. Level of congestion	High level of traffic congestion (commercial speed < 3 km/h)	Reasonable (high on peak hours)	Low (fluid traffic – commercial speed >12 km/h)
1.4. Restriction applied	*Yes* Off-peak hours, week days, …	*No*	

Table 20.2b: Product features. Variables used to determine the logistic profiles (TURBLOG, 2011)

Characteristics	Classification		
2.1. Easiness of handling	*Difficult*	*Reasonable*	*Easy*
2.1.1. Size	Large (wheelbarrow, crane)	Medium (>1 person to carry one unit)	Small (>1 unit per person to carry)
2.1.2. Weight	Heavy (wheelbarrow, crane)	Medium (>1 person to carry one unit)	Light (>1 unit per person to carry)
2.1.3. Holding conditions	Difficult	Reasonable	Easy
2.2. Special conditions	*Special needs* (e.g., valuable products and frozen products)	*Might have special needs* (e.g., open packages, if food handled ambient temperature, and chilled)	*No special needs*
2.2.1. Fragility	Fragile	Might have special needs	No special needs
2.2.2. Perishability	Perishable	Not perishable	

Table 20.2c: Agents profile/deliveries profile. Variables used to determine the logistic profiles (TURBLOG, 2011)

Characteristics	**Classification**		
3.1. Urgency of deliveries	*Irrelevant*	*Relevant*	*Urgent*
3.2. Frequency of deliveries	*Low*	*Medium*	*High*
	< once a week	Several days per week	Daily
3.3. Amounts to be delivered	*Few*	*Several*	*Many*
3.3.1. Number of shops	One shop	Several shops	Retail center/ big shops
3.3.2. Vehicles weight and size	Light goods vehicle or smaller vehicles	Van/small truck	Heavy goods vehicles
3.4. Planned deliveries	*No defined routine*	*Defined routine* (e.g., after hours deliveries, 8–10 a.m., …)	

To have a more in-depth approach and get consistent guidelines for the creation of an urban logistics master plan the following requirements must also be addressed:

- Legal and regulation framework for offered services;
- Regulation of motorized vehicles access and temporary urban space occupancy, for loading and unloading activities;
- Road infrastructures with appropriate loading/unloading bays for the activities served by those bays;
- Logistic infrastructures and information systems that support the goods distribution;
- Regulation of security requirements for the transport and handling of goods;
- Incorporation of the needs and specificities of urban logistics in the urban planning activity, joining the environmental and the urban planning sectors;
- Energy policies aimed at the optimization of the urban logistic system fostering the utilization of nonpolluting modes in the inner city short-range distribution services that represent a significant part of the urban logistics movements

From the analysis of the characteristics reflects in Tables 20.2a, 2b, 2c and 20.3 we have extracted five logistic profiles (TURBLOG, 2011) that have been confronted with the reality of several cities across the world and have successfully survived the test. Table 20.4 summarizes the description of these profiles.

Table 20.3: Matching of the logistic profiles with the appropriate delivery solution (Macário, Filipe, Reis, Martins, 2007b)

Delivery solution/Logistic profile description	**Hierarchic terminals with bicycle or other eco-friendly, light delivery service**	**Cooperative services**	**Safe deposit boxes**
Area profile: Lower commercial density and homogeneity, lower logistic accessibility; any hourly restrictions; Product profile: No special needs nor perishable; Delivery profile: Small amount, frequency and urgency irrelevant	+ + + + +	+	+ + + + +
Area profile: Lower commercial density and homogeneity, lower logistic accessibility; any hourly restrictions; Product profile: No special needs nor perishable; Delivery profile: **Big amounts**, frequency and urgency irrelevant	+ + +	+ +	**0**
Area profile: **Higher commercial density**, lower homogeneity, lower logistic accessibility; any hourly restrictions; Product profile: No special needs nor perishable; Delivery profile: Medium/big amounts, frequency and urgency irrelevant	+ +	+ +	**0**
Area profile: Higher commercial density and **homogeneity**, lower logistic accessibility; any hourly restrictions; Product profile: No special needs nor perishable; Delivery profile: Medium/big amounts, frequency and urgency irrelevant	+ +	+ + + +	**0**

Notes: Highlighted in bold are the main characteristics that change between profiles. The appropriateness of the services ranges from: + + + + + "Service highly suited for the profile" to 0 "Service not appropriate for this profile."

Table 20.4: Logistic profiles (TURBLOG, 2011)

1. City Area Features	Profile A	Profile B	Profile C	Profile D	Profile E
	Cluster of shops specialized in one specific type of service/product: ex. a neighborhood that is known for furniture stores, craft or art pieces, technological pole.	Hotels, restaurants, small grocery stores, neighborhood markets	Business center (courier, small deliveries, B2C)	Large commercial (retail, shopping centers, distribution warehouses)	Residential areas with local trade
1.1. Commercial density	High	Low/Medium/High	High	High	Low/Medium
1.2. Homogeneity	High	Low/Medium/High	Low	Low	Low/Medium
1.3. Logistic accessibility	Good/Reasonable	Bad/Reasonable/Good	Reasonable/Bad	Good	Reasonable/Bad
1.4. Restriction applied	Yes/no	Yes/No	Yes	No	Yes
2. Product Characteristics	**Profile A**	**Profile B**	**Profile C**	**Profile D**	**Profile E**
	Cluster of shops specialized in one specific type of service/product: ex. a neighborhood that is known for furniture stores, craft or art pieces, technological pole.	Hotels, restaurants, small grocery stores, neighborhood markets	Business center (courier, small deliveries, B2C)	Large commercial (retail, shopping centers, distribution warehouses)	Residential areas with local trade
2.1. Easiness of handling	Easy/Reasonable/ Difficult	Easy/Reasonable/ Difficult	Easy	Easy/reasonable/ Difficult	Easy/reasonable/ Difficult

Table 20.4: (*Continued*)

2. Product Characteristics	**Profile A**	**Profile B**	**Profile C**	**Profile D**	**Profile E**
2.2. Special conditions	No special needs/ special needs	Special needs	No special needs	Might have special needs	Might have special needs
2.2.1. Fragility	No special needs	Fragile	No special needs	No special needs	No special needs
2.2.2. Perishability	Not perishable	Perishable	Not perishable	Not perishable	Not perishable
3. Agent Profile/ Deliveries Profile	**Profile A**	**Profile B**	**Profile C**	**Profile D**	**Profile E**
	Cluster of shops specialized in one specific type of service/product: ex. a neighborhood that is known for furniture stores, craft or art pieces, technological pole.	Hotels, restaurants, small grocery stores, neighborhood markets	Business center (courier, small deliveries, B2C)	Large commercial (retail, shopping centers, distribution warehouses)	Residential areas with local trade
3.1. Urgency of deliveries	Irrelevant/Relevant/ Urgent	Urgent	Relevant/Urgent	Relevant	Irrelevant/Relevant/ Urgent
3.2. Frequency of deliveries	Low/Medium/High	High	High	Medium/High	Low/Medium
3.3. Amounts to be delivered	Few/Several/Many	Several	Few/Several	Many	Few/Several/Many
3.4. Planned deliveries	No defined routine/ Defined routine	Defined routine	No defined routine/ Defined routine	Defined routine	No defined routine

Grey areas – Features that are not considered relevant for definition of the logistics profile.

20.4. Applied Concept: Case Analysis

As previously stated, the third step of the ongoing research is to test the impacts achieved with the implementation of the different logistic solutions. After generally surveying Lisbon's commercial areas and trading practices, and assessing the availability of information databases, it was decided to choose as study area for this third step – modeling effects of implementation of logistic profiles-one of the traditional neighborhoods of Lisbon – "Alvalade," a planned area from the 1950s that reflects the concept of multifunctional area gathering residence from different social and economic segments, as well as services, leisure activities and a few industry cases.

Following the previous consideration, the solutions/services considered for the modeling exercise (later referred as scenario 2) were the following:

- Two hierarchic levels of freight terminals inside the urban territory, which might act in conjunction with freight yards located outside city boundaries; the higher level, bigger terminals, acting as freight hubs, collecting freight coming from the freight yards or directly from the producers, and distributing it to smaller level terminals, which shall be fully integrated within the urban tissue; the exact locations and sizes of these terminals is to be planned according to the different logistic needs of the areas they are to serve; these terminals are associated with delivery services responsible for the "last-mile" distribution; these shall be performed by "eco-friendly" vehicles, such as bicycles, electric tricycles and so on, according to the different LP needs;
- Cooperative services like car pooling and car sharing, which can be put together by groups of agents (e.g., shop owners) with the same delivery needs. Deliveries done this way will directly link any of the terminals mentioned above with the recipients' locations;
- Safe deposit boxes, in which medium sized lockers are placed on key points, like parking lots or freight terminals; these boxes work like normal postal boxes, in which parcels are dropped and then picked by the recipients (repair technicians, services or even end costumers); these boxes shall compliment other solutions to be implemented. They do not have the goal of reducing the number of deliveries, but to minimize the effects caused by the randomness of small and urgent deliveries, and to avoid the failure of such deliveries due to the absence of the final recipient;
- Collective and regular services using the facilities used by regular public transport in nonoperating hours (e.g., light rail, tram, or underground networks); none of these types of services was explored in the current work, due to the absence of such networks in the area chosen as case study.

From the conceptual point of view, the aforementioned services will form a network in which the nodes are the terminals (distribution centers) and the links are the transport services between them. The outset of the urban logistic system consists in defining and determining the levels of service in the different terminals, as well as its main characteristics in terms of localization, operation and capacity. These characteristics are largely determined by the specifications associated to each type of good and, consequently, to certain commercial activities.

We need to consider different logistic services, or combinations of services, aimed at serving different delivery needs. The services considered here arise not only from the literature review, but also from the observations done, namely of the experiences implemented in several cities.[4] They were selected in a way to serve the study area with the aim of creating services that work in an integrated way, so to cover all the logistic needs of the area being served.

20.4.1. The Case Study Area[5]

The ''Bairro de Alvalade'' is a multifunctional area, which planning started in the 1930s. One of the first and most interesting experiences in modern urban planning is the city of Lisbon. This neighborhood is a central and consolidated residential area, which has very intensive commercial streets, mixing traditional and modern shops, as some warehouses and a market; the many restaurants and cafés represent the major share of the commercial activity.

In order to briefly characterize the logistic activity on this neighborhood, an observation survey was done on its main avenue (Avenida da Igreja), in which the majority of shops is located. This was done by observing, for a whole week, the delivery of goods to three kinds of shops: home textiles, clothing & shoes shops, restaurants & cafes, and markets & small food shops. These clusters of activities were chosen due to their representativity of a large share of the shops within the neighborhood.

The data was then analyzed concerning deliveries per day/time period, per type of product and per type of shop. The graphics on Figures 20.2–20.5 show the main results of these analyses.

A general problem within the city of Lisbon is the use of traffic lanes for loading and unloading activities due to the lack of appropriate bays. This obstructs the normal traffic flow and poses difficulties and risks for the loading and unloading activities. Therefore, the duration and place in which these tasks were collected and analyzed. The results are shown in Table 20.5.

The analysis of the graphs and numbers above clearly reveals that shops dealing with food and related products are highly delivery demanding, and seem to be the ones causing the most of the problems associated with urban logistics. However, it was not possible to distinguish which of the following two factors justifies this: the type of business-small shops without storing facilities, which implies a high rotation

4. For more detail on cities and cases read TURBLOG reports.
5. A preliminary version of this study was developed in the LOGURB project (2007). Currently extensive surveys are ongoing (January 2013) in Lisbon to transfer the concept to downtown area and other parts of the city in parallel to in-depth academic research.

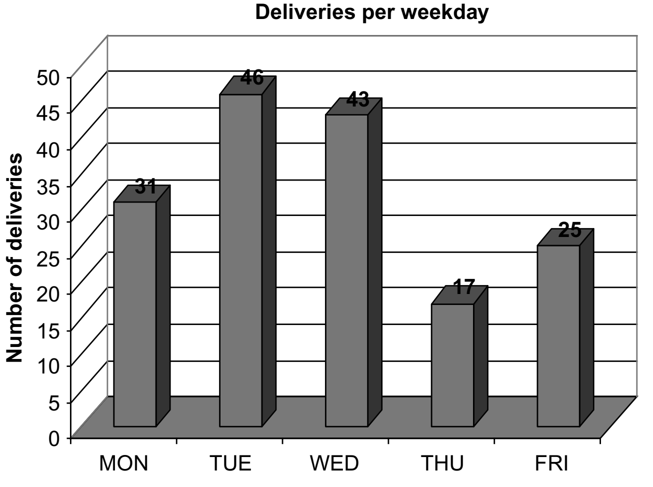

Figure 20.2: Deliveries per weekday (2007).

of stocks; or the specificities of the products-most of them being perishable and unable to be stored for long periods. A more detailed survey is needed to this segment to understand the contribution of the different attributes of the business to the stress of the urban logistic function.

20.4.2. Limitations of the Modeling Tools

The level of detail required for modeling the different logistic solutions is very high. So, the logistic schemes where only tested in a small area within ''Bairro de Alvalade,'' more specifically, the two blocks more commercially active.

The road network model was developed using a road inventory from GIS ArcInfo 9.2. The accessibility analysis was done using ArcInfo Network Analyst[6] extension that provided the capacity to develop network-based spatial analysis, such as routing, travel directions, closest facility, service area, and location-allocation. Using ArcGIS Network Analyst, we could model the realistic network conditions, including one-way streets, turn and height restrictions, speed limits, and variable travel speeds based on traffic. The road network characteristics were calibrated with more specialized software based on LoGit model approach after a thorough review of available models (Cascetta, 1995) that allowed the calculation and assignment of dynamic variation and equilibrium of flows, and other mobility key indicators, like speed and levels of service.

6. For more details on this software please, see www.esri.com

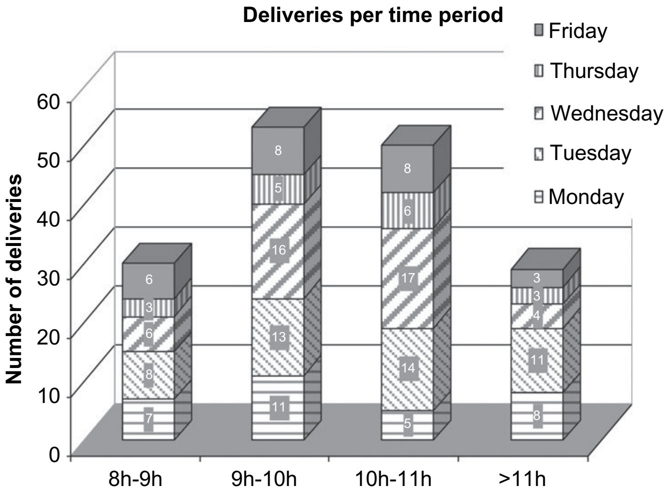

Figure 20.3: Deliveries per time period (2007).

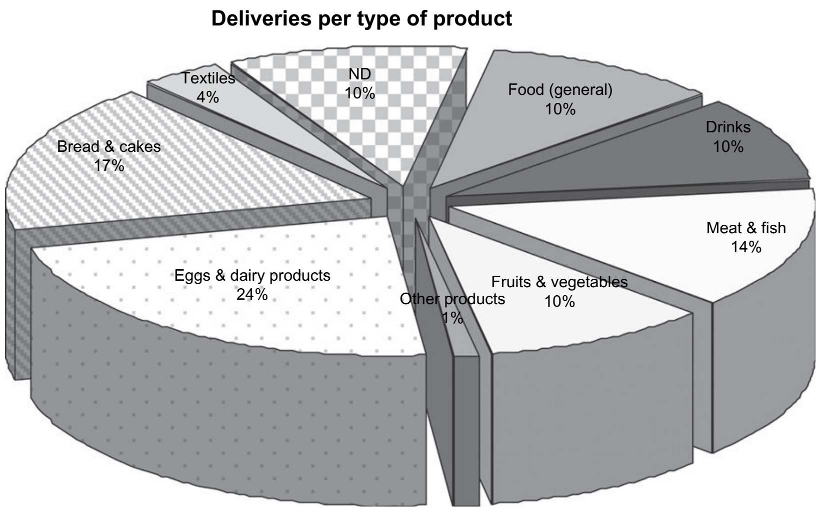

Figure 20.4: Deliveries per type of product (2007).

The path choice model was calibrated based on previous studies done for the city of Lisbon and interviews with drivers were made to validate model properties. Average travel times were computed considering the functions reported in the literature. In order to consider the level of service attribute vehicles were classified according to their dimension. The data was imported to the GIS database using ‘‘static’’ morning peak data. This allowed the simulation of morning peak ‘‘real’’ conditions, under the assumption of unchanged flows. This assumption is acceptable

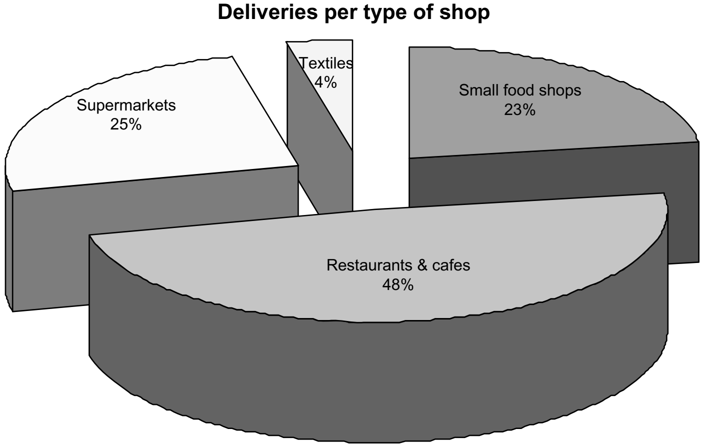

Figure 20.5: Deliveries per type of shop (2007).

Table 20.5: Number of deliveries according to the parking of the vehicles (2007)

Place of parking during loading and unloading activities	Number	%	Average parking time (seconds)
Parking lot/appropriate delivery places	50	31	15
Traffic lane	112	69	12
Total	162	100	13

for this case study, because the focus of analysis lies on the changes operated only over a small local area. The purpose was to provide exploratory evidence on the merit of the "3-step" approach proposed and open the door for a more extensive pilot simulation.

The analytical methodology proposed for urban freight distribution takes as a departure point the systemic character of mobility and its dependence from land-use as well as the subordination of mobility to urban socioeconomic dynamics. The underlying rational supports the view that urban freight is highly segmented and requires a diversified system organized according to a semihierarchical regime.

The aim of this work was not to obtain comprehensive O/D information of the urban flows, neither to provide a fully fledge organization of the urban freight systems, instead the objective is to show the potential effects the above expressed rational can have in an urban environment.

The development of a global urban freight model required a detailed definition of current transport network and the establishment of a zoning system to discretize the demand for movement of goods in space in the different urban contexts. The logistics profiles, at the root of this approach, are materialized by zones that coherently represent the land use, the space characteristics, and the product characteristics (i.e., land-use utility).

The exploratory analysis and modeling done, and the review of existing models allowed us to confirm what has been put forward by some other authors (Cascetta, 2001) that most of the urban freight models developed are not integrated in the other components of urban mobility (Macário, 2011) and so have no connection with measures taken at urban scale neither any relation with the passenger models that have been much more developed in the last decades. The immediate consequence is that such models are incapable of providing feedback or prospective capacity for decision makers.

For diverse reasons (Russo & Comi, 2004), urban freight models develop in a restocking logic (from warehouses to shop retailers) that does not consider neither the movement to end-consumer that represents a very significant part of urban logistics nor the restructuring of the chain with the introduction of intermediate terminals and the three tier process of distribution. The next steps of this research will focus on this two tier approach, from warehouses to intermediate terminals, from terminal to shop-retailer and, finally from shop-retailers to end-consumers. Each tier has different characteristics and requires a different type of model but in each tier a link must ensure the relation with urban measures and interaction with passenger. Only this complex architecture can assure the adequate feedback for policy making.

20.4.3. Results of Analysis

The analysis done consisted in comparing of two scenarios: today's Business As Usual situation (Scenario 1-BAU), and a future situation, in which the concepts are directly applied to the logistic system and simulated, in a way as close as possible to reality (scenario 2). For the implementation of both scenarios, the development of an extensive and complete inventory of local commerce on the two analyzed blocks was needed. A micro-distribution freight model was built over this inventory. This model was directly linked with the Lisbon road network, and with the main metropolitan area road network.

In order to simulate the diverse origins (generation) of the goods flowing into Alvalade, the analysis was further segmented into three geographical rings. The first one has been defined by the border of the neighborhood itself, and the "inner model" simulates the local impacts and changes at micro (street) level. The second ring is limited by the city boundaries, and includes Lisbon major inner road systems (e.g., a simplified simulation of the road network); it allows quantifying the impacts caused by goods entering Alvalade from other neighborhoods inside Lisbon. The third ring, outside the city boundaries, allows the analysis of changes achieved in the delivery of goods received directly from outside the metropolitan area.

One of the crucial points for this analysis was the construction of a solid and somehow significant characterization of the attraction of flows caused by the commercial activity within the neighborhood. Since no statistical information was found for any region in Portugal, concerning the attraction of commercial activity, the development of a methodology to determine those values was required. This methodology is depicted in Figure 20.6, and explained below.

For the first step, information from the aforementioned inventory was used (types of commerce within the study area and number of employees per store). The number of employees was used as a proxy for calculating the movements generated by each type of store. Finally, the total number of movements was obtained (Laboratoire d'Economie de Transport (LET), 2000).

The study area has fourteen segments of commerce and services. Table 20.5 presents those segments, along with the total number of employees and the total number of movements, calculated using the previous methodology. It is worth mentioning that the parameters given by the LET study do not follow a linear function in respect with the total number of employees; instead, they show a decreasing return: the higher the number of employees, the smaller the number of movements generated by each one.

Concerning scenario 2, it is assumed that the total number of movements per segment is the same as the current one. The only change is just the introduction of the new distribution solutions presented in previous chapter. Additionally, it was considered that less polluting vehicles were responsible for a share of the deliveries, alongside with vehicles using the present technology.

To keep this situation as close to reality as possible, it was considered that a significant share of the shop keepers would not be willing to change their current logistic solution. Thus, the minimum percentage of movements using road vehicles considered is 50%. The exceptions to this are bookstores and other retail stores where a higher penetration of the proposed new logistic solutions is expected. By applying the new distribution patterns to the different business segments the following Table 20.6 was produced with the estimated number of weekly movements, and later distributed by the different logistic solutions as presented in Table 20.7.

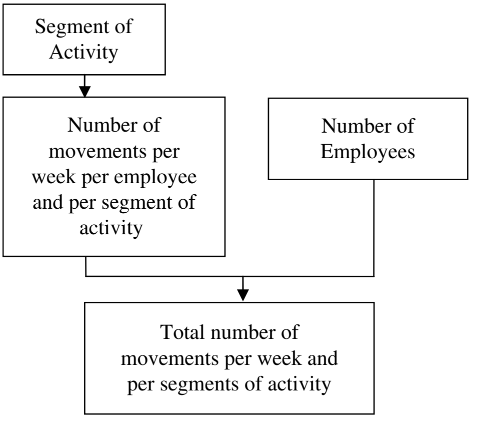

Figure 20.6: Methodological framework for the definition of logistic deliveries.

Table 20.6: Movements by commercial segments

Segment	Total number of employees	Total number of movements (per week)
Cafés, restaurants	30	61
Hardware stores	3	18
Pharmacies	8	65
Textile and shoe stores	36	71
Bookstores	3	39
Small markets	20	172
Market	19	104
Butcheries	14	49
Furniture stores	5	21
Normal services	20	12
High flow services	5	21
Other retail	27	117

Table 20.7: Total number of movements/week per segments and per logistic solution

Segment	Bicycle	Box	Car pooling	Sustainable cars	As usual
Cafés, restaurants	15.3	0.0	0.0	18.3	27.5
Hardware stores	3.6	2.7	0.0	3.6	8.1
Pharmacies	6.5	0.0	0.0	19.5	39.0
Textile and shoe stores	21.3	0.0	0.0	17.8	32.0
Bookstores	15.6	3.9	5.9	7.8	5.9
Small markets	34.4	0.0	51.6	25.8	60.2
Market	20.8	0.0	10.4	15.6	57.2
Butcheries	0.0	0.0	0.0	24.5	24.5
Furniture stores	4.2	2.1	0.0	4.2	10.5
Normal services	48	0.0	0.0	3.6	3.6
High flow services	8.4	0.0	0.0	6.3	6.3
Other retail	29.3	11.7	23.4	23.4	29.3

The work already done allows the comparison of both scenarios for the first (local) ring and shows some interesting conclusions. The results illustrate only changes related with the improvement in local efficiency. One can expect to achieve bigger efficiency gains with the concentration of goods and optimization of movements within a larger geographic area.

The implementation of scenario 1 (BAU situation) implies the overall modeling of movements for each commercial or service unit of the neighborhood (see Figure 20.7)

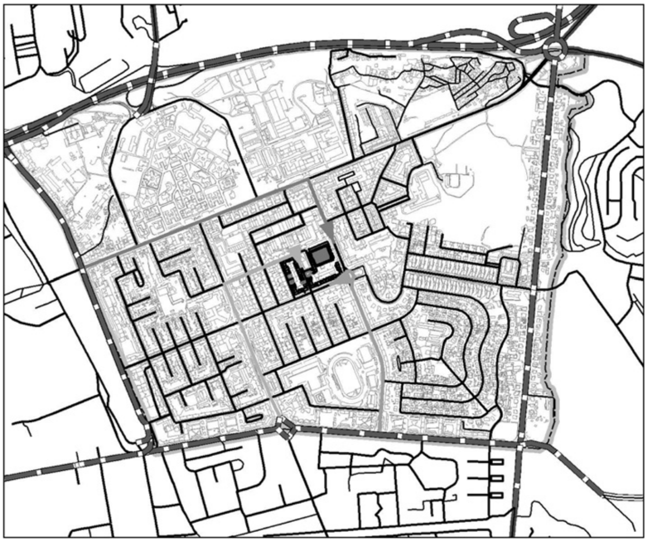

Figure 20.7: Operational details for scenario 1 using the First Ring model.

and the construction of a local O/D freight matrix. This allows the computation of local mileage and trip time. A simulation of the morning peak was used, which represents a "near congestion" situation.

The deliveries done inside the First Ring represent the consumption of 356 weekly hours of average trucks (with a driver and sometimes an assistant) in the neighborhood's leg of the delivery chain, and on the activities related with unloading and delivering. The total mileage done inside this ring is about 1650 km (per week). A rough estimate of the annual impact can be done, assuming an annual equivalence factor of 52 weeks. This implies a total of 18,500 truck hours per year and a mileage of more than 85,000 km of impacts only inside the First Ring. The assessment of the impacts of the model in a larger geographic area certainly will represent higher costs due to the caption of the overall urban transport production (mileage and trip time), which is somehow minimized in the context of the analysis made only for the area of Alvalade.

Scenario 2 implies the existence of a Local Terminal for the distribution of goods using the new delivery solutions previously appointed. A preliminary analysis of the consolidated area revealed two candidate places for its location, both of which were further analyzed. Concerning the availability of space, the local impacts generated by this terminal (quality of accesses, emerging freight traffic, etc.) and the desired proximity with the commercial blocks, the choice for the terminal location was an area near the Market, as shown on Figure 20.8.

The analysis using the Local Terminal implies a twofold and two step deployment of the freight assignment submodel. First, the freight trips were divided in two distinct groups: one in which all the unchanged trips will remain, and a second one in which the movements using new solutions are contained. The first group will have no

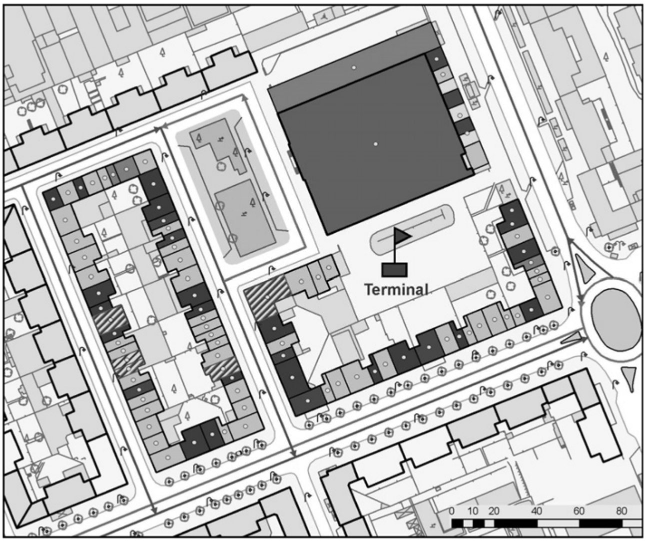

Figure 20.8: Location of the local terminal for scenario 2.

change from scenario 1. The second group is going to feed the "two steps" implementation developed using the terminal. This may be seen as an inter-modal trip chain, in which the freight mode changes, but with a more complex situation. The generation factors do not remain constant and the capacity of the final end delivery mode is obviously lower, and there is a need for more "end trips."

The comparison of results between the two scenarios show an increase of around 7% (about an extra 35 truck hours per week) for the time spent in the delivery chain. Besides, there is a mileage reduction of about 17% (less than 277 km per week). A first conclusion can be directly taken from these figures: the local environmental impacts are undoubtedly positive. The mileage reduction of 17% in Alvalade is an important factor for the reduction in air pollution, noise, accidents and other local urban impairments.

It is expected that the results to be achieved with the model using two larger rings-including larger areas of Lisbon city and the whole metropolitan area- will invert the operational balance. The benefits obtained from the larger reduction in mileage and environmental impacts will overcome the increase in truck operational cost. In that case, the path for the assessment of economical feasibility and autonomy for the terminals and for setting the new delivery schemes is fulfilled.

20.5. Decision-Making Concerns: Need for a Driver

The case presented confirms the potential to generate benefits that can result from the adoption of a logistic profile concept. But it confirms also the need to have initiative in the introduction of policies and measures as well as innovative solutions to overcome space limitations.

From the cities observed we conclude that urban logistics systems suffer from a defensive position assumed by the main stakeholders in the process. Decision makers and authorities do not take the risk of accepting any responsibility in urban logistics. Their approach to the problem tends to remain in the restriction stage. Operators in turn avoid interference from authorities in their businesses, which is historically undertaken by private entrepreneurs, and refrain from being involved in solutions where responsibilities and functions are not clearly stated from the outset. They welcome only a facilitator role from the authorities. There is thus a consensus that relative positions of these stakeholders must be kept.

From the analysis of different examples of successful urban logistic measures (TURBLOG, 2011) it is possible to conclude that most of the innovative business concepts presented rely on partnerships other than the typical buyer-supplier relationship, with the expectation to improve performance (efficiency) and accessibility of their services as core value propositions. Moreover, some business concepts were only effectively implemented because they were sustained by public administration policies, which provided availability of warehouse spaces or accessibilities and, in some cases, financial incentives, resulting in partnerships with the municipality or other government administrations (e.g., Monoprix, Chronopost, La Petite Reine). In order to meet the municipal environmental requirements and restrictions and also looking towards improving service performance, some companies developed joint ventures to develop these new services (e.g., La Petite Reine, that developed the tricycle needed for its business with a local manufacturer, or the Beer Boat that is operating in the City of Utrecht).

The analysis done is based on the idea that the most suitable logistic solution is defined not only by the business characteristics but also by the delivery, the product and city area features (logistic profile), as well as the policies adopted/to be adopted for the city. It is the combination of these three pillars that constitute the backbone of the decision making for best urban logistics solutions. In most cities the key requirements are:

- Integrated approach to land use and urban living: first, last and transit mile;
- Tailored path dependent approach (each city has its own evolutionary path);
- Private initiative with collaborative solutions;
- Low public intervention but public policies acting as facilitators for private entrepreneurship and innovation;
- High technical standards (including environmental and energy);
- Innovation to solve complex problems.

In the end the process is expected to identify the most suited practice for a given reality considering the logistic profile, the business models and the expected impacts. Four different "urban logistic tools" must be put in place to facilitate entrepreneurs in the process of innovation that is required to conceive urban logistic services that bring added value to the supply chain:

- Perpetual information data base on activities and respective logistic needs;
- Strategic evaluation for the organization of urban logistic activities. This tool provides key stakeholders (public and private) with information about how logistics activities should be organized and the impacts that they might have;
- Delivering and Servicing Plan. This tool aims to optimize deliveries, and thus decrease the number of trips and reduce the impacts of urban freight movements;
- Loading and unloading regulations. This tool aims a set of measures including parking regulations and restrictions on movement of heavy goods vehicles (time frames, truck routs, vehicles weight, and size restrictions), in order to regulate the operations of loading and unloading.

20.6. Innovation as a Development Driver

The results gathered so far allow to conclude that the new urban logistic schemes presented and tested within the research are relevant for the design of a future master plan for urban logistics. We have made evidence that there are effective gains in the adoption of the LP concept with logistic solution addressing the different market segments. The implementation of such a plan must be promoted with the support of local authorities to guarantee its economic and social feasibility.

In parallel, the discussion about the "business plan" of these new urban logistic nodes and their juridical and legal integration must be defined. Several questions need further attention and research work:

- How to organize the funding to launch these centers until they reach acceptable economic thresholds?
- Are these suitable for adoption of public–private partnerships or even public–public partnerships?
- What is the more effective regulatory and organizational framework to stimulate private entrepreneurship in the implementation and development of these centers?
- What are the critical dimensions to support a network of logistic centers serving and urban system?
- Who are the winners and losers if a more extensive implementation of the LP concept is done and how can we mitigate those losses?

As North (1990) stated "The speed of economic change is a function of the rate of learning, but the direction of that change is a function of the expected payoffs to acquiring different kinds of knowledge." Innovation can act as a driver of this process but we need to ensure its basic pillars, being: the institutional base setting for innovation (i.e., human capital, rules); interactive processes (i.e., evolutionary processes considering path-dependent tracks); monitoring and learning structures and processes (i.e., assessing short-term outcomes and long-term impacts.

Our knowledge of the distribution system is a precondition to model innovative solutions to the benefit of the different stakeholders. The existing modeling approach despite the boost received in the last half decade still falls very short of what decision makers need to enable them a good quality of decision and policy making. As stated by Manheim (1979) "there is no rational objective way of deciding what is the best model for a particular application." The challenge is to make the real world use the transport system in the most effective and efficient way and this requires looking at the system with user eyes. That is in an holistic way, considering technical and managerial constraints as well as opportunities, and reflecting the rational of the different agents in the design of the analytical tools and processes used. Models must reflect reality as reality is never able to fit into models the more sophisticated they can be.

No doubt that there is still a long path to be developed to level up urban freight distribution with passengers. But a major milestone is the recognition that urban freight distribution is certainly a key condition for the good performance of an urban mobility system and, consequently, for the competitiveness of the city. This irreversibly opens the door to new research. We just need to be bold enough to step back, rethink the methods and processes through which we have been supporting decision making, and chose a more value-laden approach to improve the quality of decision.

References

BESTUFS. (2002). Land use planning and business models for urban distribution *centres*, BESTUFS workshop, 25-26 April 2002, La Rochelle, France. Retrieved from http://www.bestufs.net/

BESTUFS. (2008). *Deliverable D5.2 quantification of urban freight transport effects II*. http://www.bestufs.net/bestufs2_quantific.html

Cascetta, E. (1995). The Italian decision support system for transportation policies and investments: General architecture and development status. *Proceeding of The World Conference on Transportation Research*, Sydney.

Cascetta, E. (2001). *Transportation systems engineering: Theory and methods*. Boston, MA: Kluwer Academic.

CEC (Commission of the European Communities). (2007a). Communication from the Commission: – Freight Transport Logistics Action Plan COM(2007) 607.

CEC (Commission of the European Communities). (2007b). Communication from the Commission: – The EU's freight transport agenda: Boosting the efficiency, integration and sustainability of freight transport in Europe COM(2007) 606.

DETR. (1999). Department of the Environment, Transport and the Regions (DETR) Sustainable Distribution: A Strategy, DETR.

DfT (Department for Transport). (2003). A guide on how to set up and run freight quality partnerships, transport energy good practice guide 335, DfT. Retrieved from http://www.freight.dft.gov.uk/gpg355/pdf/gpg335-final.pdf

ECOTEC Research and Consulting Ltd, in cooperation with NordRegio and Eurofutures. (2007). The 'State of European Cities Report', within European Urban Audit, which is coordinated by Eurostat with National Statistical Offices.

FHA, Federal Highway Administration. (2009). —*Freight Facts and Figures 2009*. Retrieved from http://ops.fhwa.dot.gov/freight/freight_analysis/nat_freight_stats/docs/09factsfigures/index.htm

Freight Transport Association. (1998). Delivering the goods, FTA, Tunbridge Wells.

Hensher, D., & Puckett, S. (2005). Refocusing the modelling of freight distribution: Development of an economic-based framework supply chain behaviour in response to congestion charging. *Transportation*, *32*, 573–602.

Laboratoire d'Economie de Transport. (2000). Diagnostic du Transport de Marchandises dans une agglomération. Programme National Marchandises en Ville. Direction de la Recherche et des Affaires Scientifiques et Techniques. Paris.

Macário R. et al. (2007a). Logurb-Optimização de Sistemas Logísticos de Distribuição de Mercadorias em Meio Urbano-state of the art da logistica urbana, Fundação de Ciência e Tecnologia, Lisboa, Portugal.

Macário, R. (2011). Managing urban mobility systems. ISBN 9780857246110. Bingley, UK: Emerald Group Publishing Limited.

Macário, R., & Caiado, G. (2004). The Portuguese experience: Lisbon. In G. Ambrosino, M. Boero & M. Romanazzo (Eds.), *Policies and technologies for sustainable freight distribution in urban area*: *Towards the eLogistics Agency*. Lisboa, Portugal: ENEA Publication.

Macário, R., Filipe, L. N., Reis, V., & Martins, P. (2007b). Elements for a master plan in urban logistics. In E. Taniguchi & R. G. Thompson (Eds.), *Proceedings of City Logistics 2007, Institute for City Logistics*. Oxford, UK: Elsevier Publication.

Manheim, M. (1979). *Fundamentals of transportation system analysis, volume 1: Basic concepts*. Cambridge, MA: MIT Press.

Muñuzuri, J., Larrañeta, L., Onieve, L. e., & Cortés, P. (2005). Solutions applicable by local administrations for urban logistics improvement, Cities, *22*(1), 15–28.

North, D. (1990). *Institutions, institutional change and economic performance*. Cambridge, UK: Cambridge University Press.

OECD-Organisation for Economic Co-operation and Development Delivering the Goods. (2003). 21st Century Challenges to Urban Goods Transport.

Osterwalder, A. (2004). The business model ontology – a proposition in a design science approach, PhD thesis, École des Hautes Etudes Commerciales de l'université de Lausanne.

Rodrigue, J. (2006). Freight and the city: An overview of urban freight distribution and city logistics, Maritime Infrastructure Engineering and Management Program, Rutgers University, Newark, NJ, April 2006.

Russo, F., & Comi, A. (2004). A modelling system to link end-consumer and distributions logistics, European Transport No. 28, pp 6–19, ISTIEE, Trieste, Italy.

Stantchev, D., & Whiteing, A. E. (2006). *Urban freight transport and logistics: An overview of the European research and policy*. TRKC Policy Brochure. Retrieved from http://www.transport-research.info/Upload/Documents/200608/20060831_105348_30339_Urban_freight.pdf

START. (2008). *START project final report: Future solutions for goods distribution*. Retrieved from http://www.start-project.org/download/final/START%20final%20Report.pdf

Taniguchi, E., & van der Heiden, R. (2000). An evaluation methodology for city logistics. *Transport Reviews*, *20*(1), 65–90.

TURBLOG. (2011). *Transferability of urban logistics concepts and practices from a worldwide perspective, 7FP Research Program of the European Commission*. Retrieved from www.turblog.eu

Visser, J., van Binsbergen, A., & Nemoto, T. (1999). Urban freight transport policy and planning. Paper for first international symposium on city logistics Cairns, Australia. Retrieved from http://www.rstrail.nl/FTAM/FTAM_papers/FTAM2/transportpolicy.pdf

Chapter 21

Tactical and Operational City Logistics: Freight Vehicle Flow Modelling

Agostino Nuzzolo and Antonio Comi

Abstract

This chapter first classifies city logistics measures that city administrators can use to reduce the negative impacts of urban freight transport, in relation to planning (i.e. strategic and tactical/operational). The focus then shifts onto models developed to support the assessment of tactical and operational measures. The assessment procedures require the simulation of freight transport demand and hence the estimation of freight vehicle origin-destination (O-D) flows. These O-D flows can be obtained from the simulation of delivery tours. Therefore, the chapter presents a system of models able to simulate delivery tours using an aggregate approach. Such models allow us to capture actors' choices which can be influenced by tactical and operational measures. They were calibrated and tested using the surveys carried out in the inner area of Rome, where more than 500 truck drivers were interviewed. Some results of the validation of the models are also presented.

Keywords: City logistics; urban freight transport; freight vehicle flows; delivery tour simulation; aggregate model

21.1. Introduction

Urban freight transport plays an essential role in meeting the needs of citizens, but at the same time it significantly contributes to non-sustainable effects on the environment, economy and society. As would be expected, local administrators use city logistics measures to reduce the above negative impacts. Indeed, several city

Freight Transport Modelling

ISBN: 978-1-78190-285-1

logistics measures can be implemented with a view to reducing the negative effects of freight distribution, seeking to reduce the number of commercial vehicles, increase the use of light and environment-friendly vehicles and optimize loading and unloading operations in order to reduce traffic congestion and interference with the other components of urban mobility (e.g. pedestrians).

Such city logistics measures can be classified in relation to their planning level as operational, tactical and strategic. At the *operational* level, they concern freight demand management (e.g. time windows for access, loading/unloading areas), telematics (e.g. using intelligent transportation systems - ITS), while at the *tactical* level they mainly concern urban freight distribution through, for example, the implementation of nearby delivery areas (NDA), or incentives to switch to third-party (3P) vehicles or financial incentives to renew the delivery fleet. At the *strategic* level, the implementable measures concern urban freight distribution organization (i.e. a two-tier system with public urban distribution centres and/or transit points), improvements in cooperation among actors and land-use transformation governance.

Besides, city logistics planners should consider that there are no cities with exactly the same conditions (Ibeas, Moura, Nuzzolo, & Comi, 2012). Cities differ in many aspects of transport and traffic conditions (e.g. demand, supply, infrastructure, traffic control and management), and other factors including geographical, environmental, demographic and socio-economic conditions, cultural background and institutional and legal frameworks. In most cases, it will also be necessary to liaise with specific users and stakeholders, and ensure their support. Therefore, it is a challenging task to make sure that success in one city can be replicated in another (Jeffery, 2011).

For example, the increasing availability of data from cities indicates that the characteristics of urban goods movements mainly depend on the size of urban area. As city size increases, so does the length of delivery tours. In this case, measures that provide two-tier systems should be preferred, such as the creation of city distribution centres or transit points. Furthermore, in large areas, the share of 3P transport is also higher because in these areas there is a greater, ever-increasing presence of franchise stores, market chains and large-scale retail trade. Therefore, in small centres, measures that push towards 3P have to be implemented.

Since the characteristics of urban areas can differ substantially, city logistics measures have to be specifically designed and assessed in order to implement the most effective. In this process, a key role is played by demand models as they can be used to assess the effects of the scenario to be implemented. The models for scenario assessment thus have to investigate the variables that can play an important role in successful scenario implementation (Stathopoulos et al., 2011). For example, while all measures may be expected to perform well in terms of reducing the external costs of transport, some might increase internal transport costs incurred by some freight actors (e.g. carriers and hence wholesalers). Therefore, in recent years researchers have sought to develop models that support city logistics assessment (Comi, Delle Site, Filippi, & Nuzzolo, 2012). In this chapter, one such model is presented.

The chapter is structured as follows. In Section 21.2, the tactical/operational city logistics strategies and measures are briefly analysed, with assessment procedures and effects to be evaluated. In Section 21.3, the focus is on models developed to assess such city logistics measures. Given that freight vehicle origin-destination (O-D) flows are necessary and can be estimated from the simulation of delivery tours, a system of models able to simulate delivery tours, calibrated using surveys carried out in Rome, is proposed in Section 21.4. In Section 21.5, some conclusions are drawn on the performance of the model system presented and on its future possible developments.

21.2. City Logistics Strategies, Measures and Assessment

21.2.1. Strategies and Tactical/Operational Measures

Several measures have been implemented by cities to make urban mobility more sustainable and reduce the environmental impacts of freight transport due to commercial restocking in urban areas (BESTUFS, 2007; Russo & Comi, 2011). In implementing such measures, local administrators aim to:

- reduce the number of commercial vehicles on the roads;
- reduce the distance covered by heavy goods vehicles (HGVs) from warehouses and distribution centres;
- increase the share of 3P transport;
- increase the use of light and environment-friendly vehicles;
- optimize loading and unloading operations in order to reduce traffic congestion;
- reduce interference with other urban mobility components (e.g. pedestrians).

Focusing on tactical/operational planning level, the available short-term measures can be grouped into: freight traffic management, telematics and urban freight distribution organization.

Freight traffic management refers to the application of measures such as

- access time windows for limited traffic zones (LTZs) or pedestrian roads;
- reservation of loading/unloading areas for freight operations;
- vehicle type access constraints (e.g. on pollutant emissions and weight);
- route restrictions, consisting of definition of sub-networks only for trucks, or permission to use bus lanes;
- area-pricing, consisting of access charging (e.g. in LTZs);
- incentives, consisting of creating exceptions to the previous measures in order to incentivize the use of 3P transport, which should be more efficient than transport on own account, or the use of less polluting vehicles.

Telematics refers to ITS, mainly used to make enforcement more effective and broader in scope and improve delivery efficiency.

O*rganization of urban freight distribution*, at this planning horizon, concerns actions that optimize deliveries in relation to vehicle characteristics, so that

Table 21.1: Synopsis of tactical/operational city logistics measures

Planning Levels	**Strategies**	**Measures**
Operational	Interference reduction	Time windows
Operational	Interference reduction/loading-unloading optimization	Loading/unloading areas
Operational	Use of better performing vehicles	Vehicle type access constraints
Operational	Interference reduction	Route restrictions
Operational	Reduction in vehicle number	Area pricing
Tactical	Use of better performing vehicles	Incentives
Tactical	Reduction in vehicle number	Access control (telematics)
Tactical	Reduction in vehicle number	Nearby delivery area

externalities can be reduced in inner cities (e.g. city centres). Examples are given by *NDA*, consisting of urban transhipment platforms on which dedicated staff provide assistance for the dispatch of consignments for the last mile by trolleys, carts, electric vehicles and bicycles. Table 21.1 summarizes the relationships among strategies, measures and planning levels.

21.2.2. Measure Assessment and Freight Vehicle Flows

A new measure implementation process consists of several steps, which are able to:

- reveal the current critical issues through specific surveys (e.g. traffic counts, interviewees with retailers, truck drivers);
- define models to simulate the current scenario and assess the future one;
- share objectives and find an optimal set of measures that is a good compromise among the various actors involved;
- assess *ex ante* the new scenario by estimating effects and system performances, and comparing them with a set of given target values.

Therefore, an assessment methodology should consist of several stages able to point out different types of effects that can be classified as

- *internal*, that is, on the users of the system, such as retailers, wholesalers, distributors, carriers; for example, transport cost variations;
- *external*, that is, on citizens not directly involved in using the system; for example, pollutant emissions, noise, road accidents.

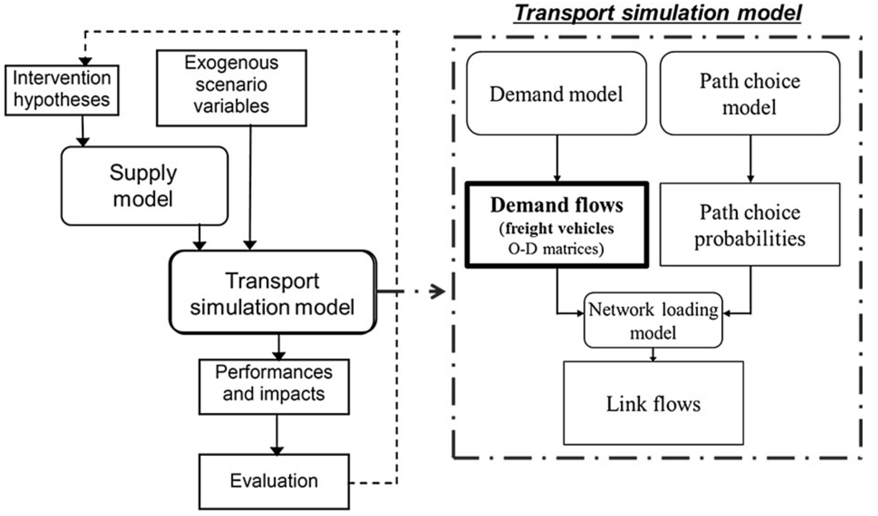

Figure 21.1: Effect assessment tool scheme.

Furthermore, within each of these two classes there may be:

- *direct* effects that are variations in transportation system costs;
- *indirect* effects, mainly cost variations induced by transport cost modifications, related to the economic and social sphere or to land-use activity sites.

This chapter refers to models that can be used to capture both the *internal and external direct effects* in relation to tactical/operational city logistics measures. The other effects, which require different specific models, will not be considered here.

In this context, the choice of a set of city logistics measures to be implemented (i.e. city logistics scenarios) has to be based on design tools by simulating the main effects of exogenously specified scenario (i.e. *what if* approach). In these tools, the freight vehicle O-D matrices, interacting within the assignment model, allow us to obtain the link freight vehicle flows and hence to estimate and evaluate link performance and the direct effects of a given city logistics scenario (Figure 21.1). As detailed in the next section, different models for the estimation of freight vehicle O-D flows have been developed.

21.3. Models for Freight Vehicle O-D Flow Estimation

Different types of demand models have been proposed to indicate the two sets of freight restocking flows taking place in a study area (e.g. from wholesalers to retailers or end-consumers): the O-D flows of *commodities* and the O-D flows of *commercial/freight vehicles*. The current literature has mainly investigated the former, and models for the estimation of the level and the spatial distribution of commodity exchanges, in terms of quantity and/or O-D delivery matrices, have been proposed (Comi et al., 2012; de Jong, Vierth, Tavasszy, & Ben-Akiva, 2012). Quantity-based models simulate the mechanism underlying the generation of freight transport demand (Gonzalez-Feliu, Toilier, & Routhier, 2010; Russo & Comi, 2010), while those using delivery units are more specific for studying the logistic process of restocking (Muñuzuri, Cortés, Onieva, & Guadix, 2012; Routhier & Toilier, 2007). Furthermore, solutions combining quantity and delivery units, and hence their advantages, have been developed (Nuzzolo, Crisalli, & Comi, 2012a).

In the case of tactical/operational measures, given the freight quantity O-D matrices, the choices that can be influenced concern the type of service (own account or using 3P vehicle), shipment size and hence number of deliveries, and delivery tour characteristics (departure time and vehicle type, number of deliveries and their sequence for each tour). Starting from the quantity of O-D flows and using models that allow characterization of O-D flow quantities by transport service type (i.e. transport service type model) and by quantity delivered (i.e. shipment size model), the delivery O-D matrices can be obtained (see e.g. Nuzzolo & Comi, 2013).

Following the aim of the chapter, in the next sections, the estimation of freight vehicle O-D flows, starting from the delivery O-D flows (Figure 21.2), is analysed in depth.

21.3.1. Freight Vehicle O-D Flow Estimation

As stated above, freight vehicle O-D flows can be estimated from delivery O-D matrices. The translation is not direct, particularly in urban areas where freight vehicles undertake complex routing patterns involving trip chains (tours). In fact, each restocker jointly chooses the number and the location of deliveries for each tour and hence defines his/her tours, trying to reduce the related costs (e.g. using routing algorithm). As pictured in Figure 21.3, the freight vehicle O-D matrices can be

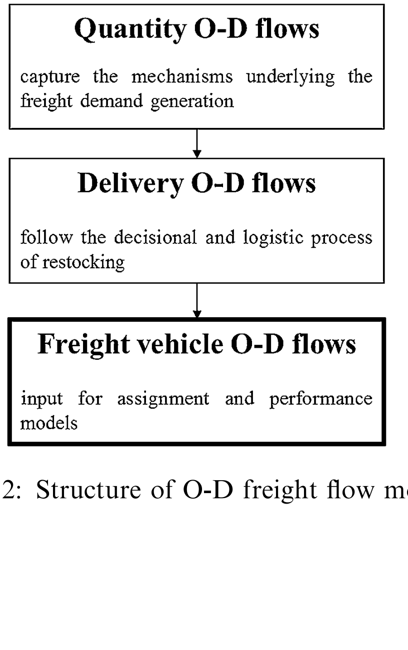

Figure 21.2: Structure of O-D freight flow modelling.

Figure 21.3: Structure of freight vehicle O-D flow modelling.

obtained using a two-step procedure: definition of delivery tours from delivery O-D matrices through delivery tour models, and definition of the freight vehicle O-D matrices from the delivery tours.

21.3.2. Delivery Tour Model Literature

Various types of models have been proposed to define delivery tours from known delivery O-D flows. The classification proposed by Wang and Holguín-Veras (2008) can be extended and subdivided into simulation and analytical models.

The *simulation models* are based on empirical relationships that allow the single trip to be combined into a tour. These models are usually implemented into an *ad hoc* Decision Support System (Ambrosini, Meimbresse, Routhier, & Sonntag, 2008; Comi et al., 2012), such as Wiver/Viseva (Lohse, 2004; Sonntag, 1985) and Freturb (Routhier & Toilier, 2007).

The *analytical* models propose mathematical relationships based on economic, statistical or behavioural assumptions within a disaggregate or aggregate approach. The *disaggregate* approach involves the use of procedures that include estimation of the delivery tours for each decision maker (e.g. carrier) with different locations of warehouses and shops to restock, different vehicles and time constraints to respect. Several principles are used to accomplish this, leading to models based on vehicle routing (Crainic, Ricciardi, & Storchi, 2009; Tamagawa, Taniguchi, & Yamada, 2010), behavioural (Ruan, Lin, & Kawamura, 2012) or activity models (Gliebe, Cohen, & Hunt, 2007) and on heuristic procedure (Figliozzi, 2006; Wisetjindawat, Sano, Matsumoto, & Raothanachonkun, 2007). However, in spite of their significant potential, implementation of these types of models has a fundamental limitation related to the amount of information required (which can only be obtained through specific large surveys) and to the expansion to the universe.

Aggregate models can be used as an alternative way to forecast urban delivery tours, given their smaller data requirements, lower computational times and less reliance on behavioural assumptions. These models consider the average behaviour of all restockers (or categories of restockers) leaving from the same warehouse zone. Within these common characteristics, three main groups of models have been proposed, as well as different specifications within each group depending on the variables explicitly simulated.

Models of the first group, known as *spatial price equilibrium models (SPE)*, are based on Samuelson's work (1952) and propose a multiple vehicle routing problem in which profit is maximized subject to competition (Thorson, Holguín-Veras, & Mitchell, 2005).

The second group of models is based on *entropy maximization theory* (Wilson, 1969). The resulting entropy formulations are aimed to find the most likely set of tour flows that meet the system's constraints such as the trips generated or attracted by each zone and the travel cost (Wang & Holguín-Veras, 2009).

The third group consists of *partial share models*. They give the probabilities that a delivery tour has a given number of stops, a given sequence of stops/deliveries and a given type of vehicle used. Within this group, two approaches have been proposed: incremental growth and multi-step.

Incremental growth studies (Hunt & Stefan, 2007; Wang & Holguín-Veras, 2008) propose to obtain, for a given type of vehicle, the number of stops per tour by incremental growth for which, at each stop, the option to come back to the base (i.e. warehouse) is considered. If the tour continues, the probability of the next destination zone is calculated (Holguín-Veras & Thorson, 2003). This approach implies major approximations because the actual choice process is not reproduced. Indeed, the choices of the number of stops and delivery zone sequence are generally pre-trip choices.

The *multi-step* approach (Nuzzolo, Crisalli, & Comi, 2011, 2012b) defines tours through joint definition of the trip chain order (that is the number of stops in a tour), the type of vehicle used and the delivery location sequence. In Section 21.4, a model of this type will be analysed and some recent advancements on calibration will also be presented.

21.3.3. Freight Vehicle O-D Matrix Definition

As each tour is undertaken by one vehicle, we can consider the sequence of deliveries as the sequence of trips belonging to the same vehicle trip-chain (tour). Once all delivery tours have been estimated, the freight vehicle O-D flows can be obtained through the aggregation of the trips of the tours, as reported in Section 21.4.

21.4. A Delivery Tour Model

This section first introduces the structure of a delivery tour model developed by the authors. Then the calibration and validation are discussed. Finally, model performance and the results of test applications are analysed.

21.4.1. Structure of the Proposed Urban Delivery Tour Model

The freight system actors considered in this chapter comprise carriers and wholesalers/distributors that transport freight on their own account (i.e. using their own vehicles). Their delivery tour choices, which can be modified by tactical/operational measures, are mainly related to tour stop/delivery number, type of vehicle to use for restocking and delivery tour to follow.

Let ND_{od} be the generic element of a given delivery O-D matrix representing the average number of deliveries departing from warehouse zone *o* and with destination zone *d*. The freight vehicle O-D matrices, satisfying the given delivery O-D matrix, can then be estimated by using an aggregate multi-step delivery tour model that considers the average behaviour of all restockers starting from the same warehouse zone (Figure 21.4).

The total number of tours T_o departing from zone *o* can be determined as follows:

$$T_o = \sum_{d'} ND_{od'} \Big/ \overline{n}_o \tag{21.1}$$

where $\overline{n}_o$ is the average number of deliveries performed by tours departing from zone *o*.

Let $p[n/o]$ be the probability that a tour departing from origin zone *o* has *n* stops/deliveries obtained by a *trip chain order model*. Therefore, $\overline{n}_o$ can be estimated as

$$\overline{n}_o = \sum_n n \cdot p[n/o] \tag{21.2}$$

Let $p[v/no]$ be the probability of using a vehicle type *v*, obtained by a *vehicle type model*. Therefore, the number of tours with *n* stops/deliveries departing from origin zone *o* and operated by vehicle type *v*, $T_o[vn]$, can be obtained as

$$T_o[vn] = T_o \cdot p[nv/o] = T_o \cdot p[n/o] \cdot p[v/no] \tag{21.3}$$

Let $p[d_j^{k+1}/d_i^k\ vno]$ be the probability of delivering in zone d_j the delivery ($k + 1$), conditional upon having previously delivered in zone d_i the delivery *k*, within a tour

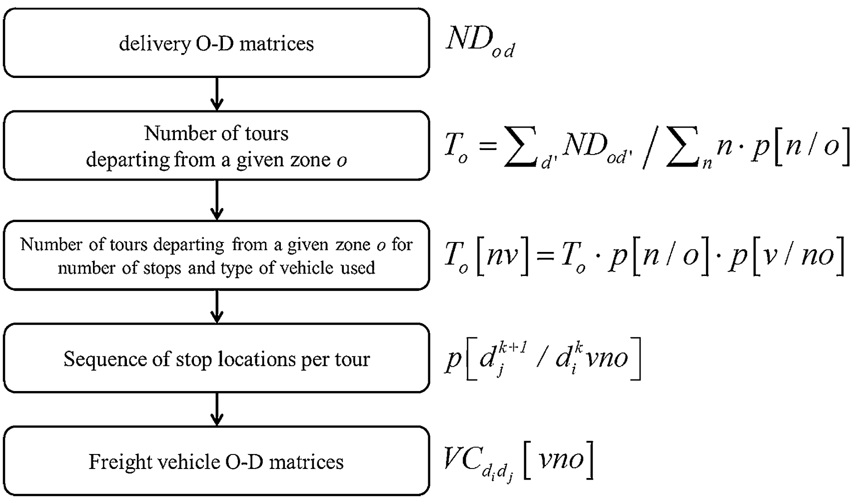

Figure 21.4: Urban delivery tour modelling.

with n stops/deliveries departing from a given zone o and using a vehicle type v, obtained by a *delivery location choice model*.

Finally, the number of vehicles $VC_{d_i d_j}$ on (d_i d_j) pair can be estimated as follows:

$$VC_{d_i d_j}[vno] = \sum\nolimits_k VC_{d_i^{k+1} d_j^k}[vno] = T_o[vn] \cdot \sum\nolimits_k p\left[d_j^{k+1}/d_i^k\, vno\right] \quad (21.4)$$

The probabilities $p[n/o]$, $p[v/no]$, $p[d_j^{k+1}/d_i^k\ vno]$ can be obtained by random utility models calibrated from survey data, as will be seen in the next sections.

21.4.2. Survey Data

Data used for this study came from surveys carried out in the inner area of Rome in 2008. This area is the historic and the most famous zone in the city, with the main historic monuments (i.e. the Colosseum) and many shopping streets reserved for pedestrians, and subjected to specific freight traffic regulations. In order to model the delivery tours, about 500 truck driver interviews were used. The driver questionnaire consisted of several sections designed to capture the characteristics of the interviewee and transport firm as well as the characteristics of transport (e.g. own account or third party, vehicle type, delivery location sequence).

It emerged that the most common vehicle type used was light goods vehicle (LGV; maximum gross laden weight less than 1.5 tonnes) characterized by an average of 2.06 deliveries per tour, close to the average number of medium goods vehicles (MGV; maximum gross laden weight less than 3.5 tonnes) which is 2.22 deliveries per tour. Besides, the survey results also evidenced that the number of stops/deliveries is related to the freight types and quantity delivered. The average delivered quantity varies from about 0.5 tonnes per retailer on own account to 0.35 tonnes for carriers, justified by the fact that retailers on own account generally made a lower number of stops/deliveries. In relation to freight types, the high revealed shares to use LGVs refer to foodstuffs, and household and personal hygiene products (daily consumption products), which amount to more than 70%. For more details on survey results, refer to Nuzzolo and Comi (2013).

21.4.3. Model Specification and Estimation

The general structure described in Section 21.4.1 is illustrated in Figure 21.5 as a tree structure with three choice dimensions: trip chain order, vehicle type and delivery location sequence. In order to avoid the computational difficulties deriving from the very large number of elementary alternatives, the first two choice dimensions (trip chain order and vehicle type) were calibrated separately from the delivery location choices.

21.4.3.1. Trip chain order and vehicle type models The choice behaviour of number of stops and vehicle type was modelled, testing both multinomial and nested logit model

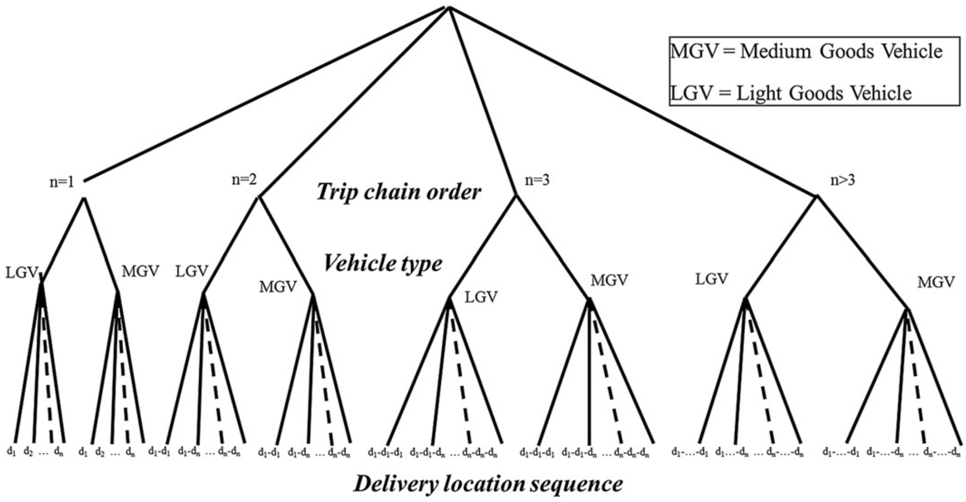

Figure 21.5: Tour delivery model structure.

forms. The calibration results for multinomial logit models are reported in Nuzzolo, Crisalli, and Comi (2012c) and Nuzzolo and Comi (2013). The nested logit was intended for testing structural correlations among choices. Furthermore, the two possible structures of the nested logit tree were considered and different attributes were taken into account during model development.

The models were developed to capture behaviour that can be influenced by tactical and operational measures. For example, the implementation of freight traffic management measures such as time windows, area pricing, route constraints or vehicle type constraints could modify delivery zone accessibility as well as shipment size with subsequent effects on the definition of delivery tours. Hence the attributes refer to level-of-service (e.g. accessibility of warehouse and retail zone), characteristics of delivery (e.g. type of freight and delivered quantity) and tours (e.g. distance covered during current tour).

Referring to random utility theory (Ben-Akiva & Lerman, 1985; Cascetta, 2009), for a given warehouse zone o (for simplicity of notation, the index o will be omitted), the perceived utility U_{nv} of the compound alternative number-of-stops (n) and type-of-vehicle (v) is given by:

$$U_{nv} = V_{nv} + \varepsilon_{vn}$$

Assume that the systematic utility can be expressed as the sum of two terms: $E[U_{nv}] = V_{nv} = V_n + V_{v/n}$, where V_n includes the attributes depending only on the number of stops per tour and $V_{v/n}$ those related to the vehicle for a given n. Similarly, the random residual is expressed as the sum of two independent random variables: $\varepsilon_{nv} = \eta_n + \tau_{v/n}$, thus yielding:

$$U_{nv} = V_n + V_{v/n} + \eta_n + \tau_{v/n}$$

The best calibrated nested logit model which yielded the identification of four classes of tours with one delivery (round trip), two, three and more than three deliveries is reported in Table 21.2 and considers the following attributes:

- Y_n is the inclusive (logsum) variable for the group n ($n \in \{1, 2, 3, >3\}$) obtained with the alternative specific systematic utilities $V_{v/n}$ ($v \in \{\text{LGV, MGV}\}$)
- KM_{o-ZTL} is the average distance from origin zone o to study area (i.e. inner zone-limited traffic zone), expressed in km;
- IAA_o is the *retailer* accessibility index of zone o, from which the tour departs (e.g. warehouse location);
- q is the average quantity of freight delivered at each stop (i.e. delivery point), expressed in tonnes;
- FGT is a dummy variable equal to *1* if the delivered freight belongs to the *foodstuffs* class, 0 otherwise.

The *retailer* accessibility index IAA_o was calculated as (Nuzzolo et al., 2011):

$$IAA_o = \left[AA_o - \min_z(AA_z)\right] \Big/ \left[\max_z(AA_z) - \min_z(AA_z)\right]$$

where AA_x is the accessibility of zone x estimated as

$$AA_x = \sum_j (UL_j)^{6.334} \cdot \exp\left[-3.913 \cdot dist_{xj}\right]$$

with
- UL_j the number of retail establishments of zone j to be restocked,
- $dist_{xj}$ the distance between zone x and j, calculated on the road network according to the path of minimum generalized travel cost, expressed in km.

There were a total of 501 valid observations for the final fitting. The log-likelihood at convergence was –885.53, and the model goodness of fit ρ^2 was about 0.15. The signs of all model coefficients are intuitively correct and the *t-student* value confirms their statistical significance.

Comparing these results with previous findings (Nuzzolo & Comi, 2013), we can see that the nested logit model yields a significant advantage over the multinomial model. Firstly, the value of the coefficient of log-sum Y indicates that the independence of alternatives hypothesis is not acceptable and that the nested logit model is to be preferred. Further, this formulation allows the choices of number-of-stops and type-of-vehicle to be combined following the revealed hierarchical delivery process, that is, the trip chain order choice is made by taking into account the alternatives available at the lower level (i.e. vehicle type). The value of goodness of fit is quite good for this type of model as demonstrated by the comparison with other similar models found in the literature (Comi et al., 2012). The sign of accessibility parameters shows that the number of deliveries per tour decreases if the accessibility of the warehouse zone increases. As emerged from surveys, this confirms that

Table 21.2: Trip chain order and vehicle type model: calibration results ($\rho^2 = 0.15$)

Attribute	**Nest $n = 1$**		**Nest $n = 2$**		**Nest $n = 3$**		**Nest n$>$3**	
	MGV	***LGV***	***MGV***	***LGV***	***MGV***	***LGV***	***MGV***	***LGV***
Logsum (Y)	0.557 (4.5)		0.656 (3.1)		0.656 (3.1)		0.656 (3.1)	
ln of Accessibility index (IAA_o)			−0.057 (−4.8)		−0.026 (−4.8)			
Distance ($KM_{o\text{-}ZTL}$)	0.131 (2.3)		0.051 (1.4)		0.096 (1.8)			
Foodstuffs (FGT, *binary variable*)		0.419 (1.6)		0.043 (1.1)				
Delivered quantity (q, in tons)	1.063 (2.8)		0.984 (3.8)		0.283 (1.1)		0.822 (1.3)	
ASC		2.767 (4.2)		1.534 (3.3)		1.985 (2.9)		0.609 (1.2)

(−) t-st value.

restockers prefer to do round trips if the warehouse is located in a zone with high accessibility as it allows them to reduce the operational complexity of tour management. The probability of using LGVs increases for foodstuffs, while it decreases if the average delivered quantity increases. Finally, as expected, the probability of using MGVs increases according to the distance of the zone to be served. Other models (e.g. mixed logit) are being tested in order to improve these first results including the availability of better quality survey data.

21.4.3.2. Delivery location choice model As regards the *delivery location choice model,* here we recall the models developed in Nuzzolo and Comi (2013). These models first simulate the choice of the first delivery zone and then the choice of the next zones by a set of multinomial logit models. The systematic utility function related to zone d, V_d, is expressed as a linear function of two sets of attributes: variables associated with a destination alternative (e.g. number of employees) and 'memory' variables representing the history of the tour (e.g. distance to be covered in order to reach the next location up to the cumulative covered distance):

- $AD_{d_j^k}$ is the number of retail employees in zone d_j;
- IAA_{d_j} is the retailer accessibility of zone d_j.
- $dist_{d_i^k d_j^{k+1}}$ is the distance between zone d_i^k and d_j^{k+1}, calculated on the road network according to the path of minimum generalized travel cost and expressed in km;
- $DS_{od_j^{k+1}}$ is the share of deliveries on od_j^{k+1} pair with respect to all deliveries departing from zone o ;
- $ASA_{d_j^{k+1}=d_i^k}$ is a dummy variable equal to *1* if the next stop is within the same current zone, *0* otherwise;
- $HT_{d_j^{k+1}}$ is the ratio between the distance to be covered to reach the next delivery location and the current distance covered.

Models that yielded the best statistical significances are presented below. Table 21.3 reports the two sets of calibrated parameters and shows that the probability of choosing for the next delivery a zone that is easy to reach and where many deliveries should be performed is higher. The sign of the memory parameter demonstrates that the systematic utility of choosing a destination is a cost function of the distance from the current stop location. It implies that destinations far away from the current stop location are those which have a lower probability of being chosen. The delivery is therefore a compound alternative composed of the aggregation of elementary alternatives. Further analysis has been developed in order to estimate other models, considering the *size function* as proposed by Ben-Akiva and Lerman (1985) and Cascetta (2009). Further developments also regard choice set modelling.

21.4.4. Model Validation

Once the proposed modelling system had been specified and estimated, it was applied to the restocking of the city centre of Rome in order to verify the model's capability

Table 21.3: Delivery location choice model: calibration results (first: $\rho^2 = 0.33$; next: $\rho^2 = 0.25$)

Attribute	**ln of retail employees (AD_d)**	**ln of accessibility index (IAA_d)**	**Distance ($dist_{didj}$)**	**Delivery share (DS_{od})**	**ln of "memory" (HT_d)**	**Same zone ($ASA_{di=dj}$)**
First delivery location model						
Value	0.213	7.84	−0.028	2.03		
	(3.8)	(9.5)	(−4.8)	(2.1)		
Next delivery location zone model						
Value	0.291		−0.325	8.408	−1.655	1.064
	(3.5)		(−5.2)	(3.2)	(−4.1)	(2.5)

(−) t-st value.

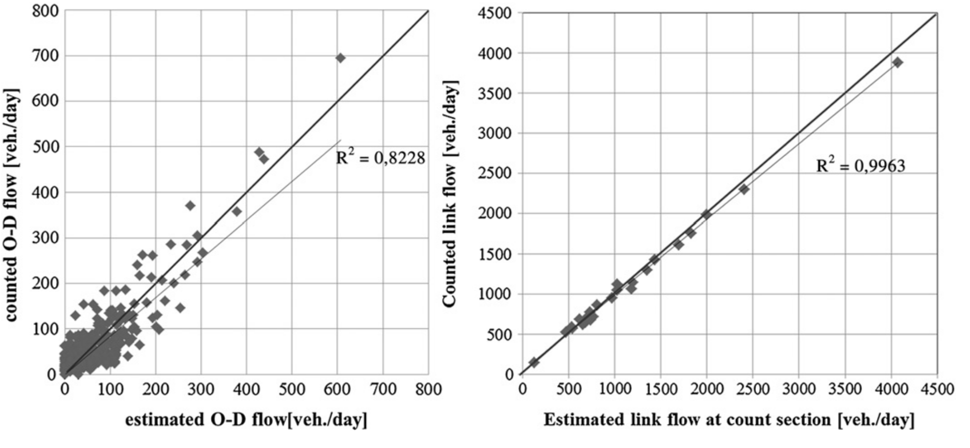

Figure 21.6: Revealed *vs* estimated O-D vehicle flows (left); revealed *vs* estimated link vehicle flows at border sections (right).

to reproduce the revealed freight vehicle O-D and link flows. The study area is about 6 km^2, has about 50,000 inhabitants and 130,000 employees, of which 24,000 are related to trade. The area is served by 14,499 tonnes/day, 34,246 deliveries/day and 18,219 tours/days.

Figure 21.6 reports the comparisons between the revealed and estimated quantity O-D flows, and between the revealed and estimated vehicle link flows. The results of both O-D and link flows confirm the goodness of model structures: the model ably reproduces the current O-D flows and the daily link flow at border sections as confirmed as compared to the 45 degree line. The goodness of fit is quite high with a value of over 0.80. Further analysis is thus in progress to define and compare the more disaggregated results (e.g. flows disaggregated for vehicle type) and the link flows counted within the study area. The latter analysis also requires the freight vehicle path choice to be investigated.

21.4.5. Model Performance

The proposed model allows us to assess how the choices of the three considered dimensions (the number of stops, the type of vehicle and the stop/delivery sequence) could be modified by the implementation of city logistics measures. For example, measures that reduce the load transported, such as vehicle weight constraints, or that modify origin or destination zone accessibility, such as area pricing or designation of delivery areas nearby, can be assessed in terms of new number of tours and relative sequence of stops. The models for the choice of number of stops and vehicle type capture the effects of city logistics solutions: time windows (modifying accessibility could increase the number of tours, which are usually non-optimized in terms of the load transported), vehicle constraints (e.g. weight; reducing the load transported could mean tours with fewer deliveries), incentives to carriers (3P transport could lead to optimized tours), freight class (city logistics measures should focus on food restocking tours).

The delivery location choice model allows us to detect the impacts due to implementation of city logistics that can modify the generalized travel cost and hence accessibility: time windows, pricing and vehicle constraints modify accessibility, leading to longer trips within the same tour; NDA allow environmental-friendly restocking tours to be increased.

21.5. Conclusions and Further Developments of the Study

The chapter presented recent developments in delivery tour modelling that allows freight vehicle O-D matrices to be obtained, hence link flows and road network performance. In particular, we:

- categorized the main measures to be implemented at urban scale in order to manage and control freight transport at different planning horizons,
- proposed a modelling system capable of assessing the short-term measures *ex ante* and predicting actors' choice behaviour.

The chapter was organized into two main parts: the first gave an overview of the urban freight system to be modelled in terms of actors, measures and city logistics strategies; the latter gave the models for freight vehicle O-D flow estimation and a modelling system to simulate delivery tours using an aggregate approach.

The modelling system is structured into two levels: tour definition and sequence of stops/delivery locations. The former focuses on modelling the flow of actual tours leaving from a given zone (e.g. a warehouse) characterized by vehicle type. This allowed the mechanisms driving tour generation to be captured more accurately. Tour definition is modelled by trip chain order and vehicle type. The former identifies the number of stops of each tour, and the latter concerns the definition of the vehicle used, thereby supporting the analysis of measures as well as the travel attributes that may influence restocking patterns. The sequence of stops allows us to model the delivery locations along the tour.

A sequence of behavioural models was formalized and calibrated. The performance and ability to reproduce the revealed flows were then tested. The calibrated models should be considered the first examples developed to test the goodness of the general architecture of the proposed modelling system. The analysis of transferability of these first results is in progress for the city of Santander (northern Spain), where similar surveys have been recently carried out (Ibeas et al., 2012).

Although the obtained statistics confirm the goodness of the approach used, further analysis is required to specify and calibrate other models in order to consolidate the results, to investigate the influence of retail size on tour definition, include size function in delivery location choice, model the choice set generation within the delivery location model and include departure time choice in order to investigate the relationship with time windows access restrictions.

References

Ambrosini, C., Meimbresse, B., Routhier, J., & Sonntag, H. (2008). Urban freight policy-oriented modelling in Europe. In E. Taniguchi & R. G. Thompson (Eds.), *Innovations in city logistics* (pp. 197–212). Hauppauge NY: Nova Science Publishers.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA: MIT Press.

BESTUFS. (2007). *BESTUFS-Good practice guide on urban freight*. BESTUFS Consortium. Retrieved from www.bestufs.net

Cascetta, E. (2009). *Transportation systems analysis: Models and applications*. New York, NY: Springer.

Comi, A., Delle Site, P., Filippi, F., & Nuzzolo, A. (2012). Urban freight transport demand modelling: A state of the art. *European Transport*, *51*. Trieste: ISTIEE. Retrieved from http://www.openstarts.units.it/dspace/bitstream/10077/6123/1/ET_2012_51_7%20-%20Comi%20et%20al.pdf

Crainic, T. G., Ricciardi, N., & Storchi, G. (2009). Models for evaluating and planning city logistics systems. *Transportation Science*, *43*(4), 432–454.

De Jong, G., Vierth, I., Tavasszy, L., & Ben-Akiva, M. (2012). Recent developments in national and international freight transport models within Europe. *Transportation*, *40*(2), 347–371.

Figliozzi, M. (2006). Modeling the impact of technological changes on urban commercial trips by commercial activity routing type. *Transportation Research Record: Journal of the Transportation Research Board, 1964*, 118–126.

Gliebe, J., Cohen, O., & Hunt, J. D. (2007). Dynamic choice model of urban commercial activity patterns of vehicles and people. *Transportation Research Record: Journal of the Transportation Research Board, 2003,* 17–26.

Gonzalez-Feliu, J., Toilier, F., & Routhier, J. L. (2010). End-consumer goods movement generation in French medium urban areas. *Procedia Social and Behavioral Sciences*, *2*(3), 6189–6204.

Holguín-Veras, J., & Thorson, E. (2003). Modeling commercial vehicle empty trips with a first order trip chain model. *Transportation Research Part B: Methodological*, *37*(2), 129–148.

Hunt, J. D., & Stefan, K. J. (2007). Tour-based microsimulation of urban commercial movements. *Transportation Research Part B: Methodological*, *41*(9), 981–1013.

Ibeas, A., Moura, J. L., Nuzzolo, A., & Comi, A. (2012). Urban freight transport demand: transferability of survey results analysis and models. Proceedings of EWGT2012 – 15th Meeting of the EURO Working Group on Transportation, September 2012, Paris. *Procedia – Social and Behavioral Sciences*, *54*, 1068–1079.

Jeffery, D. (2011). *Guidelines for assessing the transferability of an innovative urban transport concept*. NICHES+, Coordination Action funded by the European Commission under the Seventh Framework Programme for R&D, Sustainable Surface Transport. European Commission, Brussels.

Lohse, D. (2004). *Travel demand modelling with model EVA – Simultaneous model for trip generation, Trip Distribution and Mode Choice*. TU Dresden, Working paper.

Muñuzuri, J., Cortés, P., Onieva, L., & Guadix, J. (2012). Estimation of daily vehicle flows for urban freight deliveries. *Journal of Urban Planning and Development, 138*(1), 43–52.

Nuzzolo, A., & Comi, A. (2013). Urban freight demand forecasting: A mixed quantity/delivery/vehicle-base model. *Transportation Research Part E*, Elsevier, forthcoming.

Nuzzolo, A., Crisalli, U., & Comi, A. (2011). A restocking tour model for the estimation of O-D freight vehicle in urban areas. *Procedia – Social and Behavioral Sciences*, *20*, 140–149.

Nuzzolo, A., Crisalli, U., & Comi, A. (2012a). A delivery approach modeling for urban freight restocking. *Journal of Civil Engineering and Architecture, 5*(3), 251–268.

Nuzzolo, A., Crisalli, U., & Comi, A. (2012b). A system of models for the simulation of Urban freight restocking tours. In E. Taniguchi & R. G. Thompson (Eds.), *Procedia – Social and Behavioral Sciences, 39*, 664–676. doi: 10.1016/j.sbspro.2012.03.116

Nuzzolo, A., Crisalli, U., & Comi, A. (2012c). A trip chain model for simulating urban freight restocking. *European Transport, 50*. Trieste: ISTIEE. Retrieved from http://www.openstarts.units.it/dspace/bitstream/10077/6115/1/D_NUZZOLO_et%20-%20ET2012.50.07.pdf

Routhier, J. L., & Toilier, F. (2007). FRETURB V3, a policy oriented software of modelling urban goods movement. In *Proceedings of the 11th world conference on transport research*, Berkeley, CA.

Ruan, M., Lin, J., & Kawamura, K. (2012). Modeling urban commercial vehicle daily tour chaining. *Transportation Research Part E: Logistics and Transportation Review, 48*(6), 1169–1184.

Russo, F., & Comi, A. (2010). A modelling system to simulate goods movements at an urban scale. *Transportation, 37*(6), 987–1009.

Russo, F., & Comi, A. (2011). Measures for sustainable freight transportation at urban scale: Expected goals and tested results in Europe. *Journal of Urban Planning and Development*, *137*(2), 142–152.

Samuelson, P. A. (1952). Spatial price equilibrium and linear programming. *American Economic Review*, *42*(3), 283–303.

Sonntag, H. (1985). A computer model of urban commercial traffic. *Transport, Policy and Decision Making*, *3*(2), 171–180.

Stathopoulos, A., Valeri, E., Marcucci, E., Gatta, V., Nuzzolo, A., & Comi, A. (2011). Urban freight policy innovation for Rome's LTZ: A stakeholder perspective. In S. Melo & C. Macharis (Eds.), *City distribution and urban freight transport: Multiple perspectives* (pp. 75–100). Cheltenham, UK: Edward Elgar Publishing.

Tamagawa, D, Taniguchi, E., & Yamada, T. (2010). Evaluating city measures using a multi-agent model. *Procedia - Social and Behavioral Sciences, 2*(3), 6002–6012.

Thorson, E., Holguín-Veras, J., & Mitchell, J. (2005). Multivehicle routing with profits and market competition. In *Proceedings of Transportation Research Board Meeting*, Washington, DC.

Wang, Q., & Holguín-Veras, J. (2008). An investigation on the attributes determining trip chaining behavior in hybrid microsimulation urban freight models. In *Transportation Research Record: Journal of the Transportation Research Board, 2066,* 1–8.

Wang, Q., & Holguín-Veras, J. (2009). Tour-based entropy maximization formulations of urban freight demand. In *Proceeding of the 88th Transportation Research Board annual meeting,* Washington, DC.

Wilson, A. G. (1969). The use of entropy maximising models in the theory of trip distribution, mode split and route split. *Journal of Transport Economics and Policy*, *3*(1), 108–126.

Wisetjindawat, W., Sano, K., Matsumoto, S., & Raothanachonkun, P. (2007). Micro-Simulation model for modeling freight agents interactions in urban freight movement. In *Proceedings of the 86th Transportation Research Board annual meeting*, Washington, DC.

Chapter 22

Making Real-Time Fleet Management Decisions Under Time-Dependent Conditions in Urban Freight Distribution

Jaume Barceló, Jesús Arturo Orozco and Hanna Grzybowska

Abstract

The design and evaluation of city logistics applications requires an integrated framework in which all components can work together. Therefore, city logistics models should account for vehicle routing applications and fleet management models also capable of dealing with the dynamic aspects of traffic flows in the underlying road network, namely when ICT applications are taken into account. This chapter develops a methodological proposal based on an integration of vehicle routing models and real-time traffic information. In the computational experiments conducted in the chapter, a dynamic traffic simulation model has been used to emulate the actual traffic conditions providing, at each time interval, estimates of the traffic state on each link of the road network that are then used, by a real-time fleet management system, to determine the optimal dynamic routing and scheduling of the fleet.

Keywords: Decision support system; real-time freight management; time-dependent travel times; vehicle routing problem with time windows; pickup and delivery vehicle routing problem with time windows

22.1. Introduction

Fleet management in urban areas has to explicitly account for the dynamics of traffic conditions leading to congestions and variability in travel times severely affecting

Freight Transport Modelling

ISBN: 978-1-78190-285-1

the distribution of goods and the provision of services. An efficient management should be based on decisions accounting for all factors conditioning the problem: customers' demands and service conditions (time windows, service times etc.), fleet operational conditions (vehicles' positions, states, availabilities etc.) and traffic conditions.

The recent advances in information and communications technologies (ICT) have prompted the research on dynamic routing and scheduling problems. At present, easy and fast acquisition and processing of the real-time information is feasible and affordable. An information that becomes the input to dynamic models that capture the dynamic nature of the addressed problems allowing more efficient dynamic fleet management decisions.

This chapter proposes and computationally explores a methodological proposal for a decision support system (DSS) to assist in the decision making concerning the real-time management of a city logistics fleet in dynamic environments when real-time information is available. The feasibility of the proposed solutions depends on:

- the availability of the real-time information, which we assume will be provided by ICT applications, combined with the knowledge of the scheduled plan with the current fleet and customer status,
- the quality of the vehicle routing models and algorithms to efficiently tackle the available information to provided solutions.

A wide variety of vehicle routing problems with time windows becomes the engine of real-time decision-making processes to address situations where real-time information is revealed to the fleet manager and he or she has to make decisions in order to modify an initial routing plan with respect to the new needs (Barceló & Orozco, 2010). Real-time routing problems are mainly driven by events which are the cause of such modifications. Events take place in time and their nature may differ according to the type of service provided by a motor carrier. The most common type of event is the arrival of a new order. When a new order is received from a customer, the fleet manager must decide which vehicle to assign to the new customer and what is the new schedule that this vehicle must follow.

Instead of making decisions based on trial and error and dispatcher's experience a sounder procedure would be to base the decisions on the information provided by a DSS. Regan, Mahmassani, & Jaillet (1996, 1998) provide a conceptual framework for the evaluation of real-time fleet management systems (see Figure 22.1), which considers dynamic rerouting and scheduling decisions implied by operations with real-time information such as new orders or updated traffic conditions.

Unlike the classic approach, where routes are planned with the known demand and they are unlikely to be changed throughout the planning period, the real-time fleet management approach (the dotted frame in Figure 22.1) assumes that real-time information is constantly revealed to the fleet manager who has to decide whether the current routing plan should be modified or not. The logic process, depicted in the diagram in Figure 22.1, assumes partial knowledge of the demand. At the beginning of the considered time period (i.e. one work day) there is proposed an initial schedule

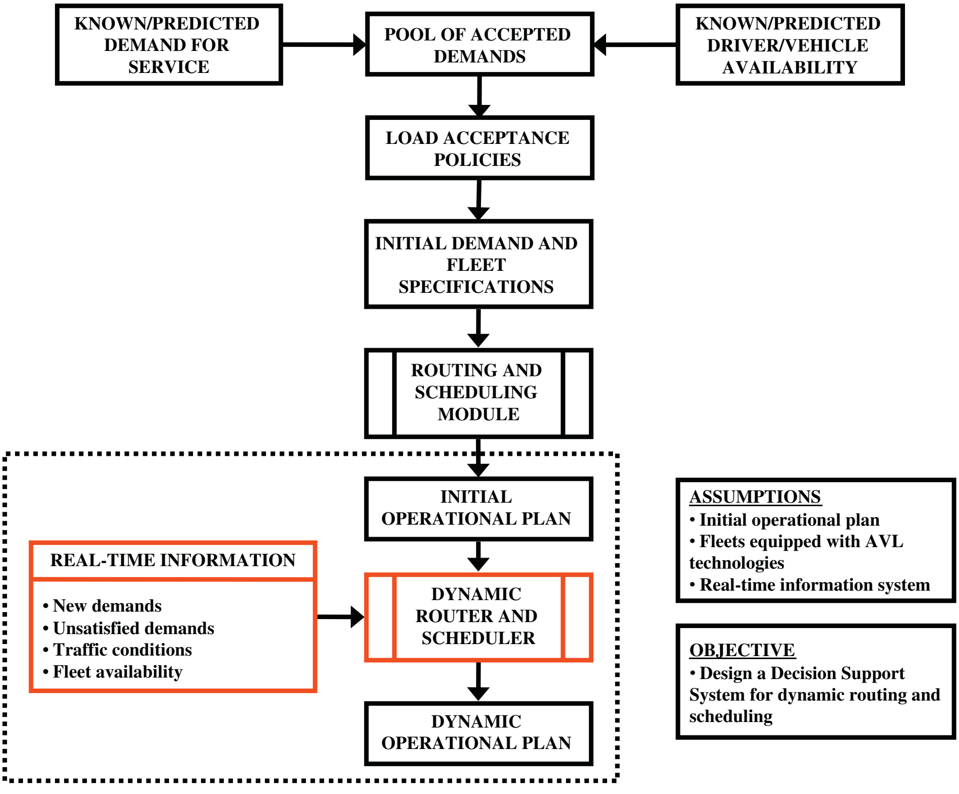

Figure 22.1: Conceptual scheme for the evaluation of real-time fleet management systems. (Adapted from Regan et al., 1998)

for the available fleet to serve the known demand. This initial operational plan can be modified later on, when the operations are ongoing and new real-time information is available. The new information can concern: new demands, unsatisfied demands, changes in the routes due to traffic conditions, changes in the fleet availability (i.e. vehicle breakdowns) etc. It constitutes an input to a *Dynamic Router and Scheduler* (DRS) which provides a proposal of a new dynamic operational plan prepared on the basis of real-time information.

In Figure 22.2, we depict an example of how the conceptual process described in Figure 22.1 can be handled by a dynamic routing and scheduling system. The routes, initially assigned to a set of five vehicles, are identified with various colours. The arrows indicate the order in which customers are to be served according to the initial schedule. We assume that vehicles can be tracked in real time. At time *t*, after the fleet has started to perform its initial operational plan, a new customer calls requiring a service which has not been scheduled. If real-time information, such as positions and states of the vehicles and current and forecasted traffic conditions, is available to the fleet manager, he or she can use it to make a better decision on which vehicle to

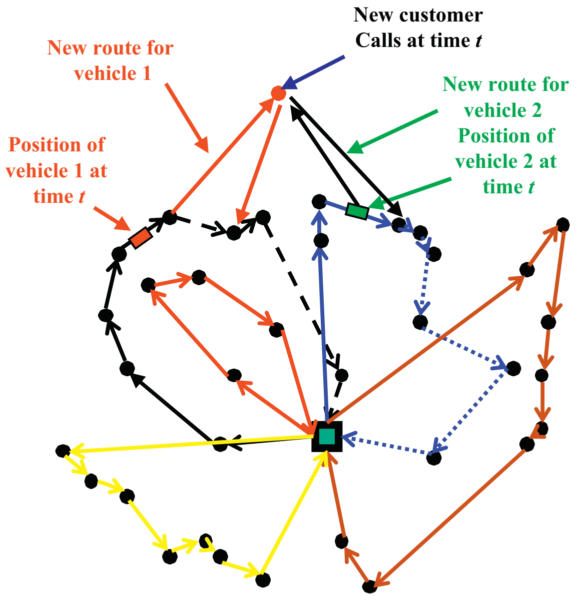

Figure 22.2: Dynamic vehicle rerouting in a real-time fleet management system.

assign to the new customer and whether the new assignment results in a direct diversion from a route (vehicle 2) or a later scheduling (vehicle 1).

Vehicle Routing Problem (VRP) techniques constitute a fundament for the transportation, distribution and *city logistics* systems modelling. The static version of the VRP has been widely studied in the literature. Among others, an extensive survey on VRP was provided by Fisher (1995) and by Toth and Vigo (2002). The static approaches reckon with all the required information to be known a priori and constant throughout the time. However, in most of the real-life cases, a large part of the data is revealed to the decision maker when the operations are already in progress. Thus, the dynamic VRP assumes partial knowledge of the demand and that new real-time information is revealed to the fleet manager throughout the operational period. A comprehensive review of the dynamic VRP can be found in Ghiani, Guerriero, Laporte, and Musmanno (2003). Psaraftis (1995) and Powell, Jaillet, and Odoni (1995) contrast the two variants of the problem and clearly distinguish the dynamic VRP from its static version.

A standard practice of introducing the dynamism into the definition of a routing problem is to determine specific features as *time-dependency*. The VRP with time-dependent travel times acknowledges the influence of traffic conditions on the routing planning. An introduction to the problem and a made heuristic development is provided in Malandraki and Daskin (1992). Further algorithm developments for the static VRP with time-dependent travel times can be found in Ichoua, Gendreau, and Potvin (2003), Fleischmann, Gietz, & Gnutzmann (2004) and Tang (2008).

To our knowledge, the dynamic VRP with time-dependent information has only been addressed by Chen, Hsueh, and Chang (2006) and Potvin, Xu, and Benyahia (2006).

Thomas and White III (2004) addressed the problem by assuming that stochastic information was represented by the time of arrival of a new request. The objective of their approach was to find the best policy for selecting the next node to visit that minimizes the total travel time. They called this problem the *anticipatory routing problem* as vehicles anticipate the arrival of a new customer by changing their path if the request is received while they are in transit. Authors assumed that a new service may be delivered only if the reward or benefit is sufficiently high. The problem was modelled as a finite-horizon Markov decision process and authors used standard stochastic dynamic programming methods to solve each one of the proposed instances. Thomas (2007) extended the results by incorporating waiting strategies with the objective of maximize the expected number of new customers served. They also modelled the problem using a finite-horizon Markov decision process. The authors proposed heuristic algorithms for real-time decision making.

In our approach, time-dependent information is generated by means of traffic simulation of a real-world urban network in order to emulate the role of a Real-Time Traffic Information System providing reliable real-time information on traffic conditions and short-term forecasts. As far as we know, the combination of a real-time traffic information system and vehicle routing models has not been addressed in the literature.

In order to test the management strategies and algorithms assuming real-time information, we have used a dynamic traffic simulation model. In this way, we were able to emulate the estimates of the current travel times as a function of the prevailing traffic conditions and the short-term forecast of the expected evolution of travel times that an advanced traffic information system would provide. This is the basis for more realistic decisions on the feasibility of providing the requested services within the perceptive time windows. Furthermore, simulation can also emulate real-time vehicle tracking giving access to positions and availabilities of the fleet vehicles, which is the information required by a DRS.

22.2. Dealing with Time-Dependent Travel Times

Typical urban networks are subject to numerous random events resulting in traffic conditions with high levels of variability. This is particularly visible during commuting hours (morning and afternoon extended rush hours) when travel times in many streets rise significantly in comparison with other time slots due to traffic congestion. Thus, the travel time of one vehicle, travelling from origin O to destination D, will likely be different leaving the origin at time t_1 than when leaving at time t_2, such that $t_1 \neq t_2$.

In order to illustrate this argument, we have calculated the travel time averages and variances, for several replications, for all links of a dynamic traffic simulation model of the downtown area of the city of Barcelona. This calibrated model was available to us as a result of previous research projects. Both the link travel time averages and variances were computed every five minutes starting from 7 am till 5 pm. The results are shown in Figure 22.3.

Since all the links of the road network don't have the same length and capacity, some variance is expected. Nonetheless, it is possible to clearly identify two time periods where the variability of link travel times is significantly different. The first period (approximately between 7:30 am and 10:00 am) corresponds to the extended morning rush hours. During this time the network passes rapidly from light traffic conditions defined at the beginning of the simulation to a heavy congestion state resulting in higher variability of travel times. The same behaviour can be observed in the extended afternoon hours (last 2 hours of the simulation). In the time slot between these two periods the variability is stationary due to normal traffic conditions.

In addition, there was found evidence that a large proportion of fleet operations take place during rush hours. The results of a survey provided by Robusté and Antón (2005) show that in 60,000 retail outlets in the city of Barcelona, at 11 am more than 50% of the ordered goods have been delivered with an average service time between 13 and 16 minutes.

The above results indicate that working exclusively with average travel times may lead to significant deviations in city logistics problems where temporality is an

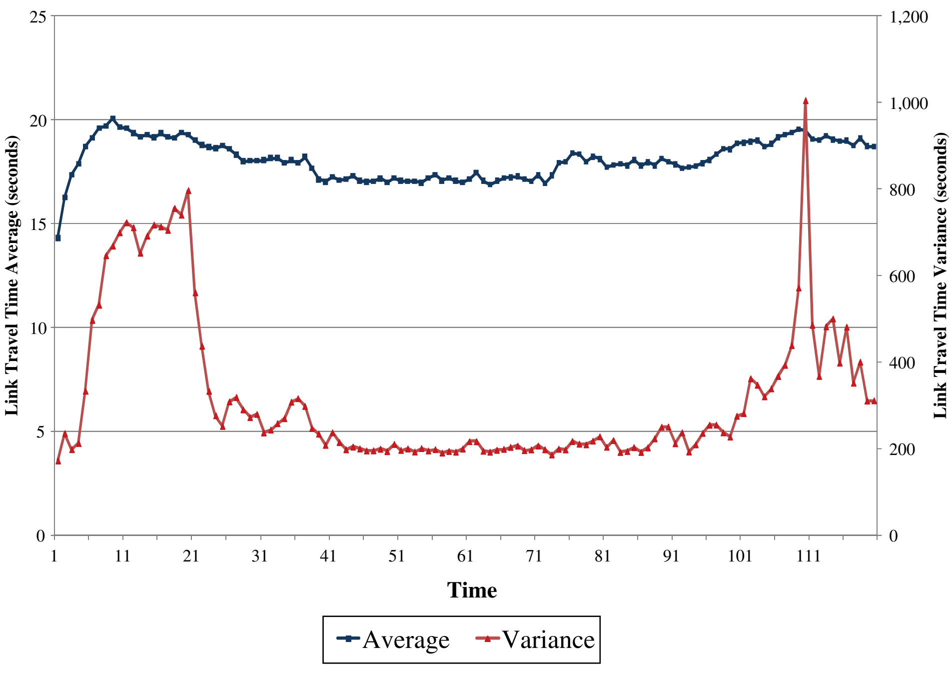

Figure 22.3: Average travel times and travel time variances in downtown area of Barcelona city.

important factor, such as the vehicle routing and scheduling problem with time windows. On that account, in our research, we propose a platform, integrating all the components of a DSS, which is able to emulate an Advanced Traffic Information System (ATIS) which provides link time-dependent data. This data is in turn used to build routes and guide vehicles through the urban network. Therefore, given that time-dependent data is available, the DSS is able to compute time-dependent shortest paths between any pair of nodes of the road network and provide them to a user, which in this case, is the DRS.

Given that our research assumes the operation of the vehicles in an urban environment, we use time-dependent travel times as the basis of computation and assignment of the routing plans.

22.3. Evaluation Platform of the Decision Support System

The evaluation of the routing policies is made through a DSS based on decision rules that use real-time information about the state of the vehicles and traffic conditions provided by a Traffic Information System. The DSS is framed within an evaluation platform that allows us to simulate with great details the performance of a fleet of vehicles as well as the impact of the decisions taken by the fleet manager in face of unknown events such as new customers requiring service, changes in the current demand or changes in the traffic conditions of the network. In essence, the DSS we propose is based on combination of dynamic traffic simulation and the algorithms defined in the DRS. Figure 22.4 shows all the components of the platform we propose and how they are connected with each other.

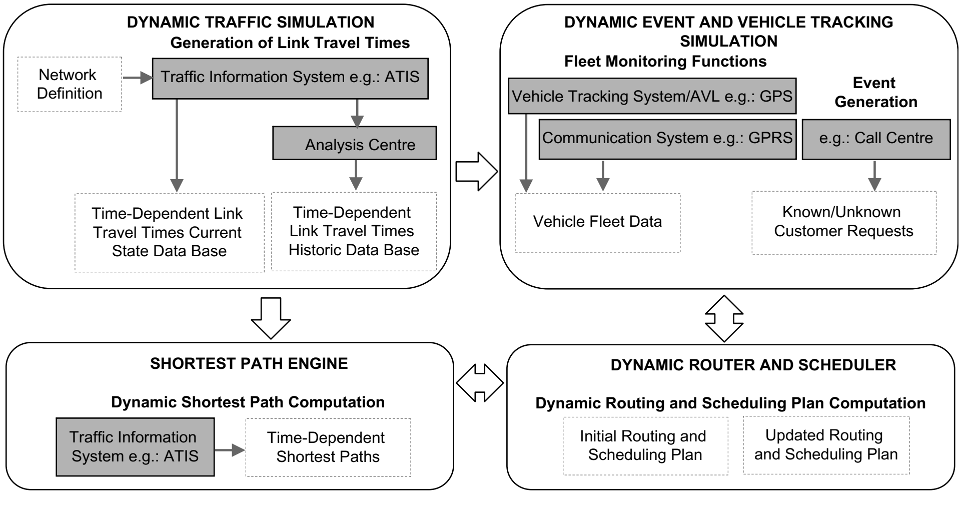

Figure 22.4: Integrated platform for a decision support system for real-time fleet management.

The components of the evaluation platform are: a) the *DRS* which is the core element of the DSS containing a set of algorithmic tools that provide an online solution to a real-time vehicle routing problem, b) the *Dynamic Traffic Simulation* model that emulates the operation of an ATIS, c) the *Dynamic Event and Vehicle Tracking Simulator* which emulates the fleet performance by guiding the vehicles through a modelled urban network using current shortest paths information and, d) a *Shortest Path Engine* that supports the capability of providing vehicle guidance through time-dependent shortest paths calculations.

It is important to remark that the proposed evaluation platform has been designed in such a way that their components can be replaced by real-time systems while keeping the logic and interactions among customers. For instance, the *Dynamic Traffic Simulation* component might be replaced by any real-time traffic information system capable of providing travel times and vehicle guidance, while the *Dynamic Event and Vehicle Tracking Simulator* can be replaced by any real-time fleet tracking and communication system. Therefore, the framework in which the DRS is developed can be implemented in any fleet management system. Next, we will briefly explain the role and the interactions between the four main components of the proposed platform.

22.3.1. Dynamic Traffic Simulation

Dynamic Traffic Simulation constitutes the first component of our approach. It is used as a supporting tool to model the road network and generate databases of time-dependent link travel times. As a modelling tool, dynamic traffic simulation helps us to build and model urban networks by defining the set of: sections, intersections, traffic controls, origin-destination matrices and other important characteristics. We also use this tool to generate two types of databases. The first one is created from several replications of a traffic simulation and it represents the historical records of a traffic information database. The second one is associated to a single replication and it is used to represent the dynamics of the traffic flows within the simulation of the fleet operational plan. It is important to remark that dynamic traffic simulation has a subsidiary role in our model and its usage is motivated by the fact that it can emulate with great detail and realism the traffic conditions of a road network.

22.3.2. Shortest Path Engine

A second component of our approach is represented by the *Shortest Path Engine*. It calculates the optimal vehicle paths among customers and the corresponding travel times. A *path* is represented by the sequence of nodes belonging to the road network that a vehicle follows when travelling among customers.

In the literature, the most studied form of shortest path problems is static. They assume that link travel times or costs are constant. When link travel times change over time, shortest paths are said to be dynamic. In this context, the employed approach to estimate the link travel times might be deterministic or stochastic.

The deterministic approach is based on the assumption that the variable of travel time depends on time. It means that for every time value there is defined a single travel time value. In this approach, Cooke and Halsey (1966) and Orda and Room (1990), Ziliaskopoulos and Mahmassani (1993), Chabini (1998), Chabini and Dean (1999), among others, have proposed algorithms that take into account the variability of link travel times.

In the stochastic approach, the link travel times are defined as random variables with corresponding probability distribution functions that vary with time. Among others, Hall (1986), Miller-Hooks and Mahmassani (1998, 2000), Ahuja, Orlin, Pallottino, and Scutella (2003), Polimeni and Vitetta (2011, 2012), and Liu, Ma, Wu, and Hu (2012) employ this approach to solve problems in stochastic, time-dependent transportation networks.

In this research, we assume deterministic time-dependent travel times and we decided to implement two efficient algorithms for calculating time-dependent shortest paths: Ziliaskopoulos and Mahmassani (1993), Chabini (1998), and Chabini and Dean (1999). The first one has been used for PDVRPTW and the second one for VRPTW. The comparison of the results allowed us to observe that the time-dependent shortest paths obtained by the two algorithms were similar and there were no significant differences in computation times.

The aim of the algorithm of Ziliaskopoulos and Mahmassani (1993) is to calculate the time-dependant shortest paths leading from all the nodes of the underlying network to one point, specified as *destination*. We extended this all-to-one approach in order to obtain the all-to-all type of solution. The change entails multiple executions of the algorithm and involves modification of the data storing structures. The number of the algorithm executions is equal to the number of the nodes in the studied graph.

The algorithm proposed by Chabini (1998) and Chabini and Dean (1999), known as *Decreasing Order of Time* (DOT), computes shortest paths from all nodes to one destination for all departure times. DOT assumes that after certain time period $T_{\max}$, all link costs remain constant. The algorithm starts calculating static shortest paths at time $T_{\max}$ and moves backwards calculating shortest paths for the previous departure time and so forth. Let $G^R = (N^R, A^R)$ be a directed graph that represents the road network where $N^R = \{1, \ldots ,n\}$ is the set of nodes and $A^R = \{(i,j)|i,j \in N^R, i \neq j\}$ is the set of links. Let $d_{ij}(t)$ be the travel time between nodes i and j when departure time is t, $t \in \{0,1, \ldots T_{\max}\}$. Let us assume that we are interested in compute the shortest paths from all nodes of the network to certain destination node q. Let C be a matrix where entry a_{ij} denotes the cost (or travel time) of an optimal path from node i to node q, when departing time is j. Let N be a matrix where entry a_{ij} represents the next node to visit on the optimal path from node i to node q departing at time j. The DOT algorithm is presented in form of a pseudo-code by Algorithm 22.1.

The output of both algorithms has two components. First, we can determine the estimated travel time from any point in the urban network to a certain customer location given that a vehicle departs at a specific time instant. Second, they provide

Algorithm 22.1. DOT algorithm.

1. Initialization

    For $i = 1$ to n **do**:

        For $t = 0$ to $T_{\max}$ **do**:

            $C[i][t] = \infty$

            $N[i][t] = 0$

    For $t = 0$ to $T_{\max}$ **do**:

        $C[q][t] \leftarrow 0$

    Compute static shortest paths at $t = T_{\max}$

2. Main Loop.

    For $t = T_{\max} - 1$ down to 0 **do**:

        For $(i,j) \in A^R$ **do**:

        $t' = \min\{T_{\max}, t + d_{ij}(t)\}$

        if $C[i][t] > C[j][t'] + d_{ij}(t)$ **then**:

            $C[i][t] = C[j][t'] + d_{ij}(t)$

            $N[i][t] = j$

the sequence of nodes (representing streets and turnings) that a vehicle must follow in order to arrive from any point of the network to a customer address. The first output helps us determine the optimal routing of the fleet. The latter emulates the use of a GPS device that guides a vehicle to a given destination point.

22.3.3. *Dynamic Event and Vehicle Tracking Simulation*

Dynamic Event and Vehicle Tracking Simulator represents the core of the simulation and constitutes the third component of our approach. It is based on the general approach of a next-event time-advance simulation by Law and Kelton (2000). It fulfils the twofold objective: (i) randomly generates the different events and (ii) moves vehicles along the road network according to the operational plan keeping track of their positions and activities at every time step.

In an initialization step, the simulator loads all the data required for the emulation of the systems, which include: road network, customers and depot attributes and, current and historical travel times generated by the dynamic traffic simulation. Once the required data has been loaded, the simulator asks the DRS to calculate an initial route assignment. It is used to set the initial vehicle attributes which are, among others, the sequence of customers to visit and the path on the road network that the vehicle will follow. Additionally, it is set the list of events with their occurrence times. A timing routine controls the following event that has been set to occur, advancing the simulation clock and executing the corresponding set of instructions. Whenever there is a change in the simulation clock, the simulator 'moves' the vehicles along the network, using the average link travel times from the current replication database. The simulation ends when the simulation clock reaches an ending time.

Discrete-event simulator emulates the dynamics of the operation of a fleet of vehicles throughout the planning period which, in this case, corresponds to one work day. The events it generates trigger the performance of the DRS. When a new event occurs, the DRS creates a new initial routing and scheduling solution.

We distinguish two types of events: *external* and *internal* (see Figure 22.5). The external events depend mostly on the demand and other factors which are independent of the traffic conditions experienced in the road network, for example, a new customer request, a cancellation of an order, changes in demand or time windows and breakdown of vehicles. Internal events, in contrast, depend on the observed traffic flow dynamics which could be the cause, for instance, of delays in the estimated arrival times or service start times.

In this chapter, we deal with two types of events: the departure of a vehicle from a customer's location and the arrival of a new customer request. When a vehicle departs from a location, the simulator records the exit time of the vehicle and computes the time-dependent shortest path to the next customer in the route. The

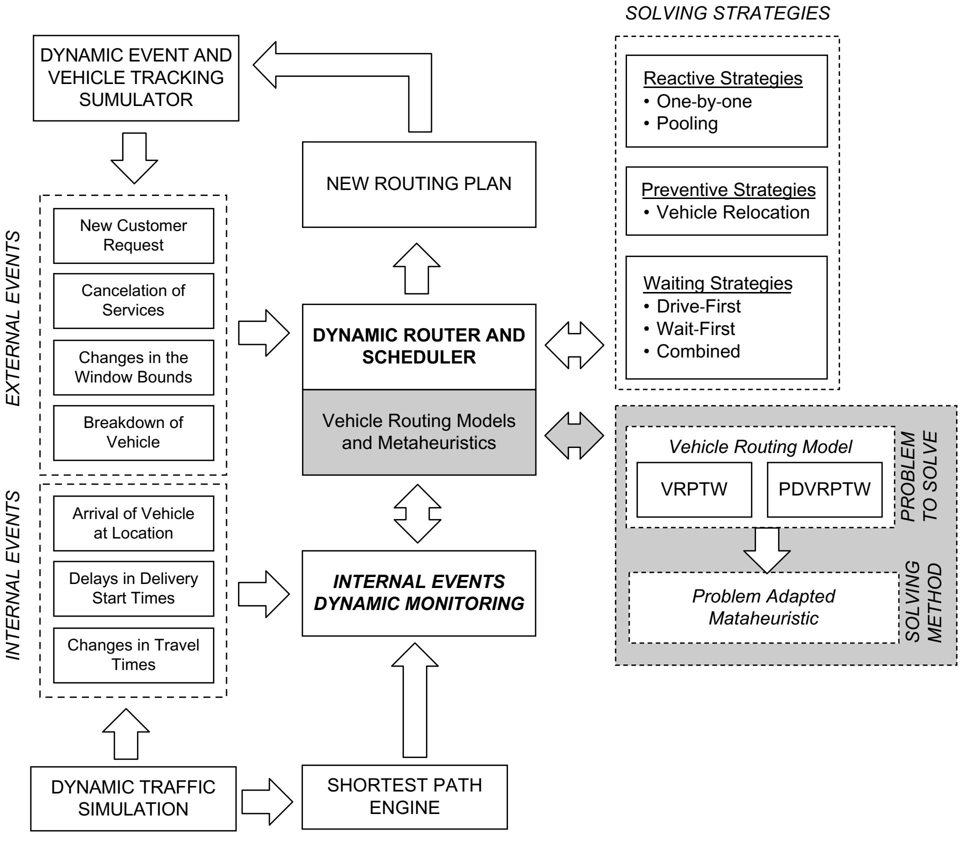

Figure 22.5: Conceptual framework of the Dynamic Router and Scheduler.

simulator then changes the vehicle's internal state to 'in transit' and updates the path on the network that the vehicle must follow to arrive to the next scheduled customer. When a new request is received, the simulator processes and passes all the required information to the DRS, which include current vehicle positions, customers that have not been served, and the time-dependent travel times computed from the current network conditions. If a solution is found, the route plan is updated accordingly and the next new customer event is scheduled. A penalty may be applied if the algorithm is not able to find a feasible solution that includes the new customer.

22.3.4. Dynamic Router and Scheduler

DRS is the last component of our platform and the core element of the proposed DSS. It interacts with the Dynamic Event and Vehicle Tracking Simulator and the Shortest Path Engine in order to recalculate routes and, if needed, propose new solutions. It consists of a set of algorithmic tools to decide which vehicle will be assigned to the new task and to provide an online solution to the underlying real-time VRP (see Figure 22.5).

In this chapter, we study two main variants of the dynamic VRP. First, we address the problem of receiving new calls when only time windows constraints are considered. Afterwards, we analyse the case of having an additional pairing and precedence constraints for pickups and deliveries. Hence, at present, the DRS collection comprises heuristic approaches to solve: the Vehicle Routing Problem with Time Windows (VRPTW) and the Pickup and Delivery Vehicle Routing Problem with Time Windows (PDVRPTW).

The objective of the VRPTW is to fulfil the customers' demands within required time windows and a scheduling horizon while minimizing the total cost of the trips. It is applicable in the cases, when a vehicle gets loaded/unloaded at the depot and distributes/recollects goods in accordance with the imposed constraints and requirements. However, in the city logistics environment it is a common request for a vehicle to collect cargo from a certain customer in order to deliver it to another, specific customer (e.g. courier and postal service). In those cases a PDVRPTW is used.

22.4. Heuristics of the Dynamic Router and Scheduler

In this section, we describe the heuristics used by the DRS in the situation when a new request is reported by a customer or a pair of pickup/delivery customers (see Figure 22.6). We assume that the initial routing plan has been previously computed using some optimization technique as *Tabu Search* (TS).

DRS starts with a review of all customers that have not been provided with the requested service at the time of the new customer call. In the particular case of PDVRPTW, delivery customers that have not been served and whose pickup partners have already been visited and the loads collected, are acknowledged as customers that need to be served obligatorily by the initially assigned vehicle. Hence,

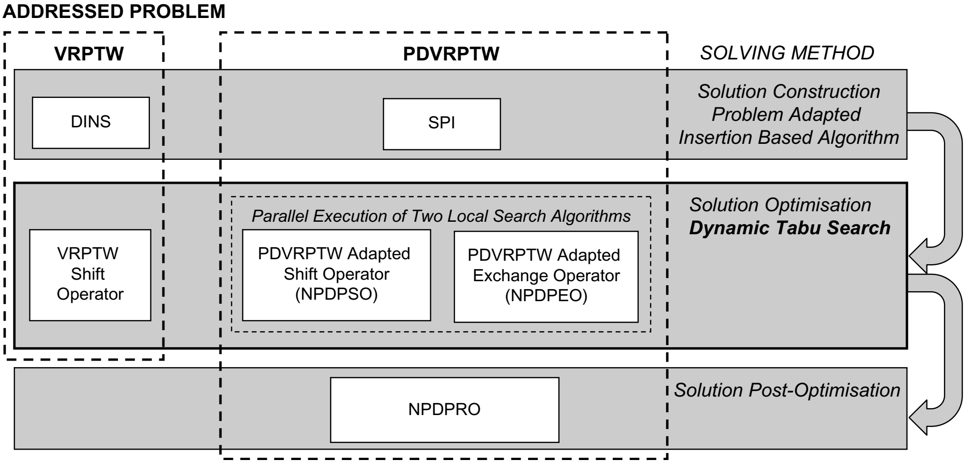

Figure 22.6: Main methodology and heuristics of the Dynamic Router and Scheduler.

they cannot be relocated in any of the subsequent processes modifying the original solution.

In the next step, the algorithm performs a new customer insertion procedure. It is adapted to the addressed problem, either VRPTW or PDVRPTW. In the case of the dynamic VRPTW, new orders are inserted with a greedy insertion heuristic called *Dynamic Insertion Heuristic* (DINS). In the case of the dynamic PDVRPTW, when a new request arrives, it is incorporated into the existing routes or, if necessary, used to create a new route with the *Dynamic Single Pair Insertion Procedure* (SPI). Both alternatives are explained in detail in the following sections. In all cases, we assume that requests are attended in a one-by-one basis.

The created routes are optimized subsequently by using a TS meta-heuristic that we have called: *Dynamic Tabu Search* (DTS). Its design allows for embedding individual algorithms, which are called when solving specific problems. On that account, DTS constitutes a common frame of the approach solving all the addressed problems. Each phase of DTS is explained in detail in the subsequent section.

In contrast to VRPTW, the definition of PDVRPTW is more restrictive which makes it more difficult to solve. Thus, in order to accelerate the computation and to improve the results of PDVRPTW, it was decided to introduce two local search operators performing in parallel within DTS. For the same purpose, a post-optimization step is added to improve the solution on the level of a single route.

22.4.1. Dynamic Insertion Heuristic (DINS)

DINS is a greedy heuristic derived from the basic insertion heuristic proposed in Campbell and Savelsbergh (2004) whereby every un-routed customer is evaluated at every insertion point. The evaluation of this movement consists in checking both feasibility and profitability of every insertion, which is the negative of the amount of additional travel time added to the route. The customer with a feasible insertion having the maximum profitability is then selected and inserted in the route. The authors proved that the total complexity for this algorithm is given by $O(n^3)$. We have adapted this heuristic for the case of the dynamic VRPTW by introducing some new elements to reflect the dynamics of the operations of a fleet in an urban environment.

The objective of DINS is to insert a new customer into the current routing and scheduling plan once the vehicles have started performing the services. The general idea of the algorithm is as follows: when a new customer arrives to the system, the algorithm checks the current state of the vehicles. Then, routes with insufficient time to visit the new customer within their schedule are rejected. Finally, the algorithm checks, on the candidate routes, for the least cost feasible insertion. If no feasible insertion is possible and there are idle vehicles at the depot, a new route is created including the new customer.

22.4.1.1. Notation We assume that there are n routed customers in R different routes at the beginning of the time planning horizon. We also suppose that there is one single *Depot* with a homogeneous fleet of vehicles with capacity Q. A route is a sequence of customers beginning and ending at the Depot. For convenience, the Depot is identified with 0 (start of route) and $n + 1$ (end of route). The following additional notation is required for the construction of the algorithm:

- $T_{i,j}(t)$: travel time between customers i and j when vehicle departs at time t,
- V_r: vehicle assigned to route r, where $r = 1, 2, \ldots, R$,
- q_r: total demand assigned to route r,
- d_i: demand of customer i,
- E_i: time window lower bound of customer i,
- L_i: time window upper bound of customer i,
- e_i: earliest service start time at customer i,
- l_i: latest service start time at customer i,
- S_i: service time at customer i.

The time window of a customer i is denoted by (E_i, L_i) and we assume that the time window of the depot is given by $(0, T)$, where T is the end of the planning period. The values e_i and l_i, refer to the earliest and latest time a service can take place at customer i, and they must satisfy the following condition: $E_i \leqslant e_i \leqslant l_i \leqslant L_i$.

When a new call is received, the system must decide where to insert the new customer. Let us assume that a new customer w arrives to our system at time instant $t > 0$. When this happens, the vehicles of the fleet may have one of the following three states:

1. The vehicle is in service at some customer i (SER).
2. The vehicle is moving to the next planned customer on the route or waiting at the customer location to start service within the time window (MOV).
3. The vehicle is idle at the Depot, without a previously assigned route (IDL).

We call this situation the *status* of a vehicle. The status will let us know whether a vehicle should: (i) be diverted from its current route, (ii) be assigned to a new route if it is idle or (iii) keep the planned trip. Whenever a new customer arrives, the status of each vehicle must be known to compute travel times from their current positions to the new customer.

22.4.1.2. Algorithm The first iteration of the algorithm consists of rejecting those routes which are definitely infeasible. In order to do this, it is necessary to compute the total available time of the routes, that is, the sum of the available time (if any) of a vehicle on the assigned route. We define the available time of a vehicle as those periods of time when: (i) a vehicle is waiting to provide service to a customer or (ii) a vehicle has finished the scheduled route and is returning to the depot. This is, in fact, the *maximum slack time* a vehicle has available to travel and service a new customer call.

The calculation of the *available route time* can be done by estimating the *arrival times* at each of the scheduled customer locations that has not been served at the time of arrival of the new request w. The arrival time a_i at customer i, is computed from the current position of the vehicle and we assume that the vehicle will start service as soon as it arrives if the customer's time window is already open; otherwise, the vehicle waits. The calculation of arrival times is done with the following procedure.

If, at instant t, the vehicle is moving towards the next customer then the estimated arrival time a_i is given by

$$a_i = t + T_{V_r,i}(t) \tag{22.1}$$

The arrival times at subsequent customers on the route are computed as follows:

$$a_i = \max\{E_{i-1}, a_{i-1}\} + S_{i-1} + T_{i-1,i}(\max\{E_{i-1}, a_{i-1}\} + S_{i-1}) \tag{22.2}$$

Hence, the available time of a vehicle in route r at some node i (including the Depot) is the difference between the associated lower limit of the time window and the arrival time of the vehicle to that location. That is

$$W_i^r = \max\{E_i - a_i, 0\} \tag{22.3}$$

Therefore, the total available time of a route r is given by

$$AT_r = \sum\nolimits_i W_i^r \tag{22.4}$$

From each route, we choose the unvisited customer i that is located the closest to the new customer w. If the travel time from i to w is greater than the total available time AT_r, then we reject the route. If every route is rejected, a new route must be created for the new customer. The routes obtained, as a result of this previous iteration, are *possibly feasible* in the sense that vehicles serving those routes have enough time to travel to the new customer. This is a necessary condition, but not sufficient, because we have to check for time windows feasibility. The second major iteration of the algorithm consists, therefore, in checking the feasibility and profitability of an insertion in every arc of the possible feasible routes. The feasible insertion with the highest profitability is then selected and the corresponding route must be updated.

Feasibility In order to evaluate the feasibility of the insertion of customer w between customer i and $i+1$, we must compute e_w and l_w, the expected earliest and latest service start times, respectively. The earliest service start time at the new customer w is given by

$$e_w = \begin{cases} \max\{E_w, t + T_{V_r,w}(t)\}, & \text{if } V_r = i \\ \max\{E_w, e_i + S_i + T_{i,w}(e_i + S_i)\}, & \text{otherwise} \end{cases} \tag{22.5}$$

The latest service start time is calculated through a backward procedure starting with the depot (the last node to be visited by the vehicle) as follows:

$$l_w = \min\{L_w, \tilde{t}_w - S_w\} \tag{22.6}$$

where

$$\tilde{t}_i = \arg\max_t \{t + T_{i,i+1}(t) - l_{i+1}\}, \forall t \leq l_{i+1} - T_{i,i+1}(t) \tag{22.7}$$

is the latest possible departure time that allows the vehicle to arrive to the next scheduled customer $i+1$ at time $l_i + 1$.

If $e_w \leqslant l_w$ and $D_w < Q - q_r$, the insertion is feasible.

Profitability The profit of an insertion is defined as the negative of the additional travel time incurred from inserting the new customer in a route. If a new customer w is to be inserted between customers i and $i+1$, the profitability of this insertion is given by

$$\begin{aligned} \mathit{profit} = -\big[&T_{i,w}(\max\{E_i, a_i\} + S_i) + T_{w,i+1}(\max\{E_w, a_w\} + S_w) \\ &- T_{i,i+1}(\max\{E_i, a_i\} + S_i)\big] \end{aligned} \tag{22.8}$$

If no insertion is possible and there are available vehicles at the depot, then a new route that includes customer w is created. If the whole fleet of vehicles is already occupied, then the call is rejected and postponed for service within the next planning period. Given the arrival of a new customer w at time t and a set R of pre-planned routes, the pseudo-code of the heuristic is shown in Algorithm 22.2.

22.4.2. *Dynamic Single Pair Insertion Procedure (SPI)*

The objective of the dynamic SPI procedure is to incorporate a newly reported pickup/delivery pair of customers in the most favourable location in one of the

Algorithm 22.2. DINS heuristic.

1. **For** each route r in R **do**:
   Compute available time AT_r
   Find travel time from w to the closest neighbour
   If available time is not enough, then reject route r
   Else, add route r to R', the set of possible feasible routes
2. **For** each route r in R' **do**:
   Check the status of the vehicle and set its position as the starting node of the route
   For each arc (i, i + 1) of the remaining route sequence **do**:
      If insertion of w is feasible and profit is improved **then**:
         Store current profit and insertion places
3. **If** insertion is possible **then**:
   Insert w in the least cost arc and update selected route.
   Else: create a new route with an available idle vehicle (if any).

existing routes without violating any of the established constraints, respecting: load, time windows, pairing and precedence.

Similarly as DINS, SPI creates a new routing plan reckoning with dynamic factors affecting the addressed problem. Hence, in contrast to the static approach, the variables specifying: travel, arrival and waiting time for all un-served customers are calculated with respect to the triggering instant t and depend on the moment a trip between two customers starts.

The value of the arrival time to the first customer i in the updated sequence, which has not been served by the assigned vehicle v before the time instant t, is estimated taking into consideration its present status.

Each route, in which it is feasible to introduce the new pair of customers, is determined as a candidate. Among all the candidates there is selected the one, which characterizes with the smallest value of the increment of the total cost. In the case when it is not feasible to introduce the new pair in any of the existing routes and there is available an idle vehicle a new route is created.

22.4.2.1. Notation We employ the notation provided in Section 22.4.1.1. However, in order to explain the functioning of the dynamic SPI process, the original version is extended as follows:

- V_S: list of dynamic pairs requesting service,
- P: dynamic pair requesting service,
- p: a pickup customer from pair P,
- d: a delivery customer from pair P,
- $q_r(j)$: current occupancy of capacity of vehicle r after serving customer j.

22.4.2.2. Algorithm The new customer pair is inserted into the existing paths fulfilling the: pairing, precedence, time windows and capacity constraints. The first two restrictions are enforced as a matter of course of the design of the insertion process. Thus, the feasibility check regards verification if both the time windows and the capacity constraints are violated.

Feasibility The route r is feasible after the insertion of the customer w between customers i and $i + 1$, if:

$$a_j \leq l_j, \text{for all customers } j \in r \text{ and } j = \{w, i + 1, \ldots, n + 1\} \tag{22.9}$$

and

$$q_r(j) \leq Q, \text{for all customers } j \in r \text{ and } j = \{w, i + 1, \ldots, n + 1\} \tag{22.10}$$

Profitability In order to assess profitability of the insertion of a new pair of customers, dynamic SPI calculates the difference between the estimated arrival time at the last node of the route a_{n+1} before and after the insertion. The insertion move resulting with the smallest increment of the total route cost is selected to be implemented.

The modus operandi of the dynamic SPI procedure is presented in form of a pseudo-code by Algorithm 22.3.

Algorithm 22.3. Dynamic single pair insertion procedure.

1. **For** each pair $P \in V_S$ **do**:
   - **For** each $r \in R$ **do**:
     - Check the status of the vehicle and set its position as the starting node of the route
     - Calculate a_i for all $i \in r$
     - **For** each $i \in r$ and $i \neq n + 1$ **do**:
       - Insert p after i
       - Update a_i for all $i = \{p, i + 1, \ldots, n + 1\}$
       - **For** each customer $j \in r$ and $j = \{p, i + 1, \ldots, n\}$ **do**:
         - Insert d after j
         - Update a_j for all $j = \{d, j + 1, \ldots, n + 1\}$
         - **If** it is feasible to insert pair P in route r **then**:
           - Store $a_n + 1$ before and after insertion of pair P
           - Store insertion positions of customers p and d in r
           - Add r to the candidate list R'
2. **If** $R' \neq \emptyset$ **then**:
   - From all $r \in R'$ pick the route which increment of $a_n + 1$ is minimal
   - Insert customers p and d in stored insertion positions
   - Update selected route
3. **Else**: create a new route with an available idle vehicle (if any).

22.4.3. Dynamic Tabu Search (DTS)

In order to optimize the solution obtained by either DINS or SPI heuristics, we have implemented DTS constituting an adaptation of the *Unified Tabu Search* (UTS) heuristic proposed by Cordeau, Laporte, and Mercier (2001) for a VRPTW. The original design of the heuristic was modified so that the dynamic aspects of the problem are taken into account. For the most part, we have changed the heuristic to perform local searches only on the unvisited customers of the scheduled routes and added a stopping criterion which takes into account the rate of improvement in the search process.

In general terms, DTS is an iterative procedure that drives the search from a current solution s to the best solution found in a subset of the solution space called the neighbourhood of s, or $N(s)$, by performing local searches. $N(s)$ consists of the set of solutions that are reached when an operator is applied to the current solution s.

As it is based on tabu search, DTS uses anti-cycling rules (tabu lists) and diversification mechanisms that allows the search to explore other areas of the solution space while escaping from local optima. During the search process, a solution s is evaluated by the following cost function:

$$f_D(s) = c_D(s) + \alpha q(s) + \beta d_D(s) + \gamma w_D(s) \tag{22.11}$$

Where:

- $c_D(s)$: estimated total travel time of the routes on the pending scheduled customers,
- $q(s)$: total capacity violation of the vehicles,
- $d_D(s)$: total duration violation of the routes,
- $w_D(s)$: total violation of the time windows constraints of the customers to be visited.

Let $B(s) = \{(i,k)$: customer i is visited by vehicle $k\}$ denote the attribute set of a solution s. Every solution s' such that $f_D(s') \geqslant f_D(s)$ is penalized with a modified function:

$$p(s') = \lambda c_D(s')\sqrt{n_u m}\sum\nolimits_{(i,k)\in B(s')}\rho_{ik} \tag{22.12}$$

where

- n_u: number of customers that have not been visited,
- m: current number of routes,
- λ: a positive parameter,
- ρ_{ik}: defined as the number of times the attribute (i,k) has been added to a solution.

The incorporation of penalization in the objective function calculation aims to diversify the search process towards novel or not very thoroughly checked areas, in the situations when the local optimum was reached. The goal of the factor λ included in the penalty calculating formula is to control the diversification's intensity.

The same objective has the addition frequency parameter, which is calculated as the sum of number of times each customer (or pair of customers in the PDVRPTW) i has been introduced into a route served by the vehicle k during the best solution search process. The value of this parameter is updated regularly after the completion of each iteration of DTS.

In every iteration, the algorithm selects the best non-tabu solution $s' \in N(s)$ that minimizes the values of the cost function $f_D(s)$ or satisfies an aspiration criteria. Our aspiration criterion accepts a solution if its cost improves the best solution found so far. In addition, as in UTS, we employed a relaxation mechanism which facilitates the exploration of the solution search space by dynamically adjusting the values of parameters: α, β and γ. When the selected solution s' is feasible, the value of the best feasible solution is updated and the values of the three parameters are reduced by dividing the current values by $1 + \delta$; otherwise, the values are increased by multiplying the current values by the factor $1 + \delta$, where δ is a parameter of fixed value. As reported by Cordeau et al. (2001), this relaxation mechanism is useful especially in the cases when tight constraints are defined.

As quick online solutions are required, we added a criterion that stops the search if a certain number of iterations κ has been reached without observing an improvement larger than ε.

The design of DTS, derived from UTS, was additionally modified so that it becomes a framework in which various local search operators can be embedded. Depending on the addressed problem, a different operator is called.

When the solved problem is a dynamic VRPTW the solution provided by DINS heuristic represents the initial solution for the TS heuristic. The $N(s)$ is generated by applying a *shift operator* to the current solution s. The shift operator removes a customer i from route k and inserts it in route k'. The possible set of insertion positions in a route starts from the current position of the vehicle. If a vehicle is waiting to start service or heading to a customer's location, the heuristic allows the diversion of the vehicle, that is, the route can be modified to visit a different customer than the originally scheduled. When customer i is removed from route k, its reinsertion in that route is forbidden for the next θ iterations. The pseudo-code for the dynamic TS procedure, when addressing a dynamic VRPTW, is described by Algorithm 22.4.

When a dynamic PDVRPTW is addressed, the solution provided by the dynamic SPI procedure constitutes the initial solution for the TS heuristic. In order to quickly solve the problem, we introduced parallelism of operations by performing synchronized and simultaneous search of the next move to execute. Hence, the routine starts with dividing the collection of possible movements into as many subsets as the number of available processors. Then in each sub-division the individual processor seeks for the best move. We were using a personal computer with two processors. Consequently, two neighbourhoods are generated by executing two local search operators. Their performance was adjusted so that each modification of a route concerns a pickup-delivery couple instead of an individual customer. This feature affects the architecture and functioning of the UTS structures such as for example the adaptive memory, as well as the general performance of the selected local search operators, which need to consider the additional constraints of pairing

Algorithm 22.4. DTS heuristic.
1. **If** solution s is feasible **then**:
   Set: $s^* = s$, $\alpha = 1$, $\beta = 1$, $\gamma = 1$
   Set: $c(s^*) = c(s)$, else set $c(s^*) = \infty$
2. **For** $i = 1, \ldots, \eta$ **do**:
   Choose solution $s' \in N(s)$ that minimizes $f(s') + p(s')$ such that s is not tabu or satisfies aspiration criteria
   If s' is feasible and $c(s') < c(s^*)$ **then**: $s^* = s'$; $c(s^*) = c(s')$
   If $q(s') = 0$ **then**: $\alpha = \alpha/(1 + \delta)$, **else** $\alpha = \alpha(1 + \delta)$
   If $d(s') = 0$ **then**: $\beta = \beta/(1 + \delta)$, **else** $\beta = \beta(1 + \delta)$
   If $w(s') = 0$ **then**: $\gamma = \gamma/(1 + \delta)$, **else** $\gamma = \gamma(1 + \delta)$
   Set $s = s'$
   Update α, β and γ
   If (number of iterations without improvement is κ) and (last improvement $< \varepsilon$) **then**: STOP

Algorithm 22.5. Parallel tabu search Meta-Heuristic.
1. **If** the solution s is feasible **then**:
   Set: $s^* = s$, $\alpha = 1$, $\beta = 1$, $\gamma = 1$
   Set: $c(s^*) = c(s)$, else set $c(s^*) = \infty$
2. **For** $i = 1, \ldots, \eta$ **do**:
   Execute in parallel operators NPDPSO and NPDPEO and create two neighbourhoods: $N(s)_S$ and $N(s)_E$ respectively
   Select solution $s_S \in N(s)_S$ that minimizes $f(s_S) + p(s_S)$ such that s_S is not tabu or satisfies aspiration criteria
   Select solution $s_E \in N(s)_E$ that minimizes $f(s_E) + p(s_E)$ such that s_E is not tabu or satisfies aspiration criteria
   If s_S is feasible and $c(s_S) < c(s^*)$ and $c(s_S) < c(s_E)$ **then**: $s^* = s_S$, $s = s_S$ and $c(s^*) = c(s_S)$
   If s_E is feasible and $c(s_E) < c(s^*)$ and $c(s_E) < c(s_S)$ **then**: $s^* = s_E$, $s = s_E$ and $c(s^*) = c(s_E)$
   Update α, β and γ
   If (number of iterations without improvement is κ) and (last improvement $< \varepsilon$) **then**: STOP
3. Apply the post-optimization procedure to s^*

and precedence (Grzybowska & Barcelo, 2011). The parallel DTS procedure for dynamic PDVRPTW is explained in a form of a pseudo-code by Algorithm 22.5.

The engaged local search operators are: the *Normal Pickup and Delivery Pair Shift Operator* (NPDPSO) and the *Normal Pickup and Delivery Pair Exchange Operator*

(NPDPEO). The objective of NPDPSO is to remove a pair of customers from the original route and feasibly insert it in another route of the current solution, in such a way that its total cost is minimized and the precedence constraint is respected. NPDPEO modifies routes by exchanging two pickup/delivery customer pairs between two routes, in such a way that the removed from their original positions customers get reinserted in new locations, providing feasibility and total cost optimization. The routs must stay connected, thus the direct predecessors and successors of the eliminated customers are properly linked.

Both NPDPSO and NPDPEO are executed simultaneously, in each iteration of the parallel DTS procedure. As a consequence, two different neighbourhoods: $N(s)_S$ and $N(s)_E$ of the same, currently best solution s are created and searched. The whole exploration space gets expanded and the probability of the complete process to be entrapped within local optima is reduced. Among the best solutions found in each neighbourhood, the one which brings larger benefits is selected for the final realization (Algorithm 22.6).

The process of shifting and exchanging of pairs of customers has been used under different names by other authors (e.g. Lau & Liang, 2002; Nanry & Barnes, 2000; Park, Okano, & Imai, 2000). It is similar in the case of the routes' rearrangement (e.g. Li & Lim 2001; Nanry & Barnes, 2000). The usage of the term normal in the operators' designations in the current chapter is to highlight their difference from their reversed versions, which permits the delivery customer to be placed in a new route before its pickup partner.

In order to improve the final solution there was introduced a post-optimization step. The *Normal Pickup and Delivery Pair Rearrange Operator* (NPDPRO) is the last method executed by the DRS when solving a dynamic PDVRPTW. Its objective is to quickly improve the solution produced by parallel DTS. It recomposes each original route, by changing the order in which the customers are to be visited, in order to find a new sequence of customers with a smaller value of the objective function. Thus, the changes are performed within the boundaries of only one route. Since it is the post-optimization routine, there are only accepted the modifications resulting with a strictly feasible solution. NPDPRO is explained in a form of a pseudo-code by Algorithm 22.6.

Algorithm 22.6. Normal pickup and delivery pair rearrange operator.

For each route $r \in R$ **do**:
- Set: $r^* = r$ and calculate: $q(r^*)$, $w(r^*)$, $c(r^*)$
- **While** best movement is not found for each customer $i \in r$ **do**:
  - Find pair partner $p(i)$ of customer i
  - Remove both customers i and $p(i)$ from route r
  - Find the best insertion of customers i and $p(i)$ in route r
  - Calculate: $q(r')$, $w(r')$, $c(r')$
  - **If** $c(r') < c(r^*)$ and route r' is feasible **then**:
    - The best movement was found $c(s^*) = c(s')$
- **If** the best movement was not found **then** set $r^* = r$

22.5. Computational Experience

A set of computational experiments has been designed and conducted to evaluate the performance of a fleet using the simulation platform to emulate daily traffic city conditions, under different scenarios.

For our computational tests, we have used a microscopic simulation model of downtown area of Barcelona (Figure 22.7). It covers 7.46 km^2 and it is one of the most important districts in the city due to the high volume of commercial activity. The graph that represents the road network consists of 1,570 nodes and 2,797 arcs. The base scenario includes a logistics centre with one single depot with a fleet of homogeneous (i.e. same capacity) vehicles. This depot provides service to a set of 100 customers located on the map. All customers have time windows with randomly generated start times and widths. It is also assumed that all customers have constant demand and service time.

22.5.1. Real-Time Traffic Information System

In order to emulate the real-time traffic information system of the urban network, we generated a database of historic travel times through microscopic traffic simulation software. This database, which represents the long-term travel time forecasts, contains the average travel times of all links on the network for every 5-minute interval of 15 replications of a simulation under normal traffic conditions. We additionally generate a set of 12 replications that are used by the dynamic event and vehicle tracking simulator to move vehicles along the network. These replications also represent the short-term travel time forecasts that combined with the long-term forecasts of the historic travel time database provides the input to the

Figure 22.7: Barcelona's downtown network.

time-dependent shortest paths calculations that are used to guide vehicles through the network.

The dynamic traffic simulation software used in the generation of the databases described above needed some specific settings in order to model as close as possible the characteristics of the road network. In the traffic simulation software that we used, route choice models are needed to describe the behaviour of drivers when deciding which path to take from a set of different alternatives in order to travel from its origin to its destination. A discrete route choice model uses the current costs to calculate the probability P_k of choosing path k. For our purposes, we have chosen the logit model with the default values provided by the software. We have also set the simulator to compute shortest paths every 10 minutes based on the statistics collected for the last two intervals.

In order to build the historical travel time database, we have set the software to store statistics every 5 minutes in each one of the replications. The total simulation period of each replication is 10 hours (from 7:00 to 17:00) resulting in a total of 120 time intervals for which we are able to obtain detailed information about link travel times. The results have been aggregated to compute averages and standard deviations for every link and time period.

22.5.2. Customer Datasets

We have generated a dataset of 100 customers which are randomly distributed across the network (see Figure 22.7 where customers are represented as black dots).

Besides their geographic location, we also distinguish customers by the characteristics of their time windows, which can be:

- *Narrow* – with time length between 0,5 and 1 hour,
- *Wide* – with time length between 4 and 6 hours.

For both datasets, we assume that there is a logistics centre with one depot and a fleet of 7 homogenous vehicles with a capacity of 1,000 units. The depot service runs from 8am to 4pm. It is also assumed that all customers have a constant demand of 10 units and a service time of 10 minutes.

We have defined two main computational experiences for evaluating new customer orders assignments. In the first set of experiments, we evaluate the case of the dynamic VRPTW while the second set of experiments address the dynamic PDVRPTW. In order to compare scenarios, the following key performance indicators have been defined:

- service level, that is, percentage of fulfilled customers,
- total travel time of vehicles,
- total waiting time of vehicles.

22.5.3. Computational Experience I: New Customer Calls in the Dynamic VRPTW

In the first set of experiments, we evaluate the performance of the fleet in the presence of new customer orders considering various degrees of dynamism. We assume that the new customer orders arriving to the fleet manager follow an exponential distribution with a mean of 5 minutes. The sum of initial and dynamic customers is assumed to be 100. For each instance, there is a set of customers whose demand is known in advance. The initial routing plan for these customers is built using the UTS heuristic. In order to achieve a 100% service level, we have used a protection or buffer factor against variability of 20% in the estimated travel times among customers. During simulation, shortest paths calculations are made after a vehicle finishes a service and when a new order assignment takes place.

We used two approaches to assign customers to the routes: the DINS and the dynamic tabu search heuristic (DTS). In preliminary tests of the DTS heuristic, we observed that best-found solutions are usually no longer improved after 1,000 iterations of the algorithm. Therefore, we set DTS to stop the search process at 1,250 iterations or, if after 500 iterations, the solutions has not been improved in at least 1%. When a new order arrives, the simulation is temporarily stopped to calculate the new assignment. Simulation is resumed once the new solution is found and routes are updated. In all our experiments, we have assumed that diversion is allowed, that is, if a vehicle is travelling or waiting at a customer location, a vehicle can be diverted to provide service to a new customer.

In order to measure the quality of the solutions obtained by these two methods, we also evaluate the performance of the algorithms against the *value of perfect information* (VPI), which can be defined as the fleet performance that would be observed if we knew the demand in advance when planning routes. Table 22.1 shows the simulation results for the different degrees of dynamism when customers are assumed to be randomly distributed in the service area. Let us first consider the case of customers with narrow time windows. When the degree of dynamism is relatively small (20%), assigning a new order with DTS reduces about 14% the total travel time of the fleet than when using a greedy heuristic as DINS. The gap between these two methods increases along with the degree of dynamism. In problems with 80% of dynamism, we observed that DTS could improve the results obtained by DINS in 36%. A similar pattern was observed in the case of wide time windows where DTS found solution that, on average, improved the total travel time experienced by the fleet in 50%.

Regarding the number of vehicles required by both methodologies, we observed that solutions proposed by DTS were always equal or lower for all degrees of dynamism. Although the objective function in DTS is in terms of travel time, the shift operation in the iterative process drives the search to ‘more balanced’ solutions which require, on average, a lower number of vehicles.

In order to test the quality of the solutions provided by the DTS we compared them with those obtained when we assume that perfect information is available. It is expected that perfect information will lead to near-optimal solutions as routes can be more carefully planned at the beginning of the operational period. From Figure 22.8,

Table 22.1: Simulation results for customers in random locations

Customer Distribution	Degree of Dynamism	Time Windows	Solving Algorithm	Initial Routes	Final Routes	Total Travel Time (secs)	Total Waiting Time (secs)	Service Level
Random	20%	Narrow	DINS	4.0	4.0	18,400	28,465	100.0%
Random	40%	Narrow	DINS	3.0	4.0	19,625	30,795	100.0%
Random	60%	Narrow	DINS	3.0	6.0	22,629	24,711	100.0%
Random	80%	Narrow	DINS	2.0	7.0	25,926	40,032	100.0%
Random	20%	Narrow	DTS	4.0	4.0	15,892	29,918	100.0%
Random	40%	Narrow	DTS	3.0	4.1	15,890	34,610	100.0%
Random	60%	Narrow	DTS	3.0	5.5	15,678	33,171	100.0%
Random	80%	Narrow	DTS	2.0	6.3	16,672	38,600	99.9%
Random	20%	Wide	DINS	3.0	4.0	11,981	19,607	100.0%
Random	40%	Wide	DINS	2.0	3.9	16,223	10,706	100.0%
Random	60%	Wide	DINS	2.0	5.0	19,020	26,241	100.0%
Random	80%	Wide	DINS	1.0	6.3	22,205	26,478	100.0%
Random	20%	Wide	DTS	3.0	3.0	9,400	10,040	100.0%
Random	40%	Wide	DTS	2.0	3.0	9,765	12,317	100.0%
Random	60%	Wide	DTS	2.0	4.0	10,456	14,329	100.0%
Random	80%	Wide	DTS	1.0	5.0	11,071	15,795	100.0%

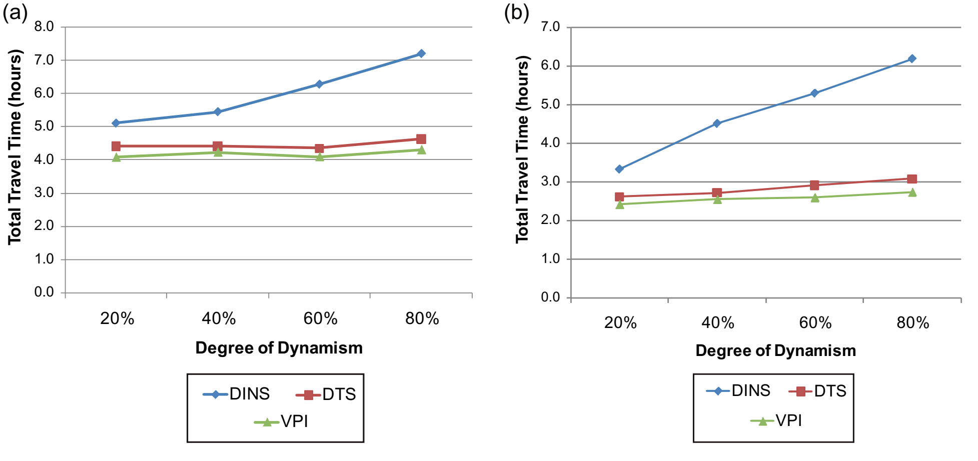

Figure 22.8: Fleet performance for different degrees of dynamism solving dynamic VRPTW (customers in random locations). (a) Narrow time windows, (b) wide time windows.

we can observe that DTS lead to relatively good results when compared to those obtained when perfect information is available. On average, DTS solutions were 6% and 8% larger than those obtained under VPI for the narrow and wide time window cases respectively. In particular, we observed that in problems with a relatively high degree of dynamism (more than 60%), DTS seems to perform better in the case of narrow time windows than in the case of wide time windows.

From the results showed above, we can state that a greedy insertion heuristic as DINS can be easily implemented and used to obtain fast results in real-time. The problem with the greedy heuristic is that the larger the degree of dynamism, the poorer is the heuristic to provide good quality solutions. DTS, on the other hand, is not as easy to implement due to the fact that it needs some calibration of parameters. Nevertheless, the quality of the solutions provided by DTS is significantly better than the one obtained with DINS when problems involve higher degrees of dynamism.

22.5.4. Computational Experience II: New Customer Calls in the Dynamic PDVRPTW

In the second set of experiments, we evaluate the performance of the fleet in the presence of new orders regarding pickup/delivery customer pairs, considering various degrees of dynamism. We assume that customers are randomly distributed across the service area. The sum of initial and dynamic customer pairs is equal to 50. For each instance, there is defined a set of pairs whose demand is known in advance. We assume that the new customer pair orders arriving to the fleet manager follow an exponential distribution with a mean of 10 minutes. As in the first set of experiments, there is used the protecting buffer factor against variability in the estimated travel times. However, since the problem is more restricted than dynamic VRPTW reaching of 100% of service level is not guaranteed. Dynamic SPI procedure was used to assign customer pairs to the routes. The routing plan is optimized by parallel DTS, which stops the search process after 500 iterations if the solution has not been improved in at least 1%. The solution is post-optimized by NPDPRO.

The time-dependent shortest path calculations take place when: a new request arrives and a vehicle finishes to service. In contrast to the first set of experiments, the shortest paths are also recalculated in the case when a vehicle arrives to a customer but cannot provide the service since the time window has been closed. Once the shortest paths are calculated and the routing plan updated the simulation is restored. Vehicle diversion is permitted.

The attributes defined in Section 22.5.2 are also valid for this set of experiments. However, due to both pairing and precedence constraints introduced in this problem, new attributes were generated. The customers were set in pairs at random. Within each pair, the customer with smaller lower-bound of the time window was set as pickup and the other as delivery. To simulate the customers' service hours, the time windows' lower bounds and durations were synchronized for each pair by applying the following rules:

- $e_p \leqslant e_d$ – the lower bound of the time window of the delivery customer e_d cannot be smaller than the lower-bound of the time window of its pickup pair partner e_p,
- $e_p + S_p + T_{p,d}(e_p + S_p) + \delta \leqslant l_d$ – the duration of the time window of the delivery customer allows for serving its pickup pair partner and directly reach the delivery location within its time window (δ is the additional time buffer value selected at random between 0 and 0.5 hour).

Table 22.2 contains the results obtained for all the tested scenarios for dynamic PDVRPTW. It shows that there is always a difference between the initially planned and finally performed routing plan. The higher is the degree of dynamism, the bigger is the difference in the number of routes and by consequence in the value of travel time.

The results also show that the smaller is the number of customers known in advance the bigger is the final travel time. This tendency is a result of the fact that as the customer pairs are added to the current routing plan the length of each route gets extended. In the scenarios with narrow time windows, when the degree of dynamism grows (from 20% to 80%), the difference in final travel time between the subsequent scenarios grows by 61%, 7% and 9%. Similarly, in the scenarios with wide time windows, the difference reaches: 31%, 5% and 23% respectively.

A decreasing tendency defines the values of total waiting time. The lower is the degree of dynamism, the higher the pressure to introduce new pairs into the existing routes. As a consequence, the routing schedules get updated and in order to achieve feasibility the total waiting times increase.

The difference in the scale of the obtained results between scenarios with narrow and wide time windows is due to the fact that it is harder to manoeuvre and relocate

Table 22.2: Simulation results for pairs of customers in random locations

Customer Distribution	Degree of Dynamism	Time Windows	Solving Algorithm	Initial Routes	Final Routes	Total Travel Time (secs)	Total Waiting Time (secs)	Service Level
Random	20%	Narrow	SPI	3.0	6.0	46,715	29,149	97%
Random	40%	Narrow	SPI	3.0	7.0	75,166	36,768	96%
Random	60%	Narrow	SPI	2.0	6.0	80,117	26,926	97%
Random	80%	Narrow	SPI	1.0	6.0	87,266	28,075	91%
Random	20%	Wide	SPI	3.0	4.0	23,104	10,376	100%
Random	40%	Wide	SPI	2.0	4.0	30,163	6,865	100%
Random	60%	Wide	SPI	2.0	4.0	31,542	6,996	100%
Random	80%	Wide	SPI	1.0	4.0	38,954	3,803	100%

the pairs of customers when the time windows width is limited. Thus, the comparison of scenarios which differ only in the time windows span specification shows, that the number of serviced customers in scenarios with wide time windows is always higher (it reaches 100% for all tested scenarios), while both the final solution cost and the total travel time values are always lower.

As before, in order to measure the quality of the solutions we evaluate the performance of the algorithms against the VPI. From Figure 22.9, we can observe that on average, in the scenarios with wide time windows, the values of travel time in SPI solutions were 13% larger than those obtained under VPI. Both SPI and VPI reached 100% of service level for all the scenarios. In the scenarios with narrow time windows, the values of travel time reached by SPI are on average 5% lower than those obtained under VPI. It is due to the fact that in most of the scenarios SPI reached lower service levels. The difference is of 3% on average. In the scenario with 40% of dynamism, both SPI and VPI reached the same service level equal to 96%. In that case, the value of travel time reached by SPI is higher than in the case of VPI. The difference is of 13%.

The above presented results show that dynamic SPI together with DTS can provide quality solutions in real-time for PDVRPTW also in the cases when addressed problems involve high degrees of dynamism. The obtained values of travel time increase in relation with growing degree of dynamism of the solved problem, since as the new customer pairs arrive the routes get naturally elongated. While in the scenarios with wide TW there is always reached 100% of service level, in the scenarios with narrow TW the large degree of dynamism affects both the service levels and the solution cost. It is due to the fact that the solved problem is challenging since it involves additional restrictions (i.e. narrow TW, pairing and precedence constraints).

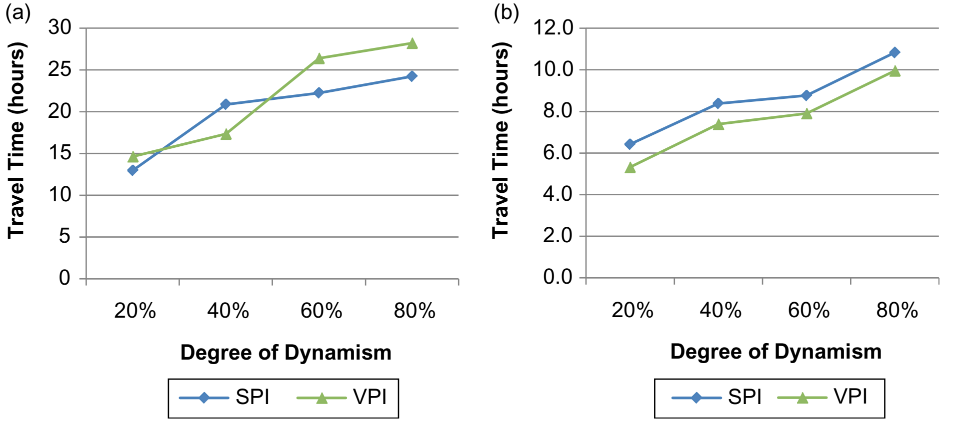

Figure 22.9: Fleet performance for different degrees of dynamism solving dynamic PDVRPTW (customers in random locations). (a) Narrow time windows, (b) wide time windows.

22.6. Conclusions and Further Research

This chapter provides an insight of the integration of vehicle routing models with real-time traffic information systems as a decision support platform. We propose a platform for the evaluation of city logistics applications in the case of the vehicle routing problems with time windows and a new approach to deal with the dynamic cases, through fast insertion heuristics and TS, when new customers' requirements arrive during the operational period. The results obtained in the simulations proved that, when creating an initial routing plan, although feasible solutions are originally obtained based on some historical average data, it may happen that not all customers will be served within their time windows due to changes in traffic conditions during the day.

Additionally, when dealing with new customer orders in the vehicle routing problem with time windows, we observe that for very low degrees of dynamism, fast insertion heuristics can be used as solving methodologies without a significant loss in the solution quality. Furthermore, when dealing with high dynamism, we observed that TS can be used to improve the solution obtained by a fast insertion heuristic. The results suggest that the proposed TS algorithms have a better performance when wide time windows are considered due to the higher degree of flexibility to shift customers from one route to another than in the case of narrow time windows.

The results obtained in this chapter assume that the underlying road network presents expected traffic conditions during the simulation. A further step on our research is the evaluation of the solving strategies when traffic conditions are suddenly altered (e.g. traffic incidents). Another issue to be investigated is the optimal calibration of parameters and data structures in order to improve the performance of the algorithms.

References

Ahuja, R. K., Orlin, J. B., Pallottino, S., & Scutella, M. G. (2003). Dynamic shortest paths minimising travel times and costs. *Networks*, *4*, 197–205.

Barceló, J., & Orozco, A. (2010). *A decision support system to assist the design and evaluation of city logistics applications with real-time information*. TRB10-3526, Presented at the 89th TRB Annual Meeting, Washington D.C.

Campbell, A. M., & Savelsbergh, M. (2004). Efficient insertion heuristics for vehicle routing and scheduling problems. *Transportation Science*, *38*(4), 369–378.

Chabini, I. (1998). Discrete dynamic shortest path problems in transportation applications. Complexity and algorithms with optimal run time. *Transportation Research Record*, *1645*, 170–175.

Chabini, I., & Dean, B. (1999). *Shortest path problems in discrete-time dynamic networks: Complexity, algorithms and implementations*. Technical Report, Massachusetts Institute of Technology, Cambridge.

Chen, H. K., Hsueh, C. F., & Chang, M. S. (2006). The real-time time-dependent vehicle routing problem. *Transportation Research Part E*, *42*, 383–408.

Cooke, K. L., & Halsey, E. (1966). The shortest route through a network with time-dependent intermodal transit times. *Journal of Mathematical Analysis and Applications*, *14*, 492–498.

Cordeau, J. F., Laporte, G., & Mercier, A. (2001). A unified tabu search heuristic for vehicle routing problems with time windows. *Journal of Operational Research Society*, *52*, 928–936.

Fisher, M. (1995). Vehicle routing. In: Ball, M.O., Magnanti, T.L., C.L. Monma, & Nemhanser, G. L. (Eds.), *Handbooks in operations research and management science. Network routing* (Vol. 8, Chapter 1, pp. 1–33). Amsterdam, the Netherlands: Elsevier.

Fleischmann, B., Gietz, M., & Gnutzmann, S. (2004). Time-varying travel times in vehicle routing. *Transportation Science*, *38*(2), 160–173.

Ghiani, G., Guerriero, G., Laporte, G., & Musmanno, R. (2003). Real-time vehicle routing: Solution concepts, algorithms and parallel computing strategies. *European Journal of Operational Research*, *151*, 1–11.

Grzybowska H., & Barcelo, J. (2011). Decision support system for real-time urban freight management. In *Proceedings of the 7th international conference on city logistics*. Retrieved from http://www.citylogistics.org/CityLogistics2011/DetailedInfoOfMallorca.html

Hall, R. W. (1986). The fastest path through a network with random time-dependent travel times. *Transportation Science*, *20*(3), 182–188.

Ichoua, S., Gendreau, M., & Potvin, J. Y. (2003). Vehicle dispatching with time-dependent travel times. *European Journal of Operational Research*, *144*, 379–396.

Lau, H. C., & Liang, Z. (2002). Pickup and delivery with time windows: Algorithms and test case generation. *International Journal on Artificial Intelligence Tools*, *11*(03), 455–472.

Law, A. M. & Kelton, W. D. (2000). *Simulation modeling and analysis*. Mc Graw Hill. Boston, New York.

Li, H., & Lim, A. (2001). *A meta-heuristic for the pickup and delivery problem with time windows*. 13th IEEE international conference on tools with artificial intelligence (pp. 160–170). November 7–9, Dallas, Texas.

Liu, H. X., Ma, W., Wu, X., & Hu, H. (2012). Real-time estimation of arterial travel time under congested conditions. *Transportmetrica*, *8*(2), 87–104.

Malandraki, C., & Daskin, M. S. (1992). Time dependent vehicle routing problems: Formulations, properties and heuristics algorithms. *Transportation Science*, *26*(3), 185–200.

Miller-Hooks, E. D., & Mahmassani, H. S. (1998). Least expected time paths in stochastic, time-varying networks. *Computers and Operations Research*, *25*(12), 1107–1125.

Miller-Hooks, E. D., & Hahmassani, H. S. (2000). Least expected time paths in stochastic, time-varying transportation networks. *Transportation Science*, *34*(2), 198–2015.

Nanry, W. P., & Barnes, J. W. (2000). Solving the pickup and delivery problem with time windows using reactive tabu search. *Transportation Research, Part B*, *34*, 107–121.

Orda, A., & Rom, R. (1990). Shortest-path and minimum-delay algorithms in networks with time-dependent edge length. *Journal of the Association for Computing Machinery*, *3*(3), 607–625.

Park, N., Okano, H., & Imai, H. (2000). A path-exchange-type local search algorithm for vehicle routing and its efficient search strategy. *Journal of the Operations Research Society of Japan-Keiei Kagaku*, *43*(1), 197–208.

Polimeni, A., & Vitetta, A. (2011). Dynamic vehicle routing in road evacuation: Route design experimentation. *WIT Transactions on the Built Environment*, *117*, 391–402.

Polimeni, A., & Vitetta, A. (2012). Optimising waiting at nodes in time-dependent networks: Cost functions and applications. *Journal of Optimisation Theory and Applications*. ISSN: 00223239.

Potvin, J. Y., Xu, Y., & Benyahia, I. (2006). Vehicle routing and scheduling with dynamic travel times. *Computers and Operations Research*, *33*, 1129–1137.

Powell, W. B., Jaillet, P., & Odoni, A. (1995). Stochastic and dynamic networks and routing. *Handbooks in Operations Research and Management Science*, *8*, 141–295.
Psaraftis, H. N. (1995). Dynamic vehicle routing: Status and prospects. *Annals of Operations Research*, *61*, 143–164.
Regan, A. C., Mahmassani, H. S., & Jaillet, P. (1996). Dynamic decision making for commercial fleet operations using real-time information. *Transportation Research Record: Journal of the Transportation Research Board*, *1537*, 91–97. doi:10.3141/1537-13
Regan, A. C., Mahmassani, H. S., & Jaillet, P. (1998). Evaluation of dynamic fleet management systems: Simulation framework. *Transportation Research Record: Journal of the Transportation Research Board*, *1645*, 176–184. doi: 10.3141/1645-22
Robusté, F., & Antón, F. R. (2005). *Logística del transporte* (Vol. 10). Barcelona: Edicions UPC.
Tang, H. (2008). Efficient implementation of improvement procedures for vehicle routing with time-dependent travel times. Transportation Research Record, 2089.
Thomas, B. (2007). Waiting strategies for anticipating service requests from known customer locations. *Transportation Science*, *41*(3), 319–331.
Thomas, B., & White, C., III. (2004). Anticipatory route selection. *Transportation Science*, *38*(4), 473–487.
Toth, P., & Vigo, D. (2002). The vehicle routing problem. Society for Industrial and Applied Mathematics, 1.
Ziliaskopoulos, A. K., & Mahmassani, H. S. (1993). Time-dependent, shortest-path algorithm for real-time intelligent vehicle highway system applications. *Transportation Research Record*, *1408*, 94–100.

Author Index

Subject Index