The Transport System and Transport Policy

The Transport System and Transport Policy

AN INTRODUCTION

Edited by

Bert van Wee
Delft University of Technology, The Netherlands

Jan Anne Annema
Delft University of Technology, The Netherlands

David Banister
University of Oxford, UK

Edward Elgar
Cheltenham, UK • Northampton, MA, USA

English language edition based on the concept and structure of, and translated in part from, Bert van Wee and Jan Anne Annema (eds) (2009), *Verkeer en Vervoer in Hoofdlijnen*, Bussum, The Netherlands: Coutinho.

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2012943188

MIX
Paper from
responsible sources
FSC® C018575
FSC
www.fsc.org

ISBN 978 0 85793 689 9 (cased)
ISBN 978 1 78195 204 7 (paperback)

Typeset by Servis Filmsetting Ltd, Stockport, Cheshire
Printed and bound by MPG Books Group, UK

Contents in brief

List of contributors xvii
List of abbreviations xix
Preface xxiii

1 Introduction 1
Bert van Wee, Jan Anne Annema and David Banister
2 The traffic and transport system and effects on accessibility, the environment and safety: an introduction 4
Bert van Wee

Part I The transport system

3 Individual needs, opportunities and travel behaviour: a multidisciplinary perspective based on psychology, economics and geography 19
Martin Dijst, Piet Rietveld and Linda Steg
4 Freight transport demand: indicators, determinants and drivers of change 51
Lóránt Tavasszy and Kees Ruijgrok
5 Land use and transport 78
Bert van Wee
6 Transport resistance factors: time, money and effort 101
Jan Anne Annema
7 Traffic flow theory and modelling 125
Serge Hoogendoorn and Victor Knoop

Part II Impacts of the transport system

8 Transport technology to reduce transport's negative impacts 163
Jan Anne Annema, Robert van den Brink and Leonie Walta
9 Accessibility: perspectives, measures and applications 207
Karst Geurs and Bert van Wee
10 Transport and the environment 227
Bert van Wee, David Banister, Jan Anne Annema and Karst Geurs

11 Traffic safety 254
Fred Wegman

Part III Transport policy and research

12 Transport policy 283
Jan Anne Annema
13 Transport futures research 305
Vincent Marchau, Jan Anne Annema, Warren Walker and Jan van der Waard
14 Appraisal methods for transport policy 329
Piet Rietveld
15 Transportation models and their applications 353
Toon van der Hoorn and Bert van Wee

Index 383

Full contents

List of contributors xvii
List of abbreviations xix
Preface xxiii

1 Introduction 1
Bert van Wee, Jan Anne Annema and David Banister

2 The traffic and transport system and effects on accessibility, the environment and safety: an introduction 4
Bert van Wee
The needs, desires, wants, preferences and choice options of people 7
Where activities take place – location 8
Transport resistance 8
Interactions between categories of factors 10
Demography 10
Travel for the fun of it 11
Goods transport 11
Technology 11
Spatial and temporal distribution of traffic and activities 12
The evaluation of policy options 13
Accessibility 13
The environment 14
Safety 14
To sum up 14
Note 15
References 15

Part I The transport system

3 Individual needs, opportunities and travel behaviour: a multidisciplinary perspective based on psychology, economics and geography 19
Martin Dijst, Piet Rietveld and Linda Steg
3.1 Introduction 19

3.2 Travel behaviour and socio-demographics 20
3.3 Conceptual model travel behaviour 24
3.4 Behavioural choice from a psychological perspective 28
3.4.1 Motivational factors: three lines of research 28
Weighing costs and benefits 28
Moral and normative concerns 30
Affect 31
An integrative perspective on environmental motivation 32
3.4.2 Contextual factors 33
3.4.3 Habitual behaviour 33
3.5 Behavioural choice from an economic perspective 35
3.5.1 Value of travel time 37
3.5.2 Price elasticity 38
3.5.3 Travel time elasticity 39
3.5.4 Income elasticity 39
3.6 Behavioural choice from a geographical perspective 40
3.6.1 Fixation in time and space 43
3.6.2 Travel time ratio 44
3.6.3 Application of geographical perspective 44
3.7 Conclusions and synthesis 45
References 47

4 Freight transport demand: indicators, determinants and drivers of change 51
Lóránt Tavasszy and Kees Ruijgrok
4.1 Introduction 51
4.2 Indicators of freight transport demand 52
4.2.1 Weight lifted 52
4.2.2 Transport performance 55
4.2.3 Traffic performance 58
4.2.4 Evolution of the different indicators 58
4.3 Logistics determinants of freight transport demand 59
4.3.1 Introduction 59
4.3.2 Production 60
4.3.3 Inventories 61
4.3.4 Transport logistics 63
4.4 Drivers of change in freight transport demand 66
4.4.1 Introduction 66
4.4.2 Economic growth 66
4.4.3 Globalization and cost changes 69
4.4.4 Mass individualization 71

4.4.5 Future logistics structures: differentiation and dynamics 73
4.5 Conclusions 75
Notes 76
References 76

5 Land use and transport 78
Bert van Wee
5.1 Introduction 78
5.2 A conceptual model for trends in passenger transport – the link with Chapter 2 79
5.3 Why should and how can land use affect travel behaviour? 79
5.3.1 The potential impacts of land use on travel behaviour: the theory of utilitarian travel demand 80
5.3.2 Key land-use variables and their impact on travel behaviour 80
Density 80
Mixed land use 81
Neighbourhood design 82
Distance to public transport connections 82
Interactions between determinants 82
5.3.3 Relationships between land-use variables, other variables and travel behaviour 83
5.4 The impact of land use on transport – a short overview of the literature 84
5.4.1 Densities 85
5.4.2 Mixed use 86
5.4.3 Neighbourhood design 86
5.4.4 Distance to public transport connections 86
5.4.5 Attitudes, lifestyles and preferences for modes, residential self-selection, and SEM models 87
5.5 Why are the conclusions different? 88
5.5.1 The research method 88
5.5.2 The level of difference in crucial factors 88
5.5.3 The geographical scale 88
5.5.4 The time horizon 89
5.5.5 Differences between countries 89
5.5.6 Indirect effects 90
5.5.7 The impact of policy 90
5.6 Evaluating the impact of land use on travel behaviour: indicators and evaluation methods 91
5.6.1 Indicators 91

Accessibility 91
The option value 92
The consumer surplus 92
Safety 92
Health impacts due to exercise 93
Environmental impacts 93
Valuation by the people 93
Financial aspects 93
Robustness 94
5.6.2 Evaluation methods 94
5.7 Evaluating the impact of land use on travel behaviour: the environment versus accessibility 95
5.8 Conclusions and discussion 97
Notes 97
References 98

6 Transport resistance factors: time, money and effort 101
Jan Anne Annema
6.1 Introduction 101
6.2 The role of travel time in passenger transport 102
Travel time components 102
Value of time 104
Constant time budgets 105
Travel time elasticities and induced traffic 108
Cross-travel time elasticities 110
6.3 The role of travel monetary costs in passenger transport 111
Constant money cost budgets 111
Price and monetary cost elasticities 112
6.4 Effort resistance factors 114
Discomfort and physical effort 115
Reliability 115
Travel information 116
Travellers' feelings of safety 116
Accident risk 117
Mental strain, stress 117
Specific constants 117
6.5 Goods transport and resistance factors 118
Transit time 118
Monetary costs 119
Transport service 120
6.6 Conclusions 121
References 121

7 Traffic flow theory and modelling 125
Serge Hoogendoorn and Victor Knoop
7.1 Introduction 125
7.2 Vehicle trajectories and microscopic flow variables 127
Gross and net headways 128
Gross and net distance headways 129
7.3 Macroscopic flow variables 129
Traditional definitions of flow, density and speed 130
Continuity equation 131
Generalized traffic flow variables 133
7.4 Microscopic and macroscopic flow characteristics 134
Headway distributions 135
Desired speed distributions 137
Gap acceptance and critical gaps 138
Capacity and capacity estimation 139
Fundamental diagrams 140
7.5 Traffic flow dynamics and self-organization 142
Capacity drop 142
Traffic hysteresis 143
Three-phase traffic flows, phase transitions and self-organization 144
7.6 Multi-lane traffic flow facilities 147
7.7 Traffic flow models 147
Safe-distance models 149
Stimulus response models 150
Psycho-spacing models 151
7.8 Macroscopic traffic flow models 151
Deterministic and stochastic queuing theory 151
Shockwave theory 153
Continuum traffic flow models 154
7.9 Network dynamics 155
Macroscopic fundamental diagram 155
Causes of network degeneration 156
7.10 Conclusions 157
References 157

Part II Impacts of the transport system

8 Transport technology to reduce transport's negative impacts 163
Jan Anne Annema, Robert van den Brink and Leonie Walta
8.1 Introduction 163
8.2 The evolution of transport technology 164

8.3 Implementing new transportation technology to solve negative impacts: a policy perspective 166
Innovation and selection towards an imperfect system 167
Lock-in and co-evolution 170
System innovations (transitions) 170
A political economy model 173
8.4 Technological innovations to reduce transport air pollution 175
The problem 175
Past and current technological changes 175
Future 179
8.5 Climate change and oil dependency 179
The problem 179
Past and current technological changes 180
Future technology 183
8.6 Noise 187
The problem 187
Road transport 187
Aviation (based on ICAO, 2008) 189
8.7 Safety 191
The problem 191
Past and current technological changes 191
Future 192
8.8 Congestion 195
The problem 195
Past and current technology 195
Pricing policies 196
Future 198
8.9 The nature of technologies and the dynamics 199
8.10 Conclusions 201
Notes 202
References 202

9 Accessibility: perspectives, measures and applications 207
Karst Geurs and Bert van Wee
9.1 Introduction 207
9.2 Perspectives on accessibility 208
9.3 Components of accessibility 209
9.4 Operationalization of accessibility measures 211
9.5 Choosing and using accessibility measures 215
Purpose of the study 215
Scientific quality 216

Operationalization (cost, ease) 217
Interpretability and communicability 217
9.6 Two examples of accessibility measures 218
Potential accessibility measures 218
Logsum accessibility measure 220
9.7 Conclusions and future trends in accessibility studies 222
References 224

10 Transport and the environment 227
Bert van Wee, David Banister, Jan Anne Annema and Karst Geurs
10.1 Introduction 227
10.2 Overview of the environmental problems of transport 228
10.3 Data and trends in transport CO_2 emissions and oil use 230
Oil dependency, peak oil 232
10.4 Data and trends in air pollution 234
10.5 Data and trends in noise 236
10.6 A comparison of environmental performance of modalities 238
10.7 Transport environmental policy instruments 240
Regulations 241
Pricing 242
Land-use planning 243
Infrastructure 244
Public transport policies 244
Marketing, education, information and communication 244
10.8 Long-term environmental challenges 244
Long-term options to reduce oil dependency and CO_2 emissions 245
Criteria to evaluate candidate policy options 247
The wider context of CO_2 policies for transport 248
10.9 Conclusions 249
Notes 250
References 250

11 Traffic safety 254
Fred Wegman
11.1 Introduction 254
11.2 Risk factors in traffic 255
Fundamental risk factors 255
Risk-increasing factors 256
11.3 Cause: ‘unintentional errors’ or ‘intentional violations’? 258

11.4 Transport modes, age groups and road types 260
11.5 Measuring safety and danger 263
11.6 Developments in road crashes 266
An example: the Netherlands 267
11.7 Shifts in road safety paradigms 269
11.8 Sustainable Safety 273
11.9 Conclusions 278
References 278

Part III Transport policy and research

12 Transport policy 283
Jan Anne Annema
12.1 Introduction 283
12.2 External effects of transport 285
External costs 285
Costs due to the use of transport means such as road vehicles, ships or aeroplanes 285
Costs due to transport means, ownership and availability 287
Costs due to infrastructure 287
External benefits 288
12.3 Maximizing welfare 288
The practice 293
12.4 Equity 293
12.5 Public choice theory 296
12.6 'Healthy' transport policies 297
12.7 Current transport policy 299
Accessibility 300
Environment and accidents 300
12.8 Conclusions 302
References 303

13 Transport futures research 305
Vincent Marchau, Jan Anne Annema, Warren Walker and Jan van der Waard
13.1 Introduction 305
13.2 Futures research and transport policy analysis 307
13.3 Level 3 approaches: forward-looking scenarios and exploratory modelling 312
Forward-looking scenarios 312
Pros and cons of the scenario approach 317
The exploratory modelling and analysis approach 317

13.4 The backcasting approach 319
13.5 Level 4 approaches: flexible and adaptive approaches 321
13.6 Conclusions 325
References 326

14 Appraisal methods for transport policy 329
Piet Rietveld
14.1 Introduction 329
14.2 Cost benefit analysis 330
14.3 Issues in SCBA 334
14.4 Valuation of environmental effects 337
14.5 Valuation of travel time 340
14.6 Multi-criteria analysis 344
14.7 Conclusions 348
Notes 350
References 350

15 Transportation models and their applications 353
Toon van der Hoorn and Bert van Wee
15.1 Introduction 353
15.2 Kinds of models 354
Descriptive versus exploratory models 354
Spatial versus non-spatial models 355
Aggregated versus disaggregated models 356
Static versus dynamic models 356
Models based on revealed preference or based on stated preference 357
Models for travel versus models for activities 357
Methods that do versus models that do not take into account the effects of transport on land use and on the economy 358
Models for passenger transport versus freight transport 359
15.3 Elasticities 359
15.4 The traditional aggregated model 361
15.5 Disaggregated models 363
15.6 Validation of models 366
15.7 Some examples of models 366
The Dutch National Model System 367
Quality of the model 372
DYNAMO 372
Activity-based model 374
Freight module in TRANS-TOOLS 375

15.8 What can and what can't a model be used for? 376
15.9 Conclusions 379
References 380

Index 383

Contributors

Jan Anne Annema, Assistant Professor of Transport Policy Analysis, Delft University of Technology, The Netherlands

David Banister, Professor of Transport Studies, University of Oxford, UK

Robert van den Brink, Senior Advisor, Goudappel Coffeng, The Netherlands

Martin Dijst, Professor of Urban Development and Mobility, Utrecht University, The Netherlands

Karst Geurs, Associate Professor of Transport Planning, University of Twente, The Netherlands

Serge Hoogendoorn, Professor of Traffic Flow Theory, Delft University of Technology, The Netherlands

Toon van der Hoorn, Emeritus Professor of Transportation Economics, Rijkswaterstaat/University of Amsterdam, The Netherlands

Victor Knoop, Assistant Professor of Traffic Flow Modelling, Delft University of Technology, The Netherlands

Vincent Marchau, Associate Professor of Transport Policy Analysis, Delft University of Technology; Managing Director, TRAIL Research School, The Netherlands

Piet Rietveld, Professor of Transport Economics, VU University Amsterdam, The Netherlands

Kees Ruijgrok, Professor of Transport and Logistics Management, TiasNimbas Business School of Tilburg University, The Netherlands

Linda Steg, Professor of Environmental Psychology, University of Groningen, The Netherlands

Lóránt Tavasszy, Professor of Freight and Logistics/Principal Scientist, Delft University of Technology/TNO, The Netherlands

Jan van der Waard, Programme Manager, KiM, Netherlands Institute for Transport Policy Analysis, The Netherlands

Warren Walker, Professor of Policy Analysis, Delft University of Technology, The Netherlands

Leonie Walta, Delft University of Technology, The Netherlands

Bert van Wee, Professor of Transport Policy, and Scientific Director of TRAIL Research School, Delft University of Technology, The Netherlands

Fred Wegman, Professor of Traffic Safety, Delft University of Technology/ Institute for Road Safety Research, The Netherlands

Abbreviations

ABS	antilock braking systems
ACC	adaptive cruise control
ADAS	advanced driver assistance systems
ANPR	automatic number plate recognition
BAC	blood alcohol content
BPR	bypass ratio
CA	capability approach
CAFE	Corporate Average Fuel Economy
CBA	cost–benefit analysis
CO	carbon monoxide
CO_2	carbon dioxide
CS	consumer surplus
DALY	disability adjusted life years
DAP	dynamic adaptive policies
dB	decibel
DSRC	dedicated short range communication
EEVC	European Enhanced Vehicle-safety Committee
EGR	exhaust gas recirculation
EMA	exploratory modelling and analysis
EU	European Union
EV	electric vehicle
FCV	fuel cell vehicle
GDP	gross domestic product
GHG	greenhouse gas
GIS	geographic information systems
GNP	gross national product
GNSS	global navigations satellite system
GPRS	general packet radio system
GTC	generalized transport costs
H_2	hydrogen
ICAO	International Civil Aviation Organization
ICE	internal combustion engine
IEA	International Energy Agency
IMO	International Maritime Organization
IPCC	Intergovernmental Panel on Climate Change

ISA	intelligent speed adaptation
ITS	intelligent transport systems
LCA	life cycle assessment
LMS	Landelijk Model Systeem [National Model System]
LUTI	land-use transportation interaction
MASTIC	model of action spaces in time intervals and clusters
MCA	multi-criteria analysis
MNL	multinomial logit
MPG	miles per gallon
NAM	norm activation model
NCAP	New Car Assessment Programme
NFD	network fundamental diagram
NO	nitrogen oxide
NOA (model)	needs, opportunities, abilities
NO_x	nitrogen oxides
NST/R	Nomenclature uniforme des marchandises pour les Statistiques de Transports, Revisée
OBD	on-board diagnostic systems
OBU	on-board unit
O–D	origin–destination
OECD	Organisation for Economic Co-operation and Development
PBC	perceived behaviour control
PCV	positive crankcase ventilation
PEM	proton exchange membrane
PESASP	programme evaluating the set of alternative sample path
PM	particulate matter
PPP	public–private partnership
PT	public transport
QALY	quality adjusted life years
RDM	robust decision making
RFID	radio-frequency identification
RP	revealed preference
SCBA	social cost benefit analysis
SCR	selective catalytic reduction
SEM	structural equation model
SO_2	sulphur dioxide
SP	stated preference
SPITS	sources, production, inventories, transport and sales
TPB	theory of planned behaviour
TTB	travel time budgets
UFOV	useful field of view
UK	United Kingdom

UMTS	universal mobile telecommunications system
US	United States
VOC	volatile organic compounds
VOR	value of reliability
VOSL	value of statistical life
VOT	value of time
VTTS	value of travel time savings
WHO	World Health Organization
WTA	willingness to accept
WTP	willingness to pay
ZEV	zero-emission vehicles

Preface

Bert van Wee, Jan Anne Annema and David Banister

To the best of our knowledge a multidisciplinary book that introduces the reader to the transport system, its effects on accessibility, the environment and safety, and policy and related research does not yet exist. Other books generally discuss transport from the viewpoint of one discipline, for example economics, or focus on one policy-relevant effect, for example the environment. This book aims to fill the gap.

The book is written primarily for educational purposes, for use either in courses at universities or in other education programmes or self-study. We realize that we do not cover all of the expectations that some readers would like to see included. We only aim to give a general introduction to the transport system, its effects on society and policy.

This book is based on a Dutch language book (*Verkeer en Vervoer in Hoofdlijnen*) that was published in 2002 and updated in 2009. The 2002 book was sold out within seven years (2500 copies) and is used at several universities and in several professional education programmes. We did not just translate the book, but replaced most Dutch cases and examples with ones from other countries, added international literature and added a chapter.

As well as the hard copy version, this book can be downloaded from the internet. We hope that this will stimulate courses to use the book and that we are thereby making a small contribution to reducing the costs of education. We are very glad that Edward Elgar has the courage to publish the hard copy version of this book. We especially thank Emily Neukomm for the very pleasant contacts. We hope that at least staff members and libraries will buy the hard copy version – and maybe lots of students as well.

This book is made possible thanks to the Dutch transport research school TRAIL (www.rstrail.nl), which supported us financially and by doing a lot of work related to layout, editing, indexing and the like. We especially thank Dr Vincent Marchau and Professor Ben Immers (directors of

TRAIL) and Conchita van der Stelt (TRAIL) for their support and for all their effort.

If you have any suggestions related to this book in general or parts of it or have questions, please contact us.

1

Introduction

Bert van Wee, Jan Anne Annema and David Banister

This book aims to give a general introduction to the transport system (the factors that drive this system), the impact the transport system has on accessibility, safety and the environment, and transport policy making and evaluations.

The target group of this book is people who want a general introduction to the transport system and its effects on society. We primarily have students at universities and colleges (bachelor's and master's level) in mind but also people who already have a degree in another area, and thus are not educated in the area of transport. In addition, we aim to introduce the field of transport to Ph.D. students who do research into transport but who do not have a background in transport.

An important basic principle is that we think that the transport system and its impacts on society can best be understood by combining the insights of multiple disciplines, including civil engineering, economics, psychology and geography. As a result this book is a multidisciplinary book. The background of the authors included in this book reflects this basic principle.

Here's a short explanation of the way we have structured this book. After this introductory chapter, Chapter 2 aims to explain the structure of the book by presenting a conceptual framework for factors having an impact on transport volumes and the impact of the transport system on the environment, accessibility and safety. It explains how factors like the wants and needs of people and companies, the land-use system, and the overall resistance to travel (cost, time and effort) interact, and how these factors, combined with technology and the way people use vehicles, drive accessibility, safety and the environment.

In Part I (Chapters 3–7) we describe how the transport system is structured. Chapter 3 explains how the wants and needs of people drive passenger transport. It explains that most people's trips result from them wanting to participate in activities located in different places. Activity and travel patterns result from the wishes, possibilities and constraints of people. Chapter 4 is its equivalent for goods transport. It describes the needs of companies to transport goods. It concludes that goods transport is very heterogeneous: all kinds of goods, ranging from computer chips to raw materials, have to be transported for all kinds of markets using modes such as lorries, vans, barges, pipelines, aircraft and sea-going ships. Those goods are produced by all kinds of producers, and bought by a wide range of clients, ranging from other producers to customers. The huge heterogeneities result in complex relationships between the economy, society and logistics. Chapter 5 focuses on the impact of the land-use system on transport. It explains that land use can have an impact on transport, dominant land-use factors being densities, the level of mixed use, neighbourhood design and distances between origins and destinations, and public transport nodal points such as railway stations. Chapter 6 introduces the reader to how transport resistance (time, costs and effort) influences transport. It explains how changes in infrastructure, prices and other resistance components have steered passenger and goods transport. Chapter 7 then focuses on one part of transport resistance by describing traffic flows in more detail. It describes traffic flows and explains how capacities of networks, demand, and congestion levels are interrelated.

Part II, Chapters 8–11, discusses the impacts of the transport system on accessibility, the environment and safety. Because such impacts not only result from the factors discussed in Chapters 2–7 but are also due to the technology used in the transport system, Chapter 8 firstly discusses this technology. Chapter 9 presents a definition of accessibility and several accessibility concepts, including the pros and cons of their use, depending on the purpose of use. Chapter 10 gives an overview of environmental impacts of the transport system, and concludes that transport probably is the sector having the largest negative impact on the environment. Chapter 11 discusses transport safety as well as factors having an impact on safety, dominant clusters of factors being infrastructure, the user of infrastructure, and vehicles.

Part III, Chapters 12–15, gives an introduction to transport policy and related research. Chapter 12 introduces the reader to the reasons why governments develop transport policies, and how policy tasks range from local to (inter) national governmental bodies. Chapter 13 gives an overview of methods to explore the future, including the area of forecasting and backcasting. Chapter 14 discusses cost–benefit analysis (CBA) and multi-criteria analysis (MCA), the two most important methods to evaluate candidate policy options *ex ante*. To conclude, Chapter 15 discusses transport models and their applications, models that are widely used *ex ante* to evaluate changes in travel demand and travel times that result from policy options such as changes in infrastructure, land use and pricing. Such changes in travel demand and travel times are input for CBAs and MCAs.

The parts and chapters are in a logical order: firstly the transport system is described, next policy relevant effects, and then research and policy. This is based on the assumption that the reader will read all the chapters in the given order. However, because we realize that not all readers will have read all the previous chapters we have tried to write the chapters in such a way that they can be read more or less independently. To avoid overlap we cross-refer between chapters.

2

The traffic and transport system and effects on accessibility, the environment and safety: an introduction

Bert van Wee

People travel because they want to carry out activities such as living, working, shopping and visiting friends and relatives at different locations. Goods are transported because several stages of production are spatially separated. For example, components for cars may be produced at different locations, whereas the assembly is in the main factory. Cars finally have to be transported to distribution centres in several countries, and to dealers where people can buy them.

Developments in transport are relevant for several reasons. Firstly, without transport modern societies would not be possible at all. Because no reasonable person would question the absolute relevance of transport, what matters more is the impact of *changes* in the transport system on *changes* in the economy or the wider society. Secondly, transport causes negative impacts: environmental pressure, safety impacts, and congestion being the three most important negative impacts. Thirdly, developments in transport trigger policies in several areas, including infrastructure planning, land-use planning, pricing policies and subsidies, and regulations with respect to safety (such as maximum speeds or the crash-worthiness of vehicles) or the environment (such as emissions standards for pollutants, CO_2 and noise). Therefore many questions are relevant for both researchers and policy makers, such as: Which determinants have an impact on traffic and transport? Which determinants have effects on the environment, accessibility and safety? This chapter deals with such questions. Its goal is to provide an overview of the subjects that are dealt with in the next chapters and of the relationships between these subjects.[1] In this chapter we firstly give a general overview of factors having an impact on transport as well as of factors having an impact on the environment, accessibility and safety. We then further elaborate on these factors.

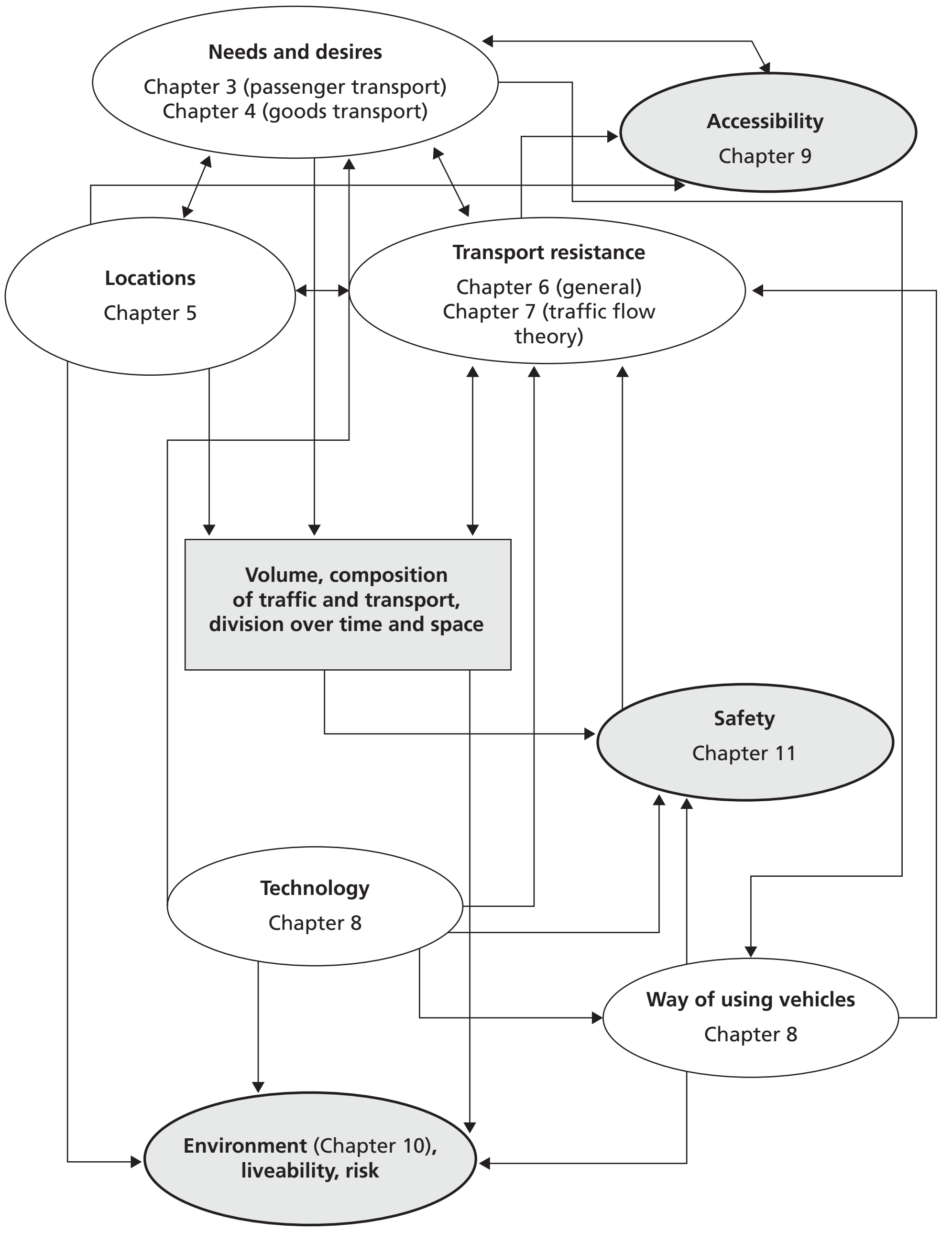
Given the population size and its decomposition by household class and age, transport volumes and their decomposition by modes and vehicle types result from:

1. the wants, needs, preferences and choice options of people and firms;
2. the locations of activities such as living, working and shopping;
3. transport resistance, often expressed in time, money, costs and other factors, which we refer to as 'effort' and which include, among others, risks, reliability of the transport system and comfort.

The three factors also have an impact on accessibility. In this book we define accessibility as:

> The extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s) at various times of the day (*perspective of persons*), and the extent to which land-use and transport systems enable companies, facilities and other activity places to receive people, goods and information at various times of the day (*perspective of locations of activities*). (See Chapter 9)

Technology and people's driving behaviour (as expressed by speed and acceleration/deceleration behaviour) have an impact on travel times and travel comfort (components of transport resistance) and therefore on accessibility, and they also have an impact on safety and the environment. Driving behaviour is influenced by people's preferences. Not only does driving fast reduce travel times but people may actually like it. The distribution of traffic over space and time also has an impact on safety, the environment and accessibility. The distribution over space includes the breakdown between traffic within and outside the built-up area and by road class. For example, traffic on a road along which hardly any houses are sited causes less noise nuisance compared to traffic on a road along which many houses are located close together. Concentrations of pollutants on pavements are higher if the pavement is located near a (busy) road. For the distribution over time, the breakdown by hour of the day is very relevant for the impact of traffic on noise nuisance, since night traffic causes much more noise nuisance than daytime traffic. On the other hand, night traffic causes hardly any congestion. Figure 2.1 visualizes these factors and their mutual relationships. Only dominant relationships are included.



Note: The corresponding chapters of this book are depicted in the figure. Thin-lined ovals with no filling are determinants for transport and traffic volume and effects. Thicker-lined ovals with shaded filling are transport and traffic effects. The rectangle represents transport and traffic volume and composition divided over time and space.

Figure 2.1 A conceptual framework for factors having an impact on transport volumes and the impact of the transport system on accessibility, the environment and safety

We will now elaborate on the factors and relationships as presented in Figure 2.1. We firstly focus on passenger transport and then briefly reflect on the transport of goods.

The needs, desires, wants, preferences and choice options of people

A part of people's wants and needs has to be fulfilled at different locations. Of course, wants, needs and preferences vary strongly between people (see Chapter 3). For example, young people may prefer to go to pubs more than older people. Economists often relate this factor to income: if someone's income rises, she or he may fulfil more needs, especially those that have a higher price. People with high incomes have more money to spend on holidays or visiting the theatre. Note that not all needs have a price. For example, a walking trip is free.

It is not only economists who pay attention to this subject; psychologists do too. They conclude that, by buying a car, people can fulfil their needs with respect to status, power and territory drifts (the desire to 'cover' a certain space). Although status might be less important now than a few decades ago, it still has an impact on vehicle choice. A new Mercedes or Lexus gives more status than an old Toyota Corolla. By pushing the throttle, a car driver controls power, which might result in a good feeling. If people park their car in front of their house, they have the feeling of expanding their territory (for the impact of symbolic and affective factors, see, for example, Steg et al., 1998; Steg, 2005).

Money poses constraints on people's choice options, as does time. Of course, all people have 24 hours a day to spend, but the time people need for different activities varies strongly between individuals and depends, among other things, on work- and family-related constraints. People working full time have less free time to spend than those working part time. People raising children may need more time when they are young. Mainly, economists and geographers pay attention to the impact of time on activity patterns and travel behaviour.

The role the transport system plays in fulfilling people's needs depends on time and space. To illustrate the impact of time: at the turn of the twentieth century, when in some countries a man with a red flag walked in front of a car, the car was a first-class status object. Now many people really need a car. To illustrate the impact of space: the aeroplane is a rather common means of transport in the US, especially for long distances, but in third world countries it is an option for only a very small fraction of the population.

Apart from generally recognized factors such as income, age, sex and household structure, lifestyle factors and related preferences and attitudes have an impact on travel behaviour (Kitamura et al., 1997; Cao et al., 2009). Finally, it should be noted that the choice options that people have vary between individuals. Not all adults have a driving licence or the physical ability to walk over longer distances. Some people have time constraints because they have to combine tasks, reducing their choice options for activities.

Where activities take place – location

Another category of factors affecting transport is the location of activities (see Chapter 5). As we have explained above, transport is needed to allow people to fulfil activities at different places or to transport goods between different locations. Therefore transport volumes depend on these 'locations of activities'. In this context only location-related activities are relevant. Some activities such as mobile phoning, brushing one's hair or thinking about the next holiday destination are not location-related and therefore not relevant for transport. It is not only the distribution of activities over space that is relevant, but also the distribution of people over houses, workplaces and other destinations. If people living in London work in Cambridge or vice versa, much more transport results than if people who live in London or Cambridge also work in the same location. Therefore, what can be seen on the map matters (land-use patterns), as well as the functional relationships between the locations of living, working, services and so on. Spatial scientists such as geographers and planners look at transport from this viewpoint.

Transport resistance

A third category of factors relevant for developments in transport is the resistance needed to travel between locations, including travel time, monetary costs and other aspects such as comfort and safety (see Chapters 6 and 7). The sum of these costs is often referred to as generalized transport costs (GTC). Lower GTC results in more transport. First, GTC depends on the quality and quantity of infrastructures of all types (roads, rail, rivers and canals, airline and port connections). Second, traffic volumes at a certain infrastructure section related to its capacity are relevant: if demand exceeds capacity, congestion occurs and this results in longer travel times. Third, infrastructure-related regulations have an impact on GTC, especially maximum speeds. Fourth, the characteristics of vehicles matter, especially the comfort levels and costs. Fifth, safety levels matter, and they depend on the infrastructure and vehicle characteristics and the way people use vehicles (driving style). Costs of private and public transport have an impact on GTC.

If we look at the time component, we see that, owing to significant motorway expansion over the last few decades, travel times between cities and towns have strongly decreased. Owing to the increase in the number of airline connections, travel times by plane between many destinations are now much shorter compared to a few years ago.

For monetary costs many people have the perception that fuel costs are dominant. These depend not only on fuel prices but also on the fuel efficiency of vehicles and on the share of fuel types (for cars: petrol and diesel). Fuel efficiency expresses how far one can drive with a certain volume of fuel (often expressed as miles to the gallon, kilometres per litre or litres/100 km). Other variable costs are maintenance and repair costs. Variable costs are related to the amount of kilometres or miles travelled. Fixed costs are independent of the amount travelled and include the purchase price of cars combined with average age at the time of scrapping, and insurance costs and taxes. The average age of cars has increased significantly during the last few decades. Whereas in many western countries in the 1970s a large majority of scrapped cars were less than 10 years old, now in the same countries cars generally last on average at least 15 years. If cars last longer, their (yearly) fixed costs decrease. In the last few decades prices of airline tickets have decreased strongly, allowing an increasing number of people to fly and allowing the same people to travel more.

Although time and costs have an important impact on transport resistance, these are not the only factors. Factors such as comfort, reliability of travel times, and safety also play a role. Cars are now much more comfortable than those in the past owing to better noise insulation, seats, handling, suspension and springs. The chance of getting killed in an accident has strongly decreased in the last few decades. Between 1990 and 2009 in the EU, the number of people killed in road accidents decreased from 76 000 to 34 500, despite the increase in road traffic (EC, 2010). Flying also is much safer now than in the past. In 2007 and 2008 worldwide, the chance of getting killed in a plane crash was about 0.01 fatality per 100 million passenger kilometres, whereas in all years from 1986 to 1998 this chance was between 0.05 and 0.09 (US Census Bureau, 2010). Internet (stationary or mobile) based travel information services make finding information about the possibilities of travelling by public transport much easier than in the past.

Several disciplines study transport resistance. Economists mainly consider time and monetary costs. On average people seem to have a constant travel time budget (see Chapter 6). Therefore, if average travel speeds double, for example because of better infrastructure, distances travelled will also double. Civil engineers focus on infrastructure and its impact on travel times and therefore transport volumes. Geographers study the impact of time- and space-related constraints and the impact transport resistance has on these constraints. Social scientists consider psychological, sociological and cultural factors in relation to transport.

Interactions between categories of factors

All three types of factors described earlier (needs, wants, desires and preferences; location of activities; and transport resistance) have an impact on transport volumes. In addition, they also have an impact on each other, in all directions (see Figure 2.1). To illustrate this: recently, in many countries, offices have relocated from central locations to the edge of town, often close to motorways. Therefore accessibility by public transport has decreased whereas accessibility by car has increased. In other words: changes in location have an impact on transport resistance of travelling by car and public transport, and this may result in an increased desire to own a (second) car. We give two more examples. Firstly, as a result of more and more flight connections to several destinations and cheaper flights, tourist facilities were developed at many locations that probably would not have been developed assuming no improvements in the airline network and no price decreases. Secondly, as the road network has strongly improved in many countries, commuting distances have also increased. To summarize: a lower transport resistance results in new locations for activities and increases in distances travelled.

Because all three categories of factors change continuously, a stable equilibrium does not exist. The relationships between factors also imply that a policy focusing on one of the factors may have several indirect effects. For example, the direct and short-term effect of higher fuel prices is that people will reduce car use, for example by changing to other modes or choosing closer destinations. An indirect effect that occurs in the longer term is that people might move to a house closer to their job.

Demography

So far we have assumed a constant population size and composition. Of course, demographic changes also have an impact on transport volumes. By composition we refer to factors such as age and household classes (for example, single-person households, a couple without children, families with children).

Travel for the fun of it

We also assumed that people travel to fulfil activities at several places. From this viewpoint travel is demand derived. But some people also travel for the fun of it (see, for example, Salomon and Mokhtarian, 1998; Mokhtarian and Salomon, 2001). For some people travel is a form of recreation, examples being recreational car trips for tourists or cycling for recreation. In this book we do not pay any further attention to this type of travel.

Goods transport

So far we have mainly paid attention to passenger transport. For goods transport, speaking in general terms, the same categories of factors that have an impact on passenger transport also have an impact on goods transport: volumes of goods transport, expressed in tonne kilometres per mode, and traffic volumes, expressed in kilometres per vehicle type, result from the locations of activities that generate goods transport, the wants and needs of producers and consumers, and transport resistance. The relationships between these factors are also relevant (see Chapter 4). For example, in many western countries the improvements in the road network have resulted in a decrease in transport costs and therefore also in other logistical concepts (such as the 'just-in-time' concept) and other location choices of firms, and other spatial patterns of origins and destinations of goods transport. Logistical choices include, amongst others, the trade-off between supplies and transport and the number and location of distribution centres for a certain firm. Spatial effects include, for example, the location of the production of car components and the assembly of the cars.

Technology

The technologies applied in transport include both those for vehicles and those for infrastructure. They may have an impact on transport volumes. For example, more fuel-efficient cars result in lower fuel costs and may therefore lead to an increase in car use (Goodwin et al., 2004). Technology also has an impact on the environment, safety and accessibility. For example, despite the growth in car use in many countries, the emissions of pollutants such as carbon monoxide (CO) and nitrogen oxides (NO_x) have decreased. Between 1990 and 2007 transport-related emissions of particulate matter decreased by 30 per cent, acidifying substances by 34 per cent and ozone precursors by 48 per cent across the 32 EEA member countries (EEA, 2010). During the last few decades the active and passive safety of cars has improved significantly, contributing to the decrease in people killed in road accidents in many countries, as mentioned above. Active safety relates to the possibilities of avoiding crashes, passive safety to the possibilities of reducing the impact of crashes once they take place. For active safety the quality of brakes and tyres is relevant; for passive safety factors such as airbags and crash performance are relevant. Technology can also have an impact on accessibility. For example, owing to traffic lights regulating the volumes and timing of cars entering the motorway network, congestion levels on motorways have decreased. In the future, technologies such as intelligent speed adaptation (ISA), lane departure warning systems, and technologies that allow cars to drive at high speeds at close distances may be introduced (see also Chapter 8). These technologies may increase the capacity of the motorway network and reduce congestion levels on these roads significantly, and may make the road system safer. Another example: porous asphalts increase visibility during rain or when surfaces are wet and thereby increase safety, and also reduce noise emissions.

Not only are the technologies used relevant, but also the way people use them. Firstly, the way vehicles are used is relevant for environmental impacts. Emissions are related to the way vehicles are used. For example, the emissions per kilometre of carbon dioxide (CO_2, which causes climate change) and nitrogen oxides (NO_x, which causes acidification and poor air quality) of an average passenger car are much higher at 140 km/h than at 80 km/h. Driving during congestion, including frequent acceleration and braking, results in lorries producing higher emissions per kilometre of NO_x and particulate matter (PM) (which also causes poor air quality) (Schallaböck and Petersen, 1999). Cars produce more noise emissions at 120 km/h than at 50 km/h. Secondly, travel times and transport resistance are related to speed. Road capacity is higher if cars drive at 90 km/h than if they drive at 140 km/h (see Chapter 7). Travel times, and thereby travel resistance, decrease if speeds are higher. Thirdly, safety is related to driving behaviour. A main effect is that accident risks increase with speed. In summary, the way vehicles are used is related to environmental and safety impacts and to transport resistance.

Spatial and temporal distribution of traffic and activities

Given a certain volume of traffic, the spatial distribution of traffic has an impact on congestion, safety and the environment. The spatial distribution includes the breakdown of road class, for example into motorways, other rural roads and urban roads. Cars and lorries driving on urban roads cause more noise nuisance and health impacts related to emissions of pollutants compared to vehicles driving in non-urban areas. And, if vehicles are travelling on urban roads, the negative impacts are related to the number of dwellings close to the roads. Therefore it is not only the spatial distribution of traffic that is of importance but also the spatial distribution of the activities of people, and how they are located in relation to the roads. Finally, the temporal distribution of traffic is relevant. Night traffic causes more noise nuisance than daytime traffic. A more balanced distribution of traffic over time causes less congestion.

The evaluation of policy options

National, regional and local authorities make transport policy, as do unions of countries, such as the European Union (EU). Many policy options are available to change the transport system, varying from building new infrastructure to changing public transport subsidies (see Chapter 12). This raises the question of how to assess these options. Because many impacts of transport are related to the location of activities, and thus land use, the assessment of transport policy options can often best include land-use patterns and transport-related land-use policies. Ex ante evaluations should include, as much as possible, all relevant positive and negative impacts, and should compare outcomes to goals and government targets (see Chapter 14). Positive impacts (benefits) include accessibility and travel time benefits. Negative impacts (costs) include both financial and non-financial costs and external effects (effects the user does not include in his or her decision), such as environmental, congestion and safety impacts. Apart from these more general costs and benefits, governments may have regional objectives. They may, for example, strive for a more equal distribution of incomes and economic development.

Accessibility

In many countries, regions, cities and towns, improving accessibility is an important government goal. Many definitions of accessibility exist. In this book we define accessibility as 'the extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s) at various times of the day' (perspective of persons), and 'the extent to which land-use and transport systems enable companies, facilities and other activity places to receive people, goods and information at various times of the day' (perspective of locations of activities) (see Chapter 9). According to this definition the level of accessibility depends on the location of activities, quality and quantity of infrastructures, and needs of people and companies. The level of accessibility has an impact on the economy, because a well-functioning transport system in combination with the land-use system is a condition *sine qua non* for economic development. Accessibility is not only relevant for the economy but also fulfils a social role. People appreciate the ability to visit relatives and friends within certain time budgets. Even though these trips do not or hardly affect GDP or unemployment levels, people value these trips positively. Welfare economics include such wider (non-GDP-related) benefits.

The environment

In many countries, including the wider EU, reducing the environmental impacts of transport is an important policy goal. Transport is a major contributor to environmental problems. In many western countries the share in CO_2 emissions is around 20–25 per cent (in the US as much as 33 per cent – see Chapter 10), and the share in other pollutants such as NO_x, CO, volatile organic compounds (VOC) and PM varies between 30 and 75 per cent (see statistics of the European Environment Agency for European data, or Davis et al., 2010, for US data; see also Chapter 10). Other environmental impacts include negative visual effects, the barrier effects of infrastructure for humans and animals, noise nuisance and local environmental (liveability) impacts resulting from moving and parked vehicles, regardless of their exhaust and noise emissions (such as the fact that in many places children cannot play on the streets any more).

Safety

In almost all countries, the safety impacts of transport are considered to be a major problem. One can distinguish between internal and external safety. Internal safety is related to the risks of being mobile. External safety refers to the risks for the population of being the victim of an aeroplane crash, explosions due to the transport of hazardous substances, or air quality problems due to accidents with vehicles transporting hazardous gases or liquids. As explained above, in most western countries accident risks have decreased sharply (see Chapter 11), more than compensating for the increased levels of mobility or vehicle kilometres. Despite the positive trends in the EU, there were still over 34 500 people killed in road accidents in 2009.

To sum up

In this chapter we have presented a general description of the transport system and impacts on accessibility, the environment and safety using a conceptual model. In the following chapters we will describe this model in more

Table 2.1 The chapters in this book related to the ovals in Oval in Figure 2.1

Oval in Figure 2.1	Chapter
Needs and desires of people (passenger transport)	3
Needs and desires of companies (goods transport)	4
Locations	5
Transport resistance – general	6
Transport resistance – traffic flow theory	7
Technology	8
Way of using vehicles	8
Accessibility	9
Environment	10
Safety	11

detail. The model forms the basis of the structure of Parts I and II of this book. Table 2.1 explains the links between the model and the book chapters.

As explained in Chapter 1, Part III of the book discusses transport policy and related research, and includes all aspects of the system as conceptualized in Figure 2.1.

NOTE

1 In this chapter we have limited the number of references. For more references relevant to the contents of this chapter we refer to the following chapters.

REFERENCES

Cao, X., P.L. Mokhtarian and S. Handy (2009), 'Examining the impact of residential self-selection on travel behavior: a focus on empirical findings', *Transport Reviews*, 29 (3), 259–395.

Davis, S.C., S.W. Diegel and R.G. Boundy (2010), *Transportation Energy Data Book*, Oak Ridge, TN: Oak Ridge National Laboratory.

EC (2010), 'Road safety', accessed 25 December 2010 at http://ec.europa.eu/transport/road_safety/specialist/statistics/index_en.htm.

EEA (European Environment Agency) (2010), 'Towards a resource-efficient transport system. TERM 2009: indicators tracking transport and environment in the European Union', *EEA Report 2010/2*, Copenhagen: EEA.

Goodwin, P., J. Dargay and M. Hanly (2004), 'Elasticities of road traffic and fuel consumption with respect to price and income: a review', *Transport Reviews*, 24 (3), 275–292.

Kitamura, R., P.L. Mokhtarian and L. Laidet (1997), 'A micro-analysis of land use and travel in five neighbourhoods in the San Francisco Bay Area', *Transportation*, 24, 125–158.

Mokhtarian, P.L. and I. Salomon (2001), 'How derived is the demand for travel? Some conceptual and measurement considerations', *Transportation Research Part A*, 35 (8), 695–719.

Salomon, I. and P.L. Mokhtarian (1998), 'What happens when mobility-inclined market segments face accessibility-enhancing policies?', *Transportation Research D*, 5 (3), 129–140.

Schallaböck, K.-O. and R. Petersen (1999), 'The spread of congestion in Europe', in *CEMT Round* Table 1*10*: *Traffic Congestion in Europe*, Paris: European Conference of Ministers of Transport, pp. 5–44.
Steg, E.M., A.B. Brand, A.J. Rooijers and C.A.J. Vlek (1998), *Diepere drijfveren van het autogebruik, Deel 2: Theoretische conclusies, doelgroepsegmentatie, en beleidsimplicaties*, Groningen: Rijksuniversiteit Groningen, Centrum voor Omgevings- en Verkeerspsychologie.
Steg, L. (2005), 'Car use: lust and must. Instrumental, symbolic and affective motives for car use', *Transportation Research Part A: Policy and Practice*, 39 (2–3), 147–162.
US Census Bureau (2010), 'The 2010 statistical abstract', accessed 25 December 2010 at http://www.census.gov/compendia/statab/2010/tables/10s1041.xls.

Part I

The transport system

3

Individual needs, opportunities and travel behaviour: a multidisciplinary perspective based on psychology, economics and geography

Martin Dijst, Piet Rietveld and Linda Steg

3.1 Introduction

Reducing the use of the private car and stimulating the patronage of public transport to combat congestion and climate change are among the objectives that have prominent positions on the policy agendas of Western countries. Stimulating sustainable development and improving accessibility often underlie these policy objectives. Despite the efforts of scientists, policy makers and public and private investors, many transport problems seem hard to solve. This might be due to the following factors. Firstly, economic growth and increase in population size and expansion of the number of households are relatively autonomous processes which are largely outside the control of policy makers. These could have a tremendous impact on the volume of traffic. Secondly, many interventions in transport are not sufficiently tuned in to each other and as a consequence suffer in their effectiveness. For example, investments in public transport are often not accompanied by pricing measures. Congestion charging in London (Litman, 2006) and Stockholm (Schuitema et al., 2010) show how effective such a combination of policies can be. Finally, owing to the routine-based character of travel behaviour (Aarts et al., 1998), changes in car use are only likely when stringent policies are implemented.

We have to realize that developments in transport are not only the sum of choices of individuals and their households but also of firms and other organizations where they are employed. To combat undesirable aggregate effects, changes in behaviour are necessary. To increase the effectiveness of policies a thorough understanding of the travel behaviour of people is needed. Often this knowledge has an ad hoc and a monodisciplinary character. However, the fields of psychology, economics and geography all have very different perspectives on travel behaviour. A complementary approach to the needs of people and their opportunities to change behaviour might lead to the development of more effective transport policies. It is the aim of this chapter to discuss the various disciplinary perspectives and to show the connections. We will ask questions like: Why do we travel? What are the drivers for travel behaviour? What is the relationship between travel behaviour and other mobility choices, such as the choice of housing, work and ownership of transport modes? Which factors influence these choices?

Based on the contributions from psychology, economics and geography, section 3.3 will provide a conceptual framework for travel behaviour research. Successively, in sections 3.4, 3.5 and 3.6 we will discuss separately the perspectives of these scientific disciplines on travel behaviour. In section 3.7 the major conclusions of this chapter will be presented. However, first we will present and discuss in section 3.2 a brief overview of some travel behaviour attributes for a selection of socio-demographics.

3.2 Travel behaviour and socio-demographics

In this section we present some general characteristics of a selection of socio-demographics on car-driving licence holders, car-ownership and travel behaviour. The data are taken from the National Travel Surveys of the United Kingdom (NTS) and the Netherlands (MON). We limited ourselves to data publicly available on websites to offer readers the opportunity to make their own (limited) calculations to update the figures in the tables and to make comparisons over the years. The Dutch website Statline from the Central Bureau of Statistics (statline.cbs.nl) especially offers good opportunities to calculate and to download a large variety of core mobility indicators. The data we present in this section should be treated with caution, since the travel surveys differ in sample size and composition, survey methods and operationalization of variables. Nevertheless they give us a good impression of (dis)similarities between socio-demographic categories in both countries.

Table 3.1 gives an overview of possession of driving licences and cars in both countries. It shows that in the UK as well as in the Netherlands a larger proportion of men are holders of a driving licence. Although, compared to the UK, in the Netherlands a larger proportion of women have a driving

Table 3.1 Holders of car-driving licences and car ownership in the United Kingdom and the Netherlands by socio-demographics

Socio-demographics		United Kingdom (%)	The Netherlands (%)
Car-driving licence holders:			
Gender			
Men		80.0	90.2
Women		65.0	75.4
Age UK in years	Age Netherlands in years		
17–20	18–19	36.0	32.6
	20–24		72.6
21–29	25–29	65.0	83.8
30–39	30–39	80.0	88.4
40–49	40–49	84.0	91.2
50–59	50–59	83.0	88.8
	60–64		84.9
60–69	65–69	78.0	73.8
≥70	≥70	54.0	46.0
Car ownership:			
Gender			
Men		61.0[1]	74.1
Women		49.0[1]	48.2
Household income: Netherlands[2]	Household income: UK		
<7500	Lowest level	48.0	45.6
7500–15 000	Second level	64.0	51.5
15 000–22 500	Third level	82.0	72.1
22 500–30 000	Fourth level	89.0	84.5
≥30 000	Highest level	90.0	96.3

Notes:

[1] For main driver.

[2] In euros yearly.

Source: UK-NTS (2009); CBS Statline (2011).

licence, this is not reflected in car ownership. These differences might be a result of the characteristics of the urbanization patterns in both countries. In the UK the urban system is dominated by large metropolitan areas and a sharp contrast between high and low population densities. As a consequence, inhabitants of the countryside have to travel large distances to reach their preferred destinations. However, in the Netherlands the urbanization pattern has a far more polycentric character, with many medium-sized cities and a relatively balanced distribution of the population (Limtanakool et al., 2006), which offers higher access to concentrated destinations. The domination of medium-sized cities at relatively short distances from each other and the existence of a relatively good rail and bicycle infrastructure in the Netherlands support the use of these sustainable modes as 'real' alternatives to car ownership and use (Dieleman et al., 2002; Schwanen et al., 2004). With the exception of young people and the elderly, the possession of a driving licence is higher in the Netherlands than in the UK. Household income in both countries has a positive impact on the ownership of a car. Good observation of the figures makes it clear that except for the highest category for all income categories, car ownership is lower in the Netherlands than in the UK. As discussed above, this might be explained by the combination of urbanization pattern and infrastructure which favours the Netherlands in sustainable transport modes.

The average number of trips and kilometres per person per day are shown in Table 3.2. In general, the average number of trips over time is a rather stable travel attribute for the UK, as observed by Metz (2010). However, the average distance travelled usually increases with the level of GDP, which allows inhabitants of a country to buy additional travel speed. In the last decade the distance travelled on overland transport (thus excluding air transport) stabilized owing to diminishing marginal utility, which saturates mobility choices (Metz, 2010). Table 3.2 shows large similarities between both countries for these travel attributes. Men and women are characterized by taking more or less the same number of trips, but resulting in men travelling larger total distances. This is largely due to the fact that more men than women are working full time and have a larger commuting distance (Schwanen et al., 2003). In the United Kingdom, men have on average a commuting distance of 17 kilometres and women 10 kilometres. For the Netherlands these figures are 15 kilometres and 6 kilometres respectively. These gender differences in working hours and commuting distance can be explained by the greater responsibility of women for maintenance of the household and care of children (McDowell et al., 2005; van der Klis and Mulder, 2008).

The most active age categories are the people who are 'middle-aged'. In this stage in their life course they usually participate in the labour market and develop their career, while most of them also raise children. These activities are accompanied by trips to reach the demanded activity places, which are usually at relatively large distances from home. At both ends of the age scale most young people and the elderly do not have paid jobs and their social networks are rather limited in geographical scale (Giuliano and Dargay, 2006). In general, the number of trips and travel distance go up with increasing income levels. However, in both countries, the second lowest income level

Table 3.2 Travel attributes in the United Kingdom and the Netherlands by socio-demographics

Socio-demographics		No. of trips per day		Kilometres per day	
		United Kingdom	The Netherlands	United Kingdom	The Netherlands
Gender					
Men		2.6	3.0	20.2	41.8
Women		2.7	3.1	17.0	28.8
Age UK in years	Age Netherlands in years				
15–24		2.5	2.6	29.9	33.0
25–44		3.0	3.3	38.0	42.7
45–64		2.8	3.3	34.4	38.7
≥65		2.0	2.4	18.3	19.9
Household income: UK	Household income: Netherlands[1]				
Lowest level	<7500	2.3	3.0	17.2	30.0
Second level	7500–15 000	2.5	2.9	22.7	25.5
Third level	15 000–22 500	2.7	3.3	28.3	37.4
Fourth level	22 500–30 000	2.9	3.4	35.2	46.9
Highest level	≥30 000	3.0	3.4	46.0	59.7
Car ownership					
No		2.0	2.4	13.5	21.2
Yes		2.8	3.3	33.6	44.1

Note: [1] In euros yearly.

Source: UK-NTS (2009); CBS Statline (2011).

deviates from this general ‘rule’. Higher-income households often include more highly educated people, who usually have a more specialized job which can be found at a larger commuting distance from home (Shen, 2000; Susilo and Maat, 2007). At the same time larger financial budgets can be spent on trips for participation in out-of-home activities. Their income allows them to buy one or more cars for the household (Table 3.1), which supports a substantially larger number of trips and travel distances in comparison with people without cars (Table 3.2).

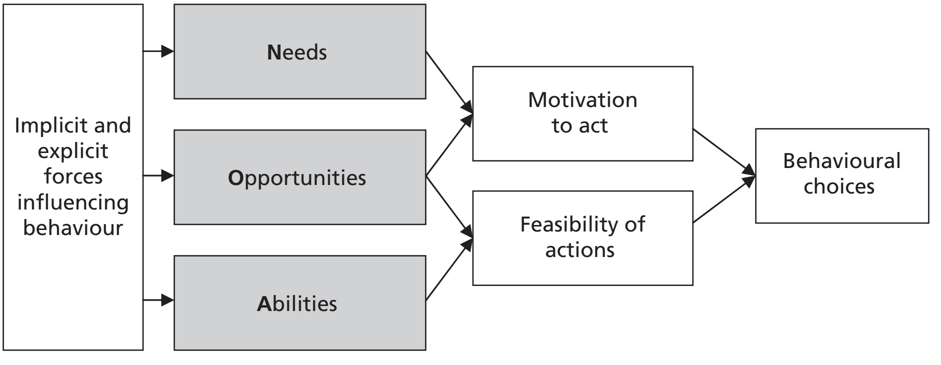
Although variation in travel behaviour is related to socio-demographics and characteristics of land uses and other urban form attributes, other contextual differences might be relevant as well. This is shown, for example, by Buehler

(2011), who made a comparison between the US and Germany in car use. In the US, 89 per cent of all trips are travelled by car, while this figure is 60 per cent for Germany. Even after controlling the impact of socio-demographics and land uses, differences between countries remained the same. This result can be explained by cultural preferences and various transport policies which favour the use of the car in the US and sustainable transport modes in Germany.

3.3 Conceptual model travel behaviour

Often we say that travel is not an aim in itself but that the utility of travel is derived from activities employed at the destinations. Pas (1980: 3–4) said: 'if all the activities in which an individual wished to participate were located at the same place, that individual would be expected to undertake little or no travel at all'. Or, in other words, if an individual can choose between an attractive destination at a distance of 1 kilometre and an equally attractive destination at 10 kilometres she will definitively opt for the first alternative. Banister (2008) states that this traditional and predominant view on transport should be rethought in order to develop effective policies to reach sustainable mobility. The fundamental assumptions of traditional transport research and planning deny that travel not only is a derived demand but also has an intrinsic value (see also Chapters 2 and 6) expressed in its symbolic and affective factors which refer to feelings or emotions evoked by travelling (Anable and Gatersleben, 2005; Ory and Mokhtarian, 2005; Steg, 2005).

Understanding travel behaviour is dependent on the activities in which individuals want to participate at their destination(s) or while travelling and the options they have to fulfil these needs. Figure 3.1 describes the NOA model from psychology. This model distinguishes three general factors that influ-



Source: Vlek et al. (1997).

Figure 3.1 Individual factors influencing travel behaviour: the NOA model ence behaviour: needs, opportunities and abilities for explaining behavioural choices. The motivation for behaviour arises from needs (N) and the presence of opportunities (O) in an individual's context to fulfil these needs, like the supply of transport alternatives and distance to destinations. The individual abilities (A) refer to the available time, money, skills and capacity for certain travel choices. These abilities, in combination with the contextual opportunities, determine the choice set of an individual or the feasibility of his behaviour. This figure also makes it clear that options for behavioural changes can be found in the needs, opportunities and individual abilities.

'Motivation' is an umbrella concept referring to different factors determining the attractiveness of travel behaviour. In the first place rational considerations are important. These are usually based on the instrumental meaning of transport modes. An individual might, for example, choose a transport mode by considering travel costs, travel time, comfort, action radius, flexibility and capacity to carry goods. Secondly, social factors are important as well. People may choose certain transport modes because they expect that other people will see that as 'normal' or 'desirable'. For example, a private car gives more status than the use of the bus, and some types of car are more representative for some groups than others (Steg and Tertoolen, 1999; Steg et al., 2001). Finally, emotions play a role. Car driving gives some people a kick but is for others a rather stressful activity. As stated before, in research and policy, rational motives for travel dominate.

The feasibility of behaviour (also called 'perceived control' or 'self-efficacy') is often not in line with the observed or 'objective' options. People often systematically overestimate the advantages of their own behaviour (like car use), while systematically underestimating the disadvantages of this behaviour. And the reverse: people tend to overestimate the negative aspects of alternative behavioural options (like the use of public transport) and downplay the positive aspects of it (Golob et al., 1979). Besides systematic biases people also lack information and knowledge on behavioural alternatives.

Feasibility and motivation for behaviour are not independent of each other. A lack of motivation for behaviour might lead to a denial of opportunities to use it. And if specific behaviour options are difficult or even not feasible, we may trivialize or deny the negative consequences caused by our behaviour. This mechanism is called 'cognitive dissonance reduction': a person may experience cognitive dissonance when her behaviour (e.g. 'I travel by car') does not match with her knowledge (e.g. 'Car use causes environmental problems'). This causes negative feelings, which can be solved by adjusting the behaviour (reduce car use) or the cognition ('Car use does not have a big

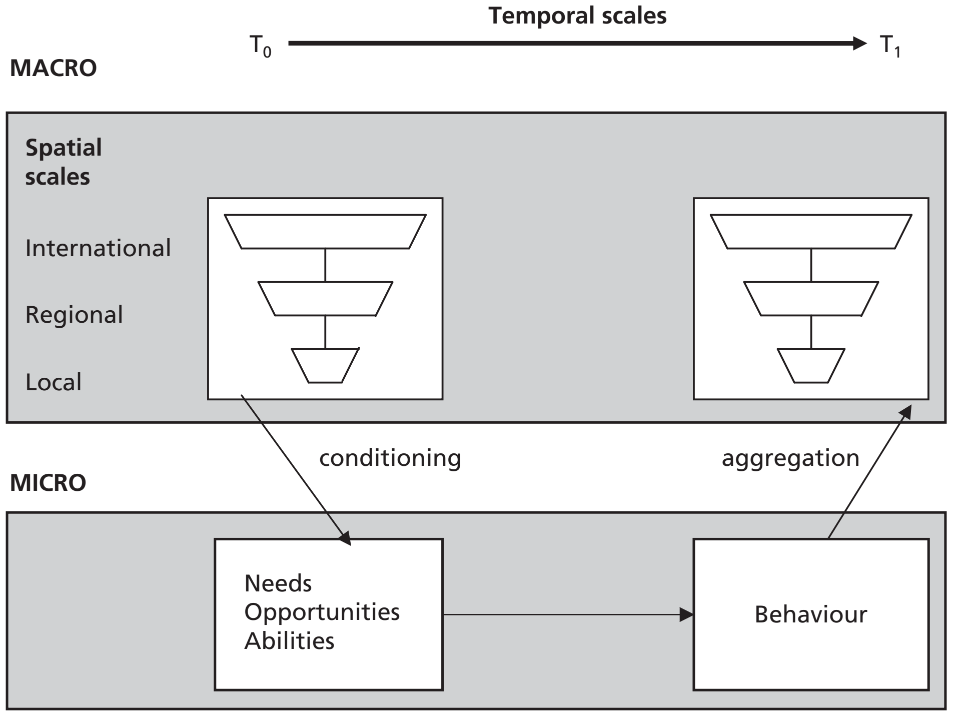


Figure 3.2 Relation between macro- and micro-developments for behavioural choices in time

negative impact on the environment'). In general, it will be easier to change one's attitudes than one's behaviour (Steg and Tertoolen, 1999).

The needs, opportunities and abilities of an individual are (in)directly related to developments in society. Changes in economic growth, changed demographic compositions of populations and households, changes in the values and norms of different (sub)groups in society, and increased diversity in ethnicity, labour participation and incomes are examples of these developments. These are taking place at the macro-level (Figure 3.2). These developments can take place on various spatial scales: a neighbourhood, a city, a region, a country, a continent or even the whole world. They determine ('condition') at T_0 the conditions for behavioural choices at the micro-level. These refer to individuals' needs and choice options defined by the NOA model. Based on this choice set, individual behavioural choices will be made which after some time at T_1 might lead to changes ('aggregation') at the various macro-levels. For example, choices for the private car as commuting mode can have an impact on congestion levels in city regions, and traffic safety in the neighbourhood, and lead to an increase of CO_2 in a country or

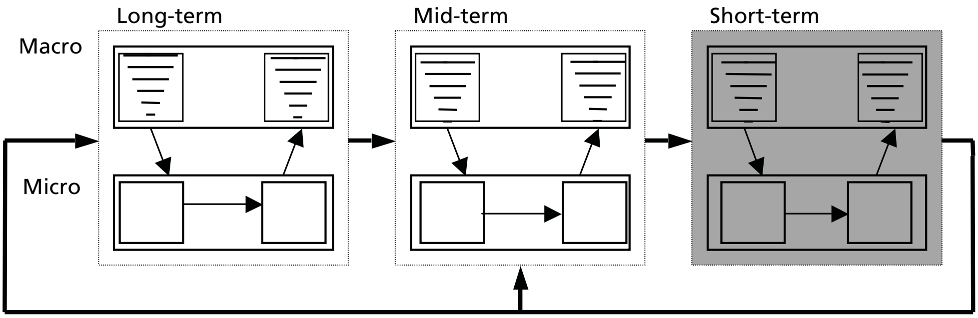


Figure 3.3 Continuum of related choices at various temporal scales

worldwide. The micro-level applies to individuals and their households as well as firms and other institutional actors. Their individual actor decisions will also have implications for the macro-level.

So far, we have been discussing the travel behaviour of people. However, travel behaviour is dependent on other mobility choices of people, like the purchase of a private car, choice of a dwelling and lifestyle choices for work, household and leisure. Cullen (1978) studied in particular the integrative character of time (see also Salomon and Ben-Akiva, 1983; Prillwitz et al., 2007). The connections between short-term choices ('travel behaviour'), mid-term choices and long-term choices are especially important. Long-term behaviour refers to household formation (e.g. getting married, having children or becoming widowed), labour participation (e.g. full-time or part-time employment and occupation) and orientation in leisure (e.g. preferences for sports or cultural activities). Choices for residential or (fixed) work location and the ownership of transport modes or ICT devices are examples of mid-term choices. Long-term and mid-term choices do not occur very frequently but strongly determine, for a relatively long time, the social and spatio-temporal contexts of the daily life of people. These decisions are usually rational decisions made by carefully weighing the pros and cons. As stated before, this is often in contrast to the habitual character of short-term behaviour.

In reality, a sharp distinction between these choices doesn't exist, but they are formed by a continuum of coherent decisions that people are making on different temporal scales. As an analogue to the choices for travel behaviour (Figure 3.2), the mid-term and long-term choices can also be explained by the NOA model. Figure 3.3 makes it clear that short-term choices are determined by mid-term and long-term choices and the reverse. A person's experiences in daily life in a geographical context can stimulate that person to move to another residential environment (mid-term decision) which better matches her needs in daily life or to change aspirations of household formation, employment and/or leisure time (long-term decision).

In various scientific disciplines, theories and models have been developed which explain the travel choices of people and the factors which influence their choice options. In the next three sections psychological, economic and geographical perspectives will be discussed. Although these disciplinary perspectives show fundamental differences in describing the behavioural mechanisms, they cannot be discussed in isolation from each other. Implicitly, they often take into account the ideas from other disciplines.

3.4 Behavioural choice from a psychological perspective

The NOA model (see Figure 3.1) distinguishes three general factors that influence behaviour: needs (or more generally motivational factors), abilities and opportunities. Psychologists are particularly interested in the role of motivational factors. Different theoretical perspectives have been put forward to study behaviour and, more particularly, travel behaviour. Below, we describe three lines of research that focus on different types of individual motivation that affect travel behaviour: perceived cost and benefits, moral and normative concerns, and affect, respectively. We also indicate how these different perspectives may be integrated into an all-encompassing framework. Next, we identify two shortcomings of these theoretical perspectives. Firstly, they do not pay due attention to the effects of contextual factors (as reflected in opportunities; see Figure 3.1) on travel behaviour. We propose ways to study individual and contextual factors simultaneously. Secondly, they imply the assumption that people make reasoned choices. However, in many cases people act habitually, which we discuss at the end of this section.

3.4.1 Motivational factors: three lines of research

Weighing costs and benefits

Various studies on travel behaviour started from the assumption that individuals make reasoned choices and choose alternatives with the highest benefits against the lowest costs (e.g. in terms of money, effort and/or social approval). One influential framework is the theory of planned behaviour (TPB) (e.g. Ajzen, 1991), which assumes that behaviour results from an intention to engage in the relevant behaviour: people plan to engage in a particular behaviour. Intentions depend on three factors: attitudes, subjective norms and perceived behavioural control. Attitudes reflect how positively

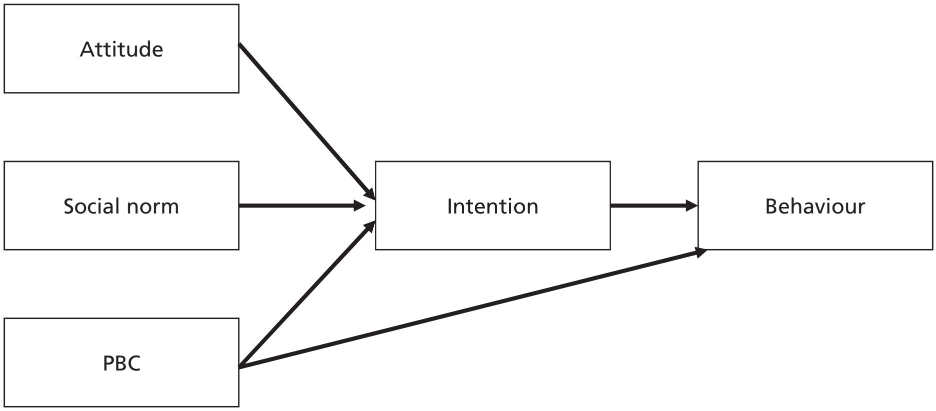


Figure 3.4 The theory of planned behaviour

or negatively people evaluate a particular action. Attitudes depend on beliefs that a behaviour will result in particular outcomes (e.g. driving a car is expensive, provides freedom or enhances one's status) and on how important these outcomes are for an individual. Social norms reflect the extent to which one believes that important others (e.g. friends, family members, colleagues) approve or disapprove the behaviour, and the motivation to comply with these expectations. Perceived behaviour control (PBC) reflects the extent to which one thinks one is capable of engaging in the relevant behaviour (see Figure 3.4). Perceived behaviour control can influence behaviour indirectly, via intentions, but also directly. For example, one can have the intention to travel by bus and feel capable of doing so (for example, because one knows the timetable and can afford to buy a ticket), but, if one then learns that the bus drivers are on strike, perceived behavioural control will affect behaviour directly.

The TPB assumes that other factors, such as demographics and general values, affect behaviour indirectly, via attitudes, subjective norms and perceived behaviour control. For example, men may travel more by car because they like driving (a positive attitude), low-income groups may drive less because they have a lower perceived behaviour control (e.g. they cannot afford to drive more) and people with strong environmental values may drive less because they are concerned about the negative environmental consequences of driving, resulting in less positive attitudes towards driving. However, as of yet, this assumption has hardly been tested explicitly in the travel domain. The extent to which attitudes, subjective norms and perceived behaviour influence intentions and behaviour differs across different types of behaviour. For example, subjective norms will be less influential when the particular behaviour is private and hardly visible to others (e.g. your friends are unlikely to observe which route you take to your holiday in France). In such cases, attitudes and perceived behavioural control are likely to exert a stronger influence on behaviour than do subjective norms.

The TPB has proven to be successful in explaining travel mode choice (Verplanken et al., 1998; Harland et al., 1999; Heath and Gifford, 2002; Bamberg and Rölle, 2003). Generally, people have a favourable attitude towards car use, while attitudes towards the use of public transport (and especially bus use) are far less positive (Steg, 2003). The subjective norm towards car use is also generally positive: people tend to think that others expect them to travel by car. The more positive subjective norms towards driving a car, the more people actually drive their car.

Various scholars have added further factors to the TPB, such as habits (Verplanken et al., 1997; see below). It has been shown that, when habits are strong, intentions influence behaviour less strongly compared to when habits are weak. Others have added positive and negative affect as predictors, reflecting the extent to which individuals anticipate that behaviour will result in positive or negative affect (we elaborate on the role of affect below). For example, some people may anticipate positive feelings when cycling in sunny weather or when driving during rain.

Moral and normative concerns

As described above, many people evaluate car use much more favourably than using public transport. This implies that reductions in car use are not very likely when people base their decisions mainly on weighing the costs and benefits of different travel modes. They will probably only reduce their car use when they value the environment and when they are concerned about the problems caused by car use. This implies that morality may play a key role in motivation to reduce car use: people need to forgo individual benefits to safeguard collective qualities.

Various studies examined the role of moral and normative consideration underlying travel behaviour, from different theoretical perspectives. First, scholars have examined the value-basis of beliefs and behaviour in the travel domain (e.g. Nordlund and Garvill, 2003; de Groot and Steg, 2008). In general, a distinction is made between self-enhancement values, in which individuals are particularly concerned about their own interests, and self-transcendence values, in which individuals are particularly concerned with the interests of others, society and the biosphere. These studies revealed that the more strongly individuals subscribe to values beyond their immediate

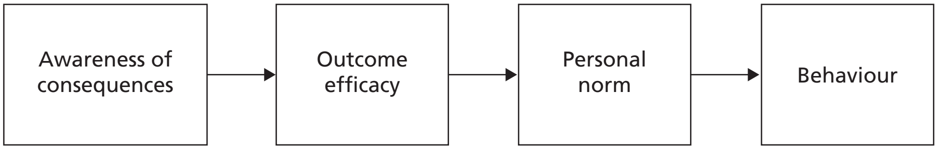


Figure 3.5 The norm activation theory

own interests, that is, self-transcendent, altruistic or biospheric values, the more favourably they evaluate reductions in car use and the more they are willing to do so.

A second line of research focuses on moral obligations to reduce the environmental impact of travel behaviour, in particular to reduce car use. These studies are based on the norm activation model (NAM) (see Figure 3.5) (Schwartz, 1977; Schwartz and Howard, 1981). The basic idea is that people will choose environmentally friendly travel options when they feel morally obliged to do so, which is reflected in personal norms. Personal norms are activated when people are aware of the problems caused by their behaviour (problem awareness, such as local air pollution caused by their car use) and when they feel responsible for these problems and think their actions can be effective in reducing these problems (outcome efficacy, e.g. 'When I drive less, local air pollution will reduce') (Bamberg and Rölle, 2003; de Groot and Steg, 2009; Steg and de Groot, 2010). The NAM appeared to be successful in explaining behaviour in the travel domain, and particularly willingness to reduce car use (e.g. Nordlund and Garvill, 2003; Eriksson et al., 2006; de Groot and Steg, 2009; Steg and de Groot, 2010). However, a study by Bamberg and Rölle (2003) revealed that the TPB predicted car use better than the NAM (see also Matthies and Blöbaum, 2007). Other studies also suggest that the NAM is particularly successful in explaining low-cost behaviour changes and good intentions, while the predictive power is less in situations characterized by high behavioural costs or strong constraints on behaviour, such as reducing car use (e.g. Hunecke et al., 2001; Bamberg and Rölle, 2003).

Affect

Various studies have explicitly examined the role of affect in explaining travel behaviour, mostly in relation to car use (see Gatersleben, 2007, for a review). These studies assume that travel behaviours are motivated not only by the (anticipated) instrumental outcomes of this behaviour (e.g. 'If I drive to work I shall get there faster than if I take the train'), but also the symbolic outcomes (e.g. 'If I take the bus to work my colleagues will think I am a loser') and the affective outcomes (e.g. 'Driving to work is much more fun than taking the bus'). So it is assumed that three types of motives may underlie travel behaviour: instrumental, symbolic and affective.

A study by Steg (2005) revealed that commuter car use was most strongly related to symbolic and affective motives, while instrumental motives (such as costs) appeared less important. This suggests that, even for highly functional trips such as commuting, affective and symbolic motives play an important role; this may be even more so for leisure trips (Anable and Gatersleben, 2005). Also, most group differences were found in the evaluation of the symbolic and affective functions of car use, while people tended to agree more on the relative importance of instrumental functions of car use. More specifically, young people and low-income groups generally valued the affective function of the car more than older respondents and higher-income groups, while male drivers valued the symbolic (and some affective) functions more strongly than female drivers did (Steg et al., 2001; Steg, 2005). Also, the car is evaluated much more favourably on these aspects than public transport (e.g. Steg, 2003).

An integrative perspective on environmental motivation

The three general lines of research just described involve rather different antecedents of travel behaviour. All three perspectives proved to be predictive of at least some types of travel behaviour. The three theoretical perspectives are not mutually exclusive, as behaviour is likely to result from multiple motivations. Goal-framing theory (Lindenberg and Steg, 2007) postulates that goals govern or 'frame' the way people process information and act upon it. Three general goal-frames are distinguished: a hedonic goal-frame 'to feel better right now', a gain goal-frame 'to guard and improve one's resources' and a normative goal-frame 'to act appropriately'. When a goal is focal (that is, when it is the 'goal-frame'), it influences how people perceive and evaluate different aspects of a situation and act upon it. The hedonic goal-frame is a priori strongest, while the normative goal-frame is especially in need of external social and institutional support in order to become focal. Goal-framing theory also proposes that, typically, multiple goals are active at a given time: one goal is focal and influences information processing the most (that is, it is a goal-frame), while other goals are in the background and increase or decrease the strength of the focal goal. The three goal-frames remarkably coincide with the three theoretical frameworks described above. That is, theories and models on affect focus on hedonic goals, the TPB focuses on gain goals, while the NAM and theories on values focus on normative goals. Thus, goal-framing theory provides an integrative framework for understanding motivations underlying travel behaviour.

3.4.2 Contextual factors

The motivational theories discussed above focus on individual motivations influencing travel behaviour, although the TPB considers individuals' perceptions of contextual factors, as expressed in perceived behavioural control. Obviously, travel behaviour does not depend on motivation alone. Many contextual factors may facilitate or constrain travel behaviour by influencing the opportunities people face. For example, the quality of public transport, or petrol price regimes can strongly affect travel behaviour (e.g. Santos, 2008; see also sections 3.5 and 3.6). In some cases, constraints may even be so severe that motivation is hardly related to travel behaviour. So it is important to consider individual motivation vis-à-vis contextual factors (as reflected in the NOA model; see Figure 3.1). This can be conceptualized in four different ways. First, contextual factors may directly affect behaviour. For example, one cannot travel by bus when no bus service is available, while a free bus ticket may result in an increase in bus ridership (e.g. Bamberg and Rölle, 2003; Fujii and Kitamura, 2004). Second, contextual factors may affect behaviour indirectly, via motivational factors such as attitudes, affect or personal norms. For example, the introduction of a cycle path may result in more positive attitudes towards cycling (e.g. because it is safer) and positive attitudes may in turn promote cycling. Third, contextual factors may moderate the relationship between motivational factors and behaviour. For example, environmental values may only result in reductions in car use when feasible alternatives are available, and cycling facilities may promote cycling only among those with strong environmental values. Fourth, related to the third point, following goal-framing theory, contextual factors may determine which type of motivation (and thus which goals) most strongly affects behaviour. For example, normative goals may be strongly related to frequency of cycling when good cycling facilities are available, while gain or hedonic goals may be prominent if cycling facilities are poor.

3.4.3 Habitual behaviour

The theoretical frameworks discussed in section 3.4.1 largely imply that individuals make reasoned choices, that is, they assume that choices are based on a careful deliberation of the pros and cons of different behavioural alternatives. However, in many cases, behaviour is habitual and guided by automated cognitive processes, rather than being preceded by elaborate reasoning. After all, we cannot possibly consider all the pros and cons of all choices that we face during a day. We simply do not have the cognitive capacity and time to do so. We just repeat the same action over and over again when we face similar choice situations. Habits are formed when behaviour results in the anticipated positive consequences over and over again. In that case, behaviour is automatically elicited by contextual cues.

Aarts et al. (1998) defined three important characteristics of habits. First, habits require a goal to be achieved. Second, the same course of action is likely to be repeated when outcomes are generally satisfactory. Third, habitual responses are mediated by mental processes. When people frequently act in the same way in a particular situation, that situation will be mentally associated with the relevant goal-directed behaviour. The more frequently this occurs, the stronger and more accessible the association becomes, and the more likely it is that an individual acts accordingly. Thus habitual behaviour is triggered by a cognitive structure that is learned and stored in and retrieved from memory when individuals perceive a particular situation.

Habits refer to the way behavioural choices are made, and not to the frequency of behaviour. Aarts and colleagues (e.g. Aarts and Dijksterhuis, 2000) developed a so-called response-frequency measure of general habit strength, relying on the assumption that goals automatically activate mental representations of habitual choices. This measure is far more accurate than simply asking people how frequently they engage in a particular behaviour, as it focuses on how choices are made. The measure has been successfully employed in various studies on travel behaviour (e.g. Aarts et al., 1998; Aarts and Dijksterhuis, 2000; Klockner et al., 2003).

Habitual behaviour may involve misperceptions and selective attention: people tend to focus on information that confirms their choices, and neglect information that is not in line with their habitual behaviour. It is also possible that people change their beliefs in line with their habitual behaviour; for example, habitual car users may evaluate driving a car even more positively and travelling by public transport more negatively to rationalize their behavioural choices.

In many cases, habits are highly functional because they enable us to cope efficiently with limited cognitive resources and time. However, when choice circumstances have changed, people may no longer make optimal decisions when they have strong habits. In general, habits are reconsidered only when the context changes significantly. For example, Fujii and colleagues found that temporarily forcing car drivers to use alternative travel modes induced long-term reductions in car use (Fujii et al., 2001; Fujii and Gärling, 2003).

The impacts of such temporary changes were particularly strong for habitual car drivers. This suggests that habitual drivers have inaccurate and modifiable perceptions of the pros and cons of alternative transport modes. Lifestyle changes may also result in reconsidering habitual behaviour, for example moving house, changing jobs or having children.

3.5 Behavioural choice from an economic perspective

In the economic discipline, actors are usually labelled as consumers instead of people or individuals. Standard economic analysis departs from the assumption that consumers base their choice on rational considerations. Preferences of consumers are the starting point of many analyses. These preferences mean that consumers – when they have the choice between options A, B and C – will arrive at a certain ranking; for example, B is preferred above A, which is preferred above C. The standard assumption is that consumers are able to arrive at a complete and consistent ranking (see, for example, Varian, 1992).

When we apply this preference-based approach to the domain of transport, the choice alternatives may be various transport modes with scores of attributes such as price, speed and comfort. A tendency can be observed that economic research focuses on the functional properties of transport modes (see also section 3.3). In economics the standard way to represent the preferences of consumers is to use utility functions (Figure 3.6). The utility functions give the summary score for the alternatives, where the various attributes are weighted according to their importance. This approach is similar to computing the attitude component in the model of planned behaviour (see section 3.4). Since people are different, it may well be that each consumer bases her choice on an individual-specific utility function.

Figure 3.6 gives an example of a utility function of a certain consumer. The curve contains the set of all combinations of attributes (for example, speed and comfort) being valued equally by the consumer, implying that he is indifferent to these alternatives. The curve contains rather different combinations of attributes, some with high speed and low comfort, and other alternatives with the opposite attribute combination, but all of them are valued equally by this consumer. Of course there will be additional relevant attributes, but for the ease of presentation we focus on these two attributes here.

Preferences are defined here in a way that is different from the NOA model (Figure 3.1). The NOA model focuses on fundamental needs such as safety,

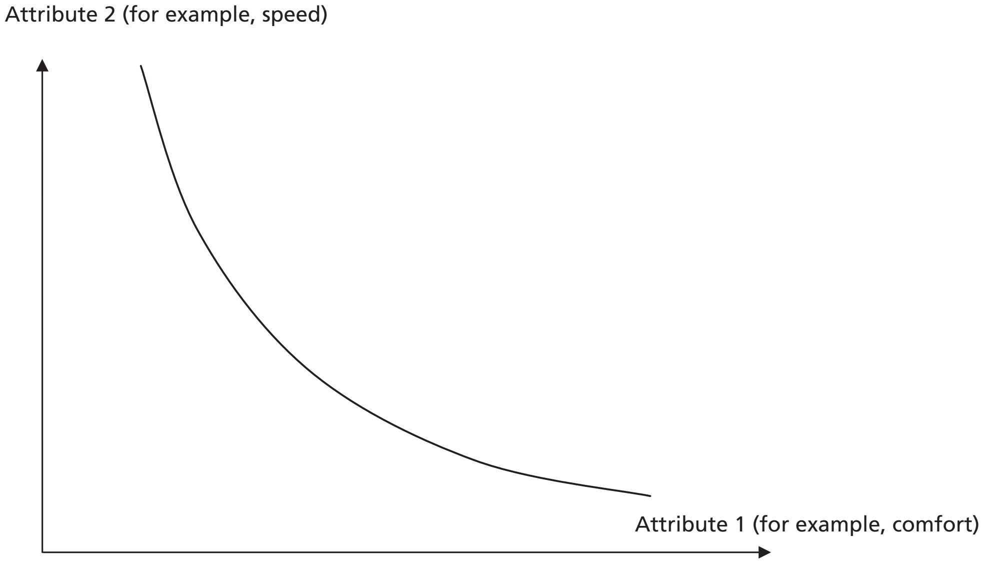


Figure 3.6 Illustration of a utility function with two attributes

health and variety, while economic models deal with 'wants'. Wants relate to preferences for ways to satisfy needs. Most economic models are based on the assumption of infinite wants: the more consumption, the better. Sometimes economic research addresses satiation phenomena. For example, people value an extra hour of sleep positively, but this value decreases as people sleep longer, and the valuation of an extra hour of sleep may even become negative.

An important observation is that wants, in an economic sense, cannot always be expressed, since the consumer lacks the financial means. Many people in the world would like to travel by car, but a considerable number do not have the money to purchase one. In economics, only the wants that are accompanied by sufficient purchasing power determine which transport services will be supplied, not the needs of the consumer. Governments can intervene to prevent fundamental needs not being satisfied, for example by subsidies for public transport in order to guarantee that long-distance trips can also be made by people with a low income.

A central theme of consumer behaviour in economic theory is the so-called optimal allocation of consumptive expenditures. This means that a consumer creates a bundle of consumption items that maximize her utility, given her income level and the prices of the various consumption items, and given her preferences as represented by the utility function. In mobility research, analysts pay attention not only to income constraints but also to time constraints. This means that consumers consider not only the price of goods, but also the time needed for consumption. For travel behaviour this means that consumers consider not only the price per kilometre, but also speed (see, for example, Becker, 1965; Golob et al., 1981; Small and Verhoef, 2007).

Optimizing utility leads to a so-called demand function for kilometres travelled where the size of demand depends on factors such as the price per kilometre, prices of other goods, income, speed of the trip and time use involved in other activities. Knowing the demand function is of great use when one wants to predict what will happen when the price per kilometre or income changes. There are various ways to express the results of economic analysis for practical purposes: valuation of travel time, price elasticity, time elasticity and income elasticity.

3.5.1 Value of travel time

An often used indicator of consumer preferences in travel is the so-called value of time (VOT) or value of travel time savings (VTTS). Chapter 12 gives a more detailed discussion of this subject. In the present chapter, we confine ourselves to a short introduction. The value of time is the core of the trade-off consumers make between price and speed when they compare various travel alternatives (Small and Verhoef, 2007). Estimations of the value of travel time usually fall within a range of 5 to 25 euros per hour. A valuation of travel time of 25 euros per hour means that, when a consumer compares a railway trip that takes 6 hours with a flight that takes only 3 hours, he will choose the train as long as it is 75 euros cheaper than the flight. Note that this is a simplified example, since other attributes such as comfort will also play a role in this trade-off. Knowledge of the value of travel time is an important tool when one wants to predict travel choices.

Different travellers will have different values of time. The value of time will obviously depend on what people would do with their time when they save travel time. If they would use this time to work longer, that would of course lead to higher income. People with a high wage rate per hour will therefore have a high value of time. The VOT also depends on the trip purpose. Business trips have the highest values of time, followed by commuting and finally other trip purposes. Situational conditions also play a role here. Someone who has an important appointment and who faces the risk of being late because of a delay will have a high value of time under that circumstance. Note also that the VOT will depend on the time that is already used for travel. Someone who already travels a lot for his work will probably have less time for other activities so that he will be prepared to pay high amounts for trip alternatives that will save time.

Another important concept in transport economics is that of generalized costs of a trip. The generalized costs are the sum of the monetary costs and the time-related costs (see, for example, Bruinsma et al., 2000). As people become richer they will be prepared to pay more to save travel time. Hence, one may expect a tendency for the share of time costs in generalized transport costs to show an increasing trend. This also means that the relevance of financial incentives will gradually decrease as people get richer (see also Chapter 6) and consumers will increasingly pay attention to attributes like quality, speed, reliability and comfort.

3.5.2 Price elasticity

The price elasticity of demand for kilometres is the percentage change in demand (q) when the price per kilometre (p) increases by 1 per cent:

$$\text{Price elasticity of demand} = [\Delta q/q]/[\Delta p/p]$$

For example, when the fuel price elasticity of demand for car kilometres is −0.2, this means that when the fuel price increases by 10 per cent the number of kilometres driven will decrease by 2 per cent. The price elasticity of demand for public transport is usually considerably closer to −1. This means that the demand for public transport is much more sensitive to price changes than is car transport to fuel price changes. One of the explanations for this is that the fuel costs have a rather limited share in total costs.

The elasticity as defined above is often called the 'own' price elasticity, since it gives the sensitivity to the price of the good or service itself. A related elasticity concept is the cross-price elasticity, implying the percentage change in the demand q for a change in the price of another good. Often the cross-price elasticity is positive. For example, a higher ticket price for rail would mean that the demand for car use will increase. This is an example of substitution between travel models. However, negative cross-price elasticities cannot be ruled out. For example, when the price of railway tickets decreases this may mean that more people will travel by train and hence also more people will travel by bus to go to the railway station. This would be an example of complementarity, a result one may expect in the case of multimodal transport chains.

3.5.3 Travel time elasticity

The demand function for transport indicates that travel behaviour depends on the duration – and hence the speed – of trips. This is expressed by the travel time elasticity. It appears that the speed does indeed have a strong influence on the demand for transport. The long-term travel time elasticity of the demand for transport is close to −1. A decrease in travel time by a certain percentage will lead to a similar increase in the total distance travelled. This elasticity of −1 would imply that the total time that is used for travelling is about constant in the course of time (Zahavi, 1979) (see also Chapter 6). Van Wee et al. (2006) even find indications that there is a gradual increase in the total time spent travelling. In the context of time elasticities, cross-travel time elasticities may also be relevant. Consider for example a multimodal trip chain with train as the main transport mode. Access modes bringing passengers to the railway station are often slow; hence the demand for rail transport may well be rather sensitive to the speed of access modes.

3.5.4 Income elasticity

The last factor we discuss here in the context of the demand function is the impact of income on the demand for transport. This can be expressed by the income elasticity. Consumer goods are defined as luxury goods when the income elasticity is higher than 1. In that case consumers spend an increasing part of their income on these particular goods. Another category of goods has an income elasticity of between 0 and 1: consumption of these goods increases with income, but at a decreasing rate. A last category of goods has a negative income elasticity: as people get richer they will consume it less and less (so-called inferior goods).

Aviation has a high income elasticity, clearly higher than 1. This is one of the reasons why aviation has grown so rapidly during the last few decades. The immediate consequence is that aviation is also a sector that will be hurt most in the case of an economic recession. Cycling and walking are transport modes with a very low income elasticity (close to zero). There is also a relationship between people's income and the trip purposes for which they would use certain transport modes. For example, people with high incomes may use the bicycle mainly for recreational activities when the weather is good, instead of using it as a transport mode for daily purposes (Rietveld, 2001).

As people have higher incomes they tend to travel more. For example, in the Netherlands, people in the lower income brackets travel about 24 kilometres daily (for children this is about 17 kilometres per day). People in the

Table 3.3 Relationship between annual income and distance travelled per person per day, the Netherlands

Annual income (euros per year, 2007)	Distance travelled per person per day (km)
Younger than 12 years, no own income	17.2
Older than 12 years, no own income	23.9
Less than 7500 euros	30.0
From 7500 to 15 000 euros	25.5
From 15 000 to 22 500 euros	37.4
From 22 500 to 30 000 euros	46.9
Above 30 000 euros	59.8

Source: CBS (2007).

high income brackets travel about 60 kilometres per day (see Table 3.3). Further, people with higher incomes choose transport modes that are relatively expensive and within a given transport mode they choose the more expensive versions (for example, first versus second class on the train).

When we consider household expenditure, the quartile (25 per cent) of the population with the lowest income spend about 14 per cent of their income on travel. For the next quartile this is about 17 per cent. This jump can mainly be explained by the difference in car ownership between these two income groups. A remarkable relationship is found between income and the use of public transport. The use of bus, tram and metro decreases (or stabilizes) when income increases. According to the definition given above, these are thus inferior goods. For the train a different pattern is found. It is high for low-income groups (in particular students, who happen to enjoy free public transport), then for median incomes it is much lower and finally it clearly increases again for higher incomes. Thus, railway trips tend to be considered luxury goods in the higher part of the income distribution.

3.6 Behavioural choice from a geographical perspective

In general geographers study behavioural choices from three perspectives: behavioural geography, utility theory and time geography. Behavioural geography focuses on the cognitive processes underlying decision-making. As such it relies heavily on psychological concepts and mechanisms (see sections 3.3 and 3.4). Utility theories are based on the economic discipline (see section 3.5). In this section time geography is focused on, since it is the theory which is most unique to the geographical perspective.

Until the 1960s transport problems were studied by means of a trip-based approach only (see also Chapter 15). This means that the basis of the analysis is the trips, which are studied independently of each other. Connections with activities and with the behaviour of other people were rarely the subject of study in this approach (Jones, 1979). Time geographer Hägerstrand (1970: 9), in his legendary paper 'What about people in regional science?', wrote the following words on these selective approaches in social sciences:

> It is common to study all sorts of segments in the population mass, such as labor force, commuters, migrants, shoppers, tourists, viewers of television, members of organizations, etc., each segment being analyzed very much in isolation from the others . . . we regard the population as made up of 'dividuals' instead of individuals.

In the 1970s, as a reaction to the shortcomings of the trip-based approach, the activity-based approach was developed. 'Activity-based approach' is a collective term for a range of studies on the trips people (want to) make. Goodwin (1983) describes this approach as a way in which observed behaviour depends on the activity patterns of people and households within their constraints in time and place. In this integral approach emphasis is put on people's needs or preferences as well as on constraints for individual choices.

Time geography was originally developed by Hägerstrand (1970) and methodologically enriched with geographic information systems (GIS) methods by Kwan (2000) and Miller (2005). This theory is based on the idea that the life of an individual, but also of other organisms and material objects, describes an uninterrupted 'path' through time and across space. The time scale of these paths can vary from a day to the whole life span. Every organism or thing is constantly in movement, even when it is itself at rest. In this manner one can see a society as being built up from a large number of webs or networks composed of uninterrupted paths that have been drawn by people, other organisms and non-living elements through time and space (Dijst, 2009). In this theory, participation in activities is not a matter of choice, but subject to three types of constraint:

1. **Capability constraints:** biological, mental and instrumental limitations. For example, people sleep on average seven to eight hours a night at home, have a certain level of intelligence and skills to carry out activities, and have transport modes at their disposal which enable them to travel through time and across space at various speeds.
2. **Coupling constraints:** encountered by people when they come together at a certain time and location with equipment and other materials for joint activities like attending a lecture at a university or buying bread in a

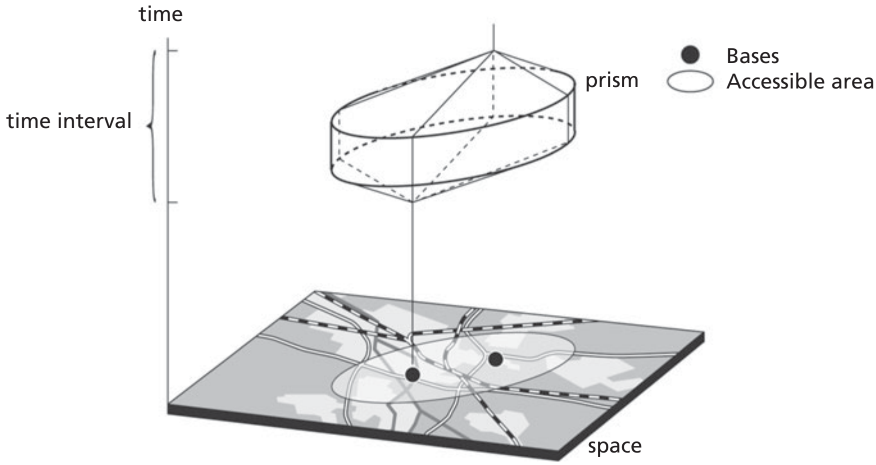


Figure 3.7 Prism and potential action space

bakery. Different time schedules and the different locations of origin can complicate this coupling of individual paths.

3. **Authority constraints:** these regulate the access of individuals to activity places through social rules, laws, financial barriers and power relationships. Business hours, the price of admission and curfews are some examples of these regulatory constraints.

In time geography, it is assumed that some activities such as work and home activities are fixed in time and in locations called 'bases'. These base locations, like fixed work locations and home, determine the opportunities to conduct more flexible activities such as buying the groceries and attending the theatre. These flexible activities can be pursued in 'time intervals', also called 'time windows', which represent blocks of time in which travel and relatively flexible activities can be carried out.

The capability, coupling and authority constraints define, for a certain time interval, a three-dimensional 'prism', which embraces the set of opportunities for travelling to activity places and to participate in activities (Figure 3.7). The projection of a prism in a two-dimensional space designates the 'potential action space'. This area contains all the activity places that can be visited within a certain time window: in other words, it represents the accessible area (Dijst, 2009).

Potential action spaces can be grouped into three basic categories according to their ideal shape: circular, linear or elliptic. Circular action space results when a journey starts and ends at the same base, like shopping trips from home or in the lunch break from work. A linear action space starts in one base and ends in another. The available time is just enough to cover the distance between the two base locations. This is often the case for daily commuting or trips to university for students. After subtracting the time necessary to travel between the bases, some extra time remains for visits to other activity places en route; the resulting action space is elliptic (Dijst et al., 2002). Each step taken in the potential action space transforms it into a path (Dijst, 2009).

Within this theoretical approach two behavioural rules can be distinguished which influence travel behaviour: fixation in time and space, and travel time ratio.

3.6.1 Fixation in time and space

As mentioned before, it is assumed that some activities are fixed in time and base locations. In section3.3, we called choices for these base locations mid-term choices. Why are these base locations so important? According to Cullen and Godson (1975: 9), 'Activities to which the individual is strongly committed and which are both space and time fixed tend to act as pegs around which the ordering of other activities are arranged and shuffled according to their flexibility ratings.' As explained above, these flexible activities, like daily shopping, can be pursued in 'time intervals', also called 'time windows', which represent blocks of time in which travel and relatively flexible activities can be carried out. In general, people have more short than longer intervals. In a working day, typical time intervals are: before the commuting trip at home, while commuting to work, during the lunch break at work, while commuting to home, after dinner. On non-working days people may use fewer but larger time intervals.

Fixations in time have meaning for the type of activity places people attend. Empirical research (e.g. Kitamura and Kermanshah, 1983; Golob, 1984; Dijst and Vidakovic, 2000) has shown that similar activities take place in the same time intervals. Mandatory activities, like daily shopping and taking children to school or day care centres, usually take place in relatively short intervals of a maximum of 1.5 hours. However, for leisure activities like social visits and attending performances in theatres or exhibitions in museums people usually need large time intervals. One of the arguments for this temporal sorting of activities is that people want to secure scarce intervals for compulsory activities as much as possible. Double-income households, especially, apply this strategy to avoid fragmentation of their leisure time.

Although time regimes show increasing levels of flexibility these can also lead to a temporal sorting of activities. For example, parents usually have to chauffeur their children two to four times a day to or from school. Often theatres cannot be visited during the daytime, and shops also have limited business hours. The duration of films, plays and sports matches are also often prescribed. Finally, the length of a time interval offers the option to take a longer time to travel to visit activity places that are at a relatively greater distance geographically.

3.6.2 Travel time ratio

The prism concept (Figure 3.7) makes it clear that individuals are constrained in their trade-offs between travel time and activity duration. Capability, coupling and authority constraints restrict the set of opportunities individuals have for travelling to activity places and to participate in activities. An individual has, in principle, three temporal choices without violating the constraints:

1. to spend the entire available time budget on travel without spending any time in an activity place;
2. to stay in a base location without travelling outside; or
3. to spend time on travel as well as on activities in one or more activity place.

To study the relationship between travel time and activity duration empirically Dijst and Vidakovic (2000) proposed the travel time ratio concept, which is defined as the ratio obtained by dividing the travel time to a particular activity place by the sum of travel time and activity duration for the same activity location. Schwanen and Dijst (2002) have shown that Dutch workers spend on average 10.5 per cent of their time available for work and travel on commuting. This corresponds to 28 minutes for an eight-hour work day. The travel time ratio for work varies systematically with socio-demographics. For example, a household type defined by the presence of a partner and children and employment status accounts for almost one-tenth of the variation. On average, daily shopping has a ratio of 0.40 and for social leisure activities 0.25 (Susilo and Dijst, 2009).

3.6.3 Application of geographical perspective

Based on the time geographical perspective for planning purposes, simulation models have been developed to assess the effects of planning measures on the choice opportunities that individuals of various types have. PESASP (programme evaluating the set of alternative sample path) and its improved version MASTIC (model of action spaces in time intervals and clusters) are good representatives of this type of model. These models facilitate the assessment of the potential impact of time policies (for example, the business hours of shops, flexible working hours and adjusted public transport schedules), transport policies (for example, new road construction and new bus stops) and spatial policies (for example, changes in the density or mixture of activity places) on the opportunities offered to people to participate in their desired activities (Dijst et al, 2002).

The MASTIC model includes four variables: the distance between the bases, the length of the available time window, the speed of the main transport mode and the travel time ratio. For these simulations accurate data on departure and arrival times of trips, the transport modes used and the street addresses of all types of activity place visited is needed. This data is usually derived from activity and travel diaries, sometimes in combination with GPS systems, and allows us to compose 'activity programmes' or the activities carried out during the various time windows in a day. In addition, a digital geographic database of the study area is necessary which includes information on activity places (like type and business hours) and a detailed street network for all transport modes.

Various scenarios can be defined to show the impact of, for example, transport and spatial policies on the feasibility of carrying out activity programmes in the available time windows. The scenarios in Table 3.4 differ in the density of dwellings, namely 15 and 35 dwellings per hectare. The results show that these spatial configurations have different impacts on people's opportunities to participate in activities outside the bases. As expected, a high suburban density offers relatively the best opportunities; on average, four-fifths of all activity programmes can be realized. In contrast, in a low suburban density with its longer travel times, only two-thirds of all activities can be carried out. In this scenario everybody in low-density areas, especially those who walk or cycle and members of two-earner households, encounters substantial difficulties in their daily lives.

3.7 Conclusions and synthesis

Below we summarize the most important conclusions of this chapter and give a synthesis of the links between the disciplines discussed:

1. In this chapter we have discussed the contribution of psychology, economics and geography to the understanding of travel behaviour. The

Table 3.4 Percentage of activity programmes realized in different spatial configurations, by household type and main transport mode

Density/household type	Car	Public transport	Walk or bicycle	Total
Suburban density scenario: high density	All differences significant at the $p < 0.01$ level			
Single-earner households with child(ren)	87	100	62	80
Single-earner households without child(ren)	83	100	100	84
Two-earner households with child(ren)	82	93	68	80
Two-earner households without child(ren)	83	86	61	78
Retired households without child(ren)	100	100	–	100
Total	83	89	65	80
Suburban density scenario: low density	All differences significant at the $p < 0.01$ level			
Single-earner households with child(ren)	72	100	41	64
Single-earner households without child(ren)	73	100	50	72
Two-earner households with child(ren)	69	87	55	68
Two-earner households without child(ren)	74	72	34	64
Retired households without child(ren)	88	100	–	83
Total	72	78	41	67

Note: – Number of cases too small.

Source: Ritsema van Eck et al. (2005).

three disciplines differ in the concepts they use, the identification of relevant determinants and behavioural mechanisms. However, each discipline explains only a part of the reality of behavioural choices. Only the combination of these disciplinary perspectives can lead to a more comprehensive understanding of travel behaviour. In this chapter we have tried to develop an integrative conceptual framework to show the coherence between these disciplines.

2. A comparison of economics and psychology shows that psychology covers a wider range of behavioural issues than economics. In economic studies the focus is on the rational thinking consumer, while in psychology habitual behaviour is also a subject of study. Besides, psychologists analyse the impact of other people on behavioural choices more than economists do. Finally, economists limit themselves to financial and temporal opportunities and constraints while psychologists also study other abilities, like skills.
3. Apart from these limitations, economics has an important advantage. This discipline is largely based on quantitative analyses which offer good opportunities to predict the impact of economic, technological and other societal trends on travel choices. The effectiveness of economic

policy measures can also be estimated. Examples are changes in income and prices, impacts of investments in new infrastructure or the choice of new residential areas. However, owing to the previously discussed limitations these predictions are not always accurate.

4. Some geographical approach uses are comparable with economics utility maximization theory. However, in this chapter we have focused on the spatio-temporal constraints people experience in daily life. Thus this geographical approach is complementary to a psychological and economic approach. Besides, geographers study short-term daily activity and travel behaviour in conjunction with the mid-term (e.g. choices of work and residential location and transport and communication modes) and long-term decisions (e.g. lifestyle choices) in the life course of individuals.

REFERENCES

Aarts, H. and A.P. Dijksterhuis (2000), 'The automatic activation of goal-directed behaviour: the case of travel habit', *Journal of Environmental Psychology*, 20, 75–82.

Aarts, H., B. Verplanken and A. van Knippenberg (1998), 'Predicting behaviour from actions in the past: repeated decision making or a matter of habit?', *Journal of Applied Social Psychology*, 28, 1355–1374.

Ajzen, I. (1991), 'The theory of planned behavior', *Organizational Behavior and Human Decision Processes*, 50, 179–211.

Anable, J. and B. Gatersleben (2005), 'All work and no play? The positive utility of travel for work compared to leisure journeys', *Transportation Research Part A: Policy and Practice*, 39, 163–181.

Bamberg, S. and D. Rölle (2003), 'Determinants of people's acceptability of pricing measures: replication and extension of a causal model', in J. Schade and B. Schlag (eds), *Acceptability of Transport Pricing Strategies*, Oxford: Elsevier Science, pp. 235–248.

Banister, D. (2008), 'The sustainable mobility paradigm', *Transport Policy*, 15, 73–80.

Becker, G.S. (1965), 'A theory of the allocation of time', *Economic Journal*, 75, 493–517.

Bruinsma, F.R, P. Rietveld and M. Brons (2000), 'Comparative study of hub airports in Europe: ticket prices, travel time and rescheduling costs', *TESG*, 91 (3), 278–292.

Buehler, R. (2011), 'Determinants of transport mode: a comparison of Germany and the USA', *Journal of Transport Geography*, 19, 644–657.

CBS Statline (2011), http://statline.cbs.nl.

Cullen, I.G. (1978), 'The treatment of time in the explanation of spatial behaviour', in T. Carlstein, D. Parkes and N. Thrift (eds), *Human Activity and Time Geography*, London: Edward Arnold, pp. 27–38.

Cullen, I. and V. Godson (1975), 'Urban networks: the structure of activity patterns', *Progress in Planning*, 4 (1), 1–96.

De Groot, J. and L. Steg (2008), 'Value orientations to explain beliefs related to environmental significant behavior: how to measure egoistic, altruistic, and biospheric value orientations', *Environment and Behavior*, 40, 330–354.

De Groot, J.I.M. and L. Steg (2009), 'Morality and prosocial behavior: the role of awareness, responsibility and norms in the norm activation model', *Journal of Social Psychology*, 149 (4), 425–449.

Dieleman, F.M., M.J. Dijst and G. Burghouwt (2002), 'Urban form and travel behavior: micro-level household attributes and residential context', *Urban Studies*, 39 (3), 507–527.
Dijst, M. (2009), 'Time geographical analysis', in R. Kitchin and N. Thrift (eds), *International Encyclopaedia of Human Geography*, 11, pp. 266–278.
Dijst, M., T. de Jong and J. Ritsema van Eck (2002), 'Opportunities for transport mode change: an exploration of a disaggregated approach', *Environment and Planning B*, 29, 413–430.
Dijst, M. and V. Vidakovic (2000), 'Travel time ratio: the key factor of spatial reach', *Transportation*, 27, 179–199.
Eriksson, L., J. Garvill and A. Nordlund (2006), 'Acceptability of travel demand management measures: the importance of problem awareness, personal norm, freedom, and fairness', *Journal of Environmental Psychology*, 26, 15–26.
Fujii, S. and T. Gärling (2003), 'Development of script-based travel mode choice after forced change', *Transportation Research Part F*, 6 (2), 117–124.
Fujii, S., T. Gärling and R. Kitamura (2001), 'Changes in drivers' perceptions and use of public transport during a freeway closure: effects of temporary structural change on cooperation in a real-life social dilemma', *Environment and Behavior*, 33, 796–808.
Fujii, S. and R. Kitamura (2004), 'What does a one-month free bus ticket do to habitual drivers? An experimental analysis of habit and attitude change', *Transportation*, 30, 81–95.
Gatersleben, B. (2007), 'Affective and symbolic aspects of car use', in T. Gärling and L. Steg (eds), *Threats to the Quality of Urban Life from Car Traffic: Problems, Causes, and Solutions*, Amsterdam: Elsevier, pp. 219–233.
Giuliano, G. and J. Dargay (2006), 'Car ownership, travel and land use: a comparison of the US and Great Britain', *Transportation Research Part A: Policy and Practice*, 40 (2), 106–124.
Golob, T.F. (1984), 'A descriptive analysis of trip-chaining behaviour in the Netherlands', in *Proceedings, Seminar H, PTRC Summer Annual Meeting*, London: PTRC, pp. 295–307.
Golob, T.F., M.J. Beckmann and Y. Zahavi (1981), 'A utility-theory travel demand model incorporating travel budgets', *Transportation Research B*, 15B, 375–389.
Golob, T.F., A.D. Horowitz and M. Wachs (1979), 'Attitude–behaviour relationships in travel-demand modelling', in D.A. Hensher and P.R. Stopher (eds), *Behavioural Travel Modelling*, London: Croom Helm, pp. 739–757.
Goodwin, P. (1983), 'Some problems in activity approaches to travel demand', in S.M. Carpenter and P.M. Jones (eds), *Recent Advances in Travel Demand Analysis*, Aldershot: Gower, pp. 470–474.
Hägerstrand, T. (1970), 'What about people in regional science?', *Papers of the Regional Science Association*, 24, 7–21.
Harland, P., H. Staats and H. Wilke (1999), 'Explaining proenvironmental behavior by personal norms and the theory of planned behavior', *Journal of Applied Social Psychology*, 29, 2505–2528.
Heath, Y. and R. Gifford (2002), 'Extending the theory of planned behaviour: predicting the use of public transportation', *Journal of Applied Social Psychology*, 32, 2154–2185.
Hunecke, M., A. Blöbaum, E. Matthies and R. Höger (2001), 'Responsibility and environment: ecological norm orientation and external factors in the domain of travel mode choice behavior', *Environment and Behavior*, 33, 830–852.
Jones, P.M. (1979), 'New approaches to understanding travel behavior: the human activity approach', in D.A. Hensher and P.R. Stopher (eds), *Behavioral Travel Modelling*, London: Croom Helm, pp. 55–80.
Kitamura, R. and M. Kermanshah (1983), 'Identifying time and history dependencies of activity choice', *Transportation Research Record*, 944, 22–30.
Klockner, C.A., E. Matthies and M. Hunecke (2003), 'Problems of operationalising habits and

integrating habits in normative decision-making models', *Journal of Applied Social Psychology*, 33, 396–417.

Kwan, M.-P. (2000), 'Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set', *Transportation Research C*, 8, 185–203.

Limtanakool, N., M. Dijst and T. Schwanen (2006), 'On the participation in long-distance travel: a decomposition analysis for the Netherlands and the UK', *Tijdschrift voor Economische en Sociale Geografie*, 97 (4), 389–404.

Lindenberg, S. and L. Steg (2007), 'Normative, gain and hedonic goal-frames guiding environmental behavior', *Journal of Social Issues*, 63 (1), 117–137.

Litman, T. (2006), *London Congestion Pricing*, Victoria, BC: Victoria Transport Policy Institute.

McDowell, L., K. Ray, D. Perons, C. Fagan and K. Ward (2005), 'Women's paid work and moral economies of care', *Social and Cultural Geography*, 6 (2), 219–235.

Matthies, E. and A. Blöbaum (2007), 'Ecological norm orientation and private car use', in T. Gärling and L. Steg (eds), *Threats to the Quality of Urban Life from Car Traffic: Problems, Causes, and Solutions*, Amsterdam: Elsevier, pp. 251–271.

Metz, D. (2010), 'Saturation of demand for daily travel', *Transport Reviews*, 30, 659–674.

Miller, H.J. (2005), 'A measurement theory for time geography', *Geographical Analysis*, 37, 17–45.

Nordlund, A.M. and J. Garvill (2003), 'Effects of values, problem awareness, and personal norm on willingness to reduce personal car use', *Journal of Environmental Psychology*, 23, 339–347.

Ory, D.T. and P.L. Mokhtarian (2005), 'When is getting there half the fun? Modelling the liking for travel', *Transportation Research A*, 39, 97–123.

Pas, E.I. (1980), *Toward the Understanding of Urban Travel Behavior through the Classification of Daily Urban Travel/Activity Patterns*, Ph.D. thesis, Evanston, IL: Northwestern University.

Prillwitz, J., S. Harms and M. Lanzendorf (2007), 'Interactions among residential relocations, life course events, and daily commute distances', *Transportation Research Record*, 2021, 64–69.

Rietveld, P. (2001), 'Biking and walking: the position of non-motorized transport modes in transport systems', in K.J. Button and D.A. Hensher (eds), *Handbook of Transport Systems and Traffic Control*, Handbooks in Transport, vol. 3, Oxford: Elsevier Science, pp. 299–320.

Ritsema van Eck, J., G. Burghouwt and M. Dijst (2005), 'Lifestyles, spatial configurations and quality of life in daily travel: an explorative simulation study', *Journal of Transport Geography*, 13, 123–134.

Salomon, I. and M. Ben-Akiva (1983), The use of the life-style concept in travel demand models, *Environment and Planning A*, 15, 623–638.

Santos, G. (2008), 'The London experience', in E. Verhoef, M. Bliemer, L. Steg and B. van Wee (eds), *Pricing in Road Transport: A Multi-Disciplinary Perspective*, Cheltenham, UK and Northampton, MA, USA: Edgar Elgar, pp. 273–292.

Schuitema, G., L. Steg and S. Forward (2010), 'Explaining differences in acceptability before and acceptance after the implementation of a congestion charge in Stockholm', *Transportation Research A*, 44, 99–109.

Schwanen, T., F.M. Dieleman and M.J. Dijst (2003), 'Car travel in Netherlands daily urban systems: does polycentrism result in lower travel times?', *Urban Geography*, 24 (5), 340–360.

Schwanen, T. and M. Dijst (2002), 'Travel-time ratios for visits to the workplace: the relationship between commuting time and work duration', *Transportation Research A*, 36, 573–592.

Schwanen, T., M.J. Dijst and F.M. Dieleman (2004), 'Policies for urban form and their impact on travel: the Netherlands experience', *Urban Studies*, 41, 579–603.

Schwartz, S.H. (1977), 'Normative influences on altruism', in L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, vol. 10, New York: Academic Press, pp. 221–279.

Schwartz, S.H. and J.A. Howard (1981), 'A normative decision-making model of altruism', in

J.P. Rushton, *Altruism and Helping Behaviour: Social, Personality and Developmental Perspectives*, Hillsdale, NJ: Erlbaum, pp. 189–211.

Shen, Q. (2000), 'Spatial and social dimensions of commuting', *Journal of the American Planning Association*, 66(1), 68–82.

Small, K.A. and E.T. Verhoef (2007), *The Economics of Urban Transportation*, London: Routledge.

Steg, L. (2003), 'Can public transport compete with the private car?', *IATSS Research*, 27 (2), 27–35.

Steg, L. (2005), 'Car use: lust and must. Instrumental, symbolic and affective motives for car use', *Transportation Research A*, 39, 147–162.

Steg, L. and J.I.M. de Groot (2010), 'Explaining prosocial intentions: testing causal relationships in the norm activation model', *British Journal of Social Psychology*, 49, 725–743.

Steg, L. and G. Tertoolen (1999), 'Sustainable transport policy: contribution of behavioural scientists', *Public Money and Management*, 19, 63–69.

Steg, L., C. Vlek and G. Slotegraaf (2001), 'Instrumental-reasoned and symbolic-affective motives for using a motor car', *Transportation Research Part F*, 4, 151–169.

Susilo, Y.O. and M. Dijst (2009), 'Behavioural decisions of travel-time ratios for work, maintenance and leisure activities in the Netherlands', *Transportation Planning and Technology*, 33 (1), 19–34.

Susilo, Y.O. and K. Maat (2007), 'The influence of built environment to the trends in commuting journeys in the Netherlands', *Transportation*, 34 (5), 589–609.

UK-NTS (2009), http://www2.dft.gov.uk/pgr/statistics/datatablespublications/nts/.

Van der Klis, M. and C.H. Mulder (2008), 'Beyond the trailing spouse: the commuter partnership as an alternative to family migration', *Journal of Housing and the Built Environment*, 23 (1), 1–19.

Van Wee, B., P. Rietveld and H. Meurs (2006), 'Is average daily travel time expenditure constant? In search of explanations for an increase in average travel time', *Journal of Transport Geography*, 14, 109–122.

Varian, H.R. (1992), *Microeconomic Analysis*, New York: Norton.

Verplanken, B., H. Aarts and A. van Knippenberg (1997), 'Habit, information acquisition, and the process of making travel mode choices', *European Journal of Social Psychology*, 27, 539–560.

Verplanken, B., H. Aarts, A. van Knippenberg and A. Moonen (1998), 'Habit versus planned behaviour: a field experiment', *British Journal of Social Psychology*, 37, 111–128.

Vlek, C., W. Jager and L. Steg (1997), 'Modellen en strategieën voor gedragsverandering ter vermindering van collectieve risico's' [Models and strategies for behaviour change aimed at reducing collective risks], *Nederlands Tijdschrift voor de Psychologie* [Dutch journal of psychology], 52, 174–191.

Zahavi, Y. (1979), *The UMOT Project*, Report DOT-RSPA-DPB-2-79-3, Washington, DC: US Department of Transportation.

4

Freight transport demand: indicators, determinants and drivers of change

Lóránt Tavasszy and Kees Ruijgrok

4.1 Introduction

Freight transport flows are a result of the interplay between economic activities such as production and consumption on the one side and on the other side the supply of logistics and transport services. This chapter discusses how freight transport flows result from the interplay between producers, consumers and providers of logistics and transport services. We examine freight transport demand from three angles:

1. **Indicators** for the description of freight transport flows: measures related to the weight of the goods moved and to transport and traffic performance.
2. **Determinants** of freight transport demand: the decisions of actors within the logistics system who create the need for freight movements.
3. **Drivers of change** in freight transport demand, that is, external forces that influence demand: economic growth, globalization and mass individualization.

The chapter is organized as follows. The indicators for freight transport demand and their interrelations are treated in section 4.2. Section 4.3 describes the determinants of demand. In section 4.4, taking an exploratory perspective, we discuss future trends and trend breaks and their influence on freight demand. We summarize and conclude this chapter in section 4.5. Figure 4.1 visualizes the structure of the chapter, expressing the causality in reverse order: drivers have an impact on determinants, which have an impact on indicators.

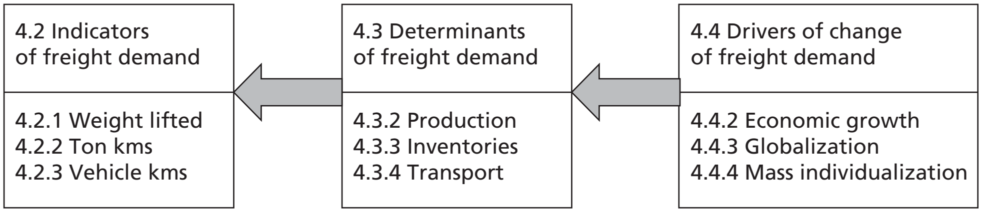


Figure 4.1 Structure of this chapter

We mostly use data material from Europe to illustrate the peculiarities of the freight transport system, also pointing out differences between Europe and other continents in the world, where the available data material allows. As the source of information we used data from Eurostat, which we have used for establishing some of the figures in the next section.

4.2 Indicators of freight transport demand

There are various ways to describe the volumes and patterns of freight transport demand and the developments in the different modes of transport. We elaborate on three key indicators:

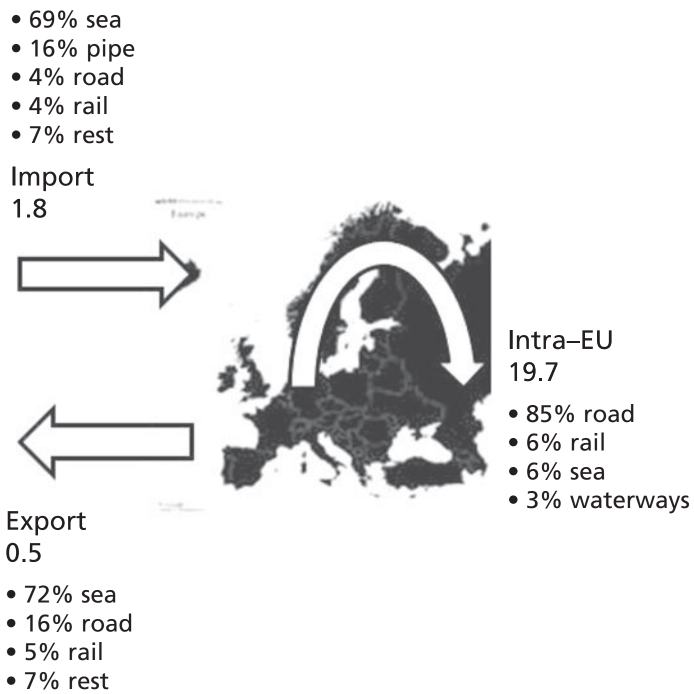
1. Volume of transport, measured by the **weight lifted** of the goods. This is measured by the mass of the goods expressed in metric tons.
2. **Transport performance**, which includes the distance moved of the freight (ton kilometres). This integrative indicator is a measure of the need for movement and used most often to characterize freight transport demand.
3. **Traffic performance**, measured in distance moved by the transport demand means in question (e.g. vehicle kilometres).

We also discuss the interrelationships between the three and conclude the section with a discussion of the economic importance of freight transport.

4.2.1 Weight lifted

Volumes of freight transport are generally expressed in transport statistics by means of the weight of the freight which is moved. Figure 4.2 shows the main volumes of transport within and to or from the European Union in 2006.

We can see that the use of transport modes depends strongly on the type of relation. Within the EU, road transport dominates, and alternative modes such as rail and inland waterways carry far less freight, with volumes about an



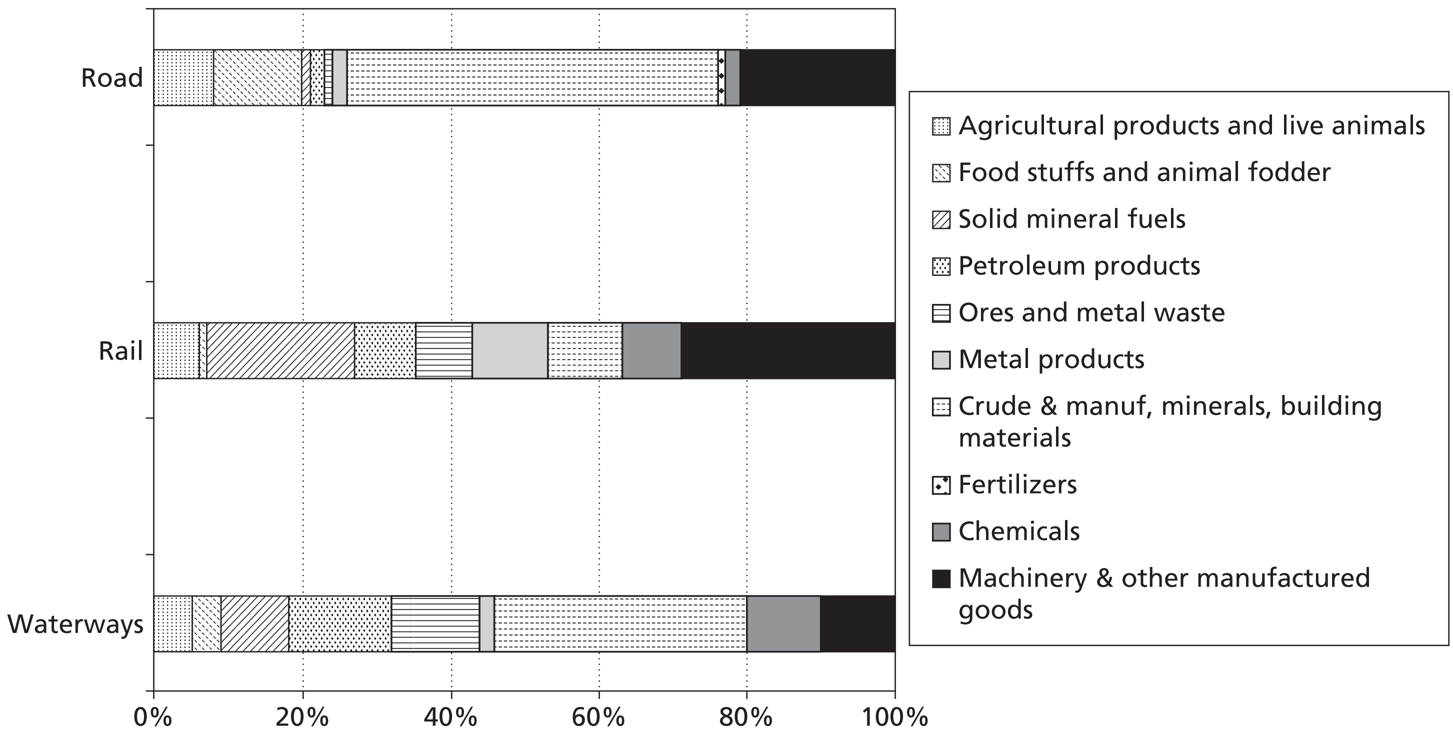
Note: All flows in billion tons, 2006; rest = other modes and unknown.

Source: Eurostat, reworking and presentation by the authors.

Figure 4.2 Freight transport flows by mode within and to or from Europe

order of magnitude lower than road. Owing to the relatively short transport distances and the fragmented national systems for rail transport, this mode has a considerably lower share than in other regions of the world, where the rail shares lie in the range between 10 per cent for the USA (FHWA, 2010) and 40 per cent in Russia (Rosstat, 2010). Sea transport dominates intercontinental flows but also has an important role for intra-European, international traffic: deep sea and short sea together carry as many tons of freight between countries as the railways. This important role for sea for continental transport is also found in Asia and much less so on other continents.

A second important characteristic of freight demand is its heterogeneity. There are many types of commodities that are associated with various sectors, which each organize their movements differently. This heterogeneity is nicely illustrated by the use of modes of transport by the different commodity groups. Figure 4.3 shows the distribution of goods types across the different inland modes of transport. The specification of goods types used here is the European NST/R[1] classification. We see that almost half of the



Note: Tonnes, EU27, 2006.

Source: Eurostat.

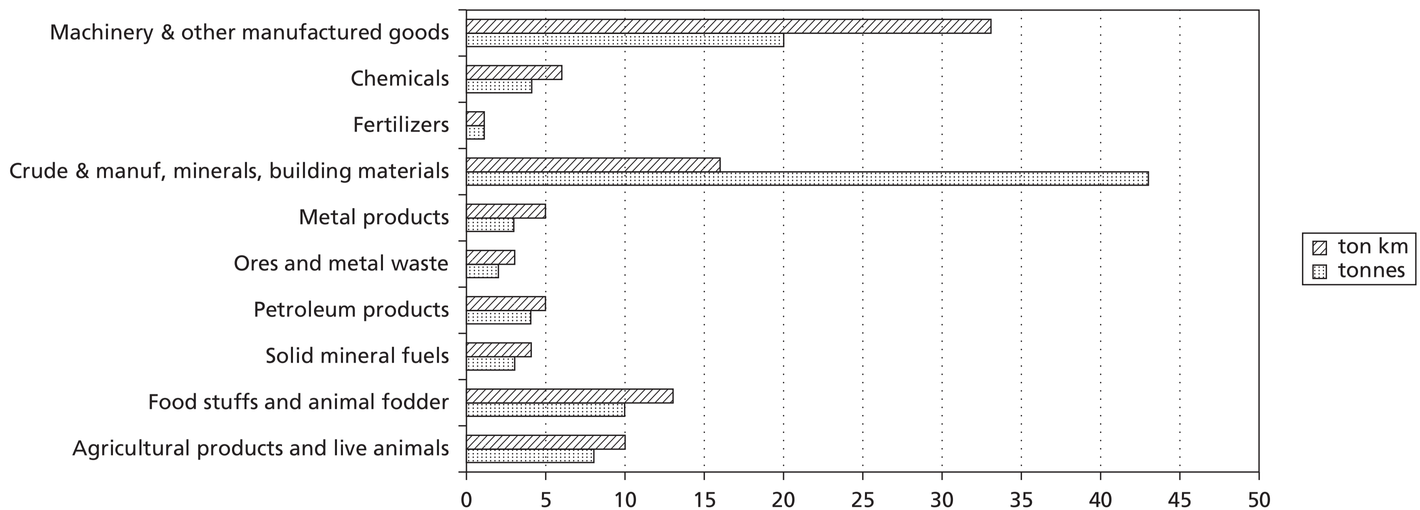
Figure 4.3 Weight moved by mode and commodity type weight of road transport is occupied by building materials. This is due to the short distances that these materials move; here, road transport offers the most flexibility in building projects. As we will see further, this does not translate into a high transport performance (in ton kilometres) because of the short distances. The weight share of the other two modes is dominated by bulk products such as solid and liquid fuels, ores and building materials (e.g. sand).

When measuring freight transport demand in tons we need to interpret statistics with care, as tons lifted is not necessarily the same as tons produced or consumed. As individual transport movements are often part of a larger transport chain, the same unit of freight may appear more than once in the statistics if the chain has several transport modes. Every time goods are unloaded and loaded on to a further mode of transport, they will be counted a second time, as registration occurs on a transport mode basis. One ton of goods produced may thus appear as two tons or more in transport statistics. This implies that the number of tons will depend on the structure of the transport chain and the transport technology used (intermodal or door-to-door unimodal transport). When we measure freight demand in ton kilometres, this risk is less severe.

Another drawback of tons lifted as a measure of freight demand is that it does not reveal much about the economic importance or social impacts of transport. When considering the economic impacts, the value density of goods may vary by several orders of magnitude, from 10 USD/ton for raw minerals such as sand and gravel to over 100 000 USD/ton for manufactured products. Secondly, indicators such as transport costs, energy use and emissions of transport are all distance related. Here, an indicator that includes distance will be a more accurate basis for examining the social costs and benefits of transport, its energy use and the environmental effects.

4.2.2 Transport performance

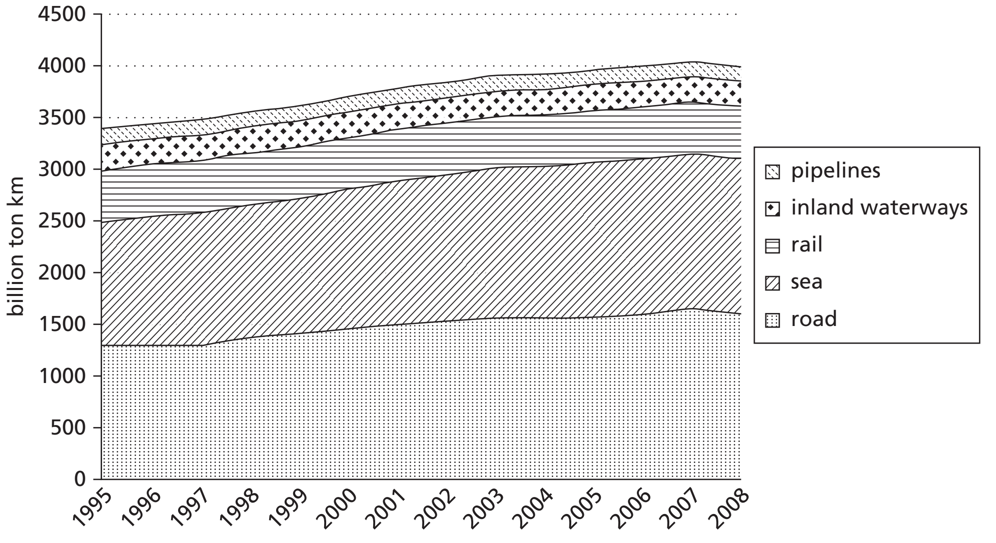
We measure transport performance in ton kilometres, the amount of tons moved multiplied by the distance that it is moved across. This measure provides a more useful measure for transport policy analysis than the tons lifted, as it is more closely related to the internal and external economic effects of freight transport (see Chapter 12 for an explanation of internal and external effects). It also provides quite a different picture from that of the tons lifted, as shown in Figure 4.4. The figure shows that crude and manufacturing minerals and building materials are transported over relatively short distances,



Note: EU27, 2006.

Source: Eurostat.

Figure 4.4 Share of commodity types moved within Europe measured by two indicators



Source: Eurostat.

Figure 4.5 Performance by mode of transport

whereas machinery and other manufactured goods are transported over relatively long distances.

The performance of the different modes of transport is shown in Figure 4.5. The figure clearly shows that the strong growth of freight transport in Europe is dominated by the road and sea modes. The other modes remain constant in terms of performance; their share in the total transport in Europe has been decreasing steadily. Air transport has been growing at the same rate as road transport but is invisible in the figure, as its yearly transport performance is about 0.15 per cent of that of road transport. The trend break in the last year (from 2007 to 2008) happened for all modes of transport and shows the effects of the global financial crisis of that time.

Although this measure for transport performance is easier to interpret from a policy perspective than the weight measure, we note that even adding ton kilometres together should be done with care, as the shipment sizes vary widely between and within modes. Taking one ton of freight over 1000 kilometres will involve a different usage of resources than 1000 tons over one kilometre because of economies of scale and density, or the technology available for that size of shipment and distance. Taking into account the number of vehicles needed to move freight eliminates part of the problem.

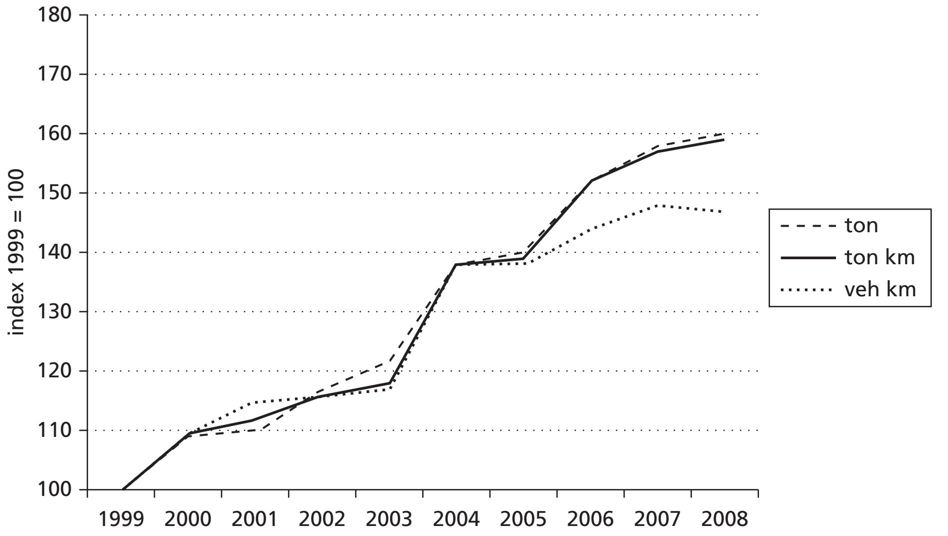
4.2.3 Traffic performance

Once we can convert tons moved into the number of vehicles, we can express freight demand in terms of traffic, which provides us with additional information for transport network design and the measurement of transport costs, energy use and so on. Many policy measures are related to the individual vehicles and, directly or indirectly, to the distances driven (e.g. taxes, tolls, permits). Knowledge of vehicle kilometres driven makes assessment of these measures easier. Average loading factors vary by country but typically lie between 40 per cent and 60 per cent when measured in weight; empty running typically varies between 20 per cent and 40 per cent (see, for example, Eurostat, 2007). Note that a loading factor of 50 per cent may seem low, until we consider that a full truck leaving for a round trip and arriving empty will be half-full on average. The only option is to pick up freight on the way, which for many companies is difficult to organize. We note that these measures relate to the truck capacity in weight and not vehicle size. The real efficiency of the system is probably higher.

4.2.4 Evolution of the different indicators

The development over time of the three indicators – tons lifted, ton kilometres and vehicle kilometres – can shed light on structural changes in freight demand patterns. Figure 4.6 compares these three indicators for road transport within the 27 countries of the European Union. This figure shows that, through time, the three indicators have developed closely in parallel. On average, transport distances have not changed during the last decade. The final year in this figure seems to indicate that the financial crisis has had some effect on transport distance, possibly owing to the fact that imports and exports decreased. Some products are sensitive to economic cycles. If these are transported over relatively long distances, the percentage change of the indicator for total transport performance (expressed in ton kilometres) will be higher than the percentage change of the indicator 'total weight lifted', throughout the cycles. Figure 4.6 also shows that a notable improvement of efficiency in terms of vehicle loads has taken place: more ton kilometres required fewer vehicle kilometres. Even though loading factors are decreasing owing to increasing individualization of consumer markets and specialization of the industry, empty running is declining as well, owing to improved backhaul possibilities. The net balance in the recent past has been an increase of efficiency of freight transport by road.

Demand for freight transport is generated through decisions made by shippers and transport companies that operate on the market of freight transport



Note: EU27.

Source: Eurostat.

Figure 4.6 Freight transport by road by the three different indicators

as representatives of the demand side and of the supply side respectively. In the next section we focus on the decision and the decision makers from the demand side.

4.3 Logistics determinants of freight transport demand

4.3.1 Introduction

Transport demand is said to be a derived demand, as it depends on the needs for services of manufacturing sectors and consumers. We characterize the logistics system that determines the demand for transport, as one consisting of five elements: sources, production, inventories, transport and sales (SPITS). The way the elements interact was described in the SPITS model developed in the early 1990s (Kuipers et al., 1995). The elements of the system follow the supply chain from raw material (source) to the final demand by the customer (sales). The components P, I and T are central in this system and interact with each other to connect source and sales, by optimization of the supply chain, given the service requirements of the customers (Christopher, 1992). We summarize this objective into the following optimization function:

$$\text{Min GLC} \mid \text{Service}$$

$$\text{GLC} = \text{T} + \text{I} + \text{H},$$

where GLC is the generalized logistics cost (which is the sum of transport T, inventory I and handling H), to be minimized given the service requirements (Service). To the extent that service requirements do not act as a constraint on the design of logistics services, there is scope for further optimization. In most cases, however, there will be service requirements (e.g. pick-up and delivery time) that bind the solution.

The possibilities for minimizing the generalized logistics costs rely on the opportunities to use economies of scale and scope as well as smart trade-offs between the three cost components of GLC. Supply chain management aims at exploiting these optimization possibilities. In many logistics processes these scale economies are essential to achieve a lower cost per item. Scale economies can be achieved by, for example:

1. bundling individual products into larger shipments;
2. using larger-scale modes of transport;
3. combining inventories upstream into central warehouses.

In this section we describe in further detail these three main determinants of the logistic system: P(roduction), I(nventories) and T(ransport). We do this by highlighting the main trends of the last few decades in the development of these activities and the strategic decision factors that have governed the changes in their spatial and functional design.

4.3.2 Production

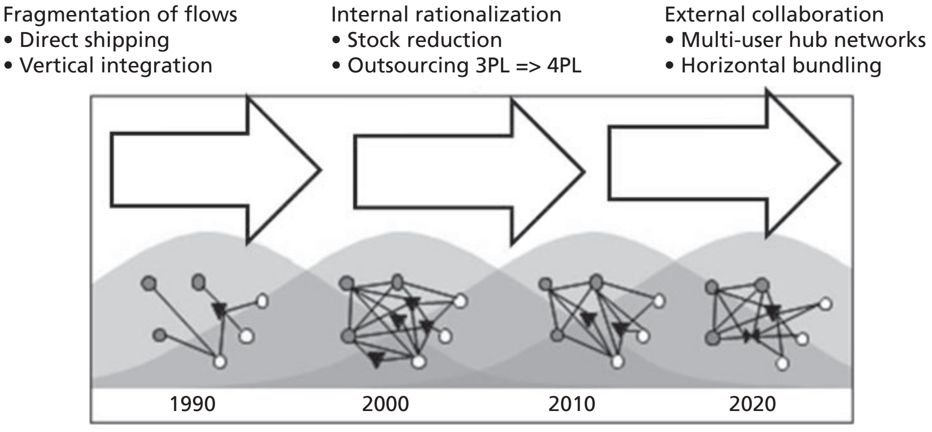
The world economy has emerged from a period with a high degree of economic protection and isolation into the present state that is characterized mainly by free trade and a high degree of specialization. One of the main drivers behind this growth of trade has been the differences in cost of producing the same type of product in different places around the world, which are due to differences in factor costs and the availability of natural resources. As a result of reduced trade barriers, production moved from the West towards Eastern Europe, Eurasia and the Far East. Together with the cost of overcoming the distance, one can determine whether it is more attractive to import the products from elsewhere and to carry the burden of transporting goods over large distances or to avoid the costs involved in transporting these goods and produce them locally. The number of production steps increases as companies focus on their core business, trying to improve their competitive power by specializing and gaining scale advantages. This disintegration of the chain applies to services as well as production steps and is also a reply to a customer base that is increasing geographically.

Both the organization of production and transport economies of scale play an important role in the choice of production and physical distribution. A well-known example is the production of automobiles. In general the assembly of automobiles takes place not too far from the final customer, but some of the parts are produced by factories that distribute their products to customers spread worldwide. The location of production plants normally is a long-term investment decision, and thus the geographic spread of production patterns used to be rather stable over time. Nowadays, the location of factories is re-evaluated more frequently, and those that are not ideally located or have a lack of governmental support are under the threat of being closed down. Assembly plants, however, are more footloose and their location can change, influenced by regulatory measures (subsidies, regulation on the share of local content), the relative importance of transport costs in the cost of final products, congestion and other capacity restrictions (Dicken, 2003).

4.3.3 Inventories

Supply chain management techniques have evolved strongly in the last few decades. In the 1980s and 1990s, companies rationalized their distribution structures by decreasing the number of warehouses. Currently supply chain management goes far beyond just-in-time deliveries between businesses to minimize on inventories. Trading partners have learned to collaborate throughout the supply chain in order to influence or respond quickly to consumer demand. Supply chain management techniques developed in the last decade go under various names which we will not treat in detail here,[2] such as quick response, lead-time management, lean logistics, agile logistics, efficient consumer response, and process and pipeline mapping. These techniques have helped firms to reduce their inventories drastically while increasing service levels. Low-volume but high-value segments have greatly expanded, benefiting from these new techniques and the opening up of new markets for customized and highly responsive services.

The advent of supply chain management, made possible by the mass introduction of personal computing and data processing facilities, was marked by a professionalization of firms' service quality improvement and cost control



Note: The dark dots (top left) indicate suppliers, the bright ones (bottom right) indicate consumers, and the black figures denote warehouses (triangles) and cross-dock locations.

Source: Own composition.

Figure 4.7 Evolution of logistics networks through time

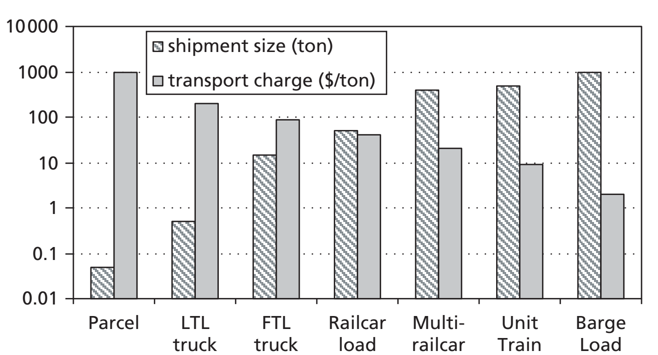
practices. Their evolution has gone through a series of cycles (see Figure 4.7). The first wave of change took place in the 1990s and involved the reduction of shipment sizes, an increase in frequencies and just-in-time transport, as a first sign of mass individualization. This caused a fragmentation of goods flows into smaller streams. Driven by the need to keep costs under control, this was followed by a second wave of development, which involved the internal rationalization of logistics processes within the company's own supply chain. In a recent third wave of change, firms are looking for economies of scale by means of external collaboration, across the company boundaries and their own supply chains. This so-called horizontal collaboration (as opposed to vertical cooperation between vendor and supplier firms within the same supply chain) is seen as one of the major innovations which will transform the logistics business landscape in the coming years (see, for example, Mason et al., 2007).

Although the increased transparency of the logistics process in supply chains and the technological means to plan and control logistics much better using advanced planning tools and IT systems have increased, we see that the practical application of these technologies is often hampered by a lack of cooperation and standardization. Real supply chain cooperation is then difficult to achieve because of countervailing powers between the supply chain partners. On the other hand, in order to achieve substantial cost efficiencies this type of cooperation is essential to synchronize the activities in the supply chain and to take full benefit of potential scale economies.

4.3.4 Transport logistics

Transport logistics concern the decision about the mode of transport to be used and, at the tactical and operational level, the planning of movements and the routeing of the freight. We limit our discussion here to the strategic-level decision of mode choice. As we have seen above, companies increasingly favour road, air and sea transport. Part of the explanation for this lies in external, socio-economic factors, like globalization and the individualization of society (we will return to these factors in the next section in more detail). In order to understand how such developments can translate into changes in mode usage we need to look into the mechanisms that govern these choices. Below, we discuss the influence of product and service characteristics on mode choice.

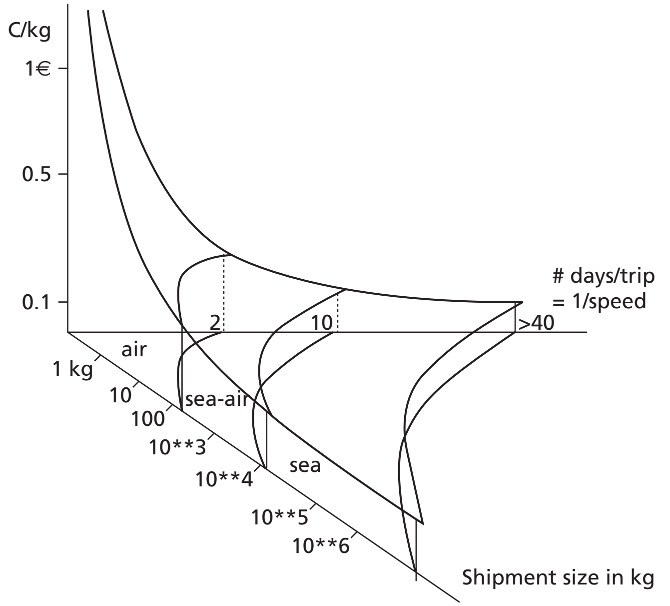
Jordans et al. (2006) found that 95 per cent of the mode choices can be related to transport distances and basic product characteristics such as value density and packaging density. For many transport flows, transport distances and basic product characteristics are given and cannot be influenced. At the same time, mode choice is also governed by preferences of firms that are more difficult to observe and relate to the performance characteristics of transport modes such as transport time reliability, lead time and prices per shipment (see, for example, Vieira, 1992). We illustrate this relationship between supply chain characteristics and mode choice by means of the element of shipment size. The size of shipment is an important degree of freedom in the minimization of GLC given required service levels. Figure 4.8 shows the average shipment sizes and related transport prices of some modes of transport. From this picture it becomes clear that huge differences exist



Note: LTL = less than truckload; FTL = full truckload.

Source: Adapted from Rodrigue (2006a).

Figure 4.8 Differences in shipment size and transport charges for different modes of transport



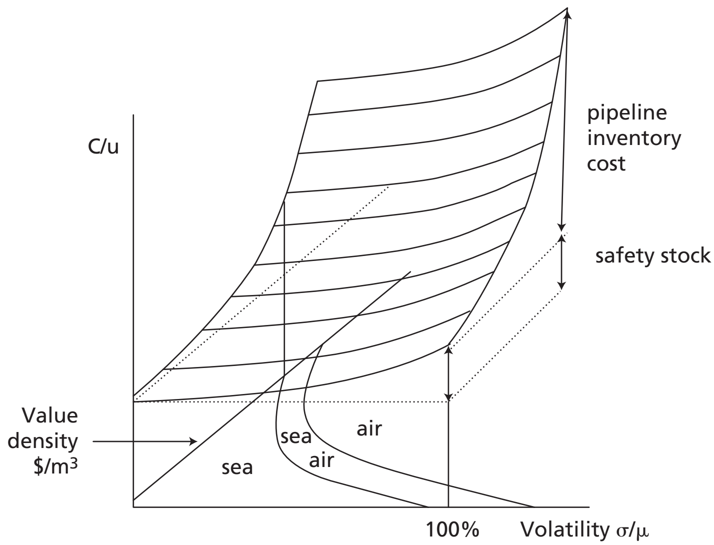
Source: Tavasszy et al. (2009).

Figure 4.9 Transport cost per unit for overseas relations, as function of speed (distance per time unit) and shipment size

between the respective modes. Mode choice will be determined in part by the size of the shipment, as economies of scale in terms of shipment sizes can easily materialize by choosing a mode that allows lower unit costs.

Because of these large differences and the limited choice flexibility, it is possible, given a limited number of exogenous factors, to specify an a-modal or mode abstract generalized logistics cost function, such as the one shown in Figure 4.9. This figure shows an example of how two long-distance overseas modes (air and sea transport, as well as combinations of sea and air) can be efficient. By specifying the weight of the shipment and the required speed, the mode choice and the transport cost per kilogram can be derived easily, if all modes are available. Thus 100 000 tons of crude oil are transported by ship, and a box of diamonds is shipped by air. There will be no discussions on the mode choice decision in these circumstances, at least in densely populated developed economies, where all above-mentioned modes of transport are present and the speed requirements of the shipments do not restrict the choice options dominantly.

Besides shipment size and speed there are two other determining factors that influence the modal choice, which is the value density of the product



Source: Tavasszy et al. (2009).

Figure 4.10 Logistics cost per unit for overseas relations, as function of value density (value per m[3]) and volatility (σ / μ)

and the level of demand uncertainty (see Figure 4.10). When taking into account the value of the product and the volatility of demand (in the figure: demand variance σ divided by mean demand μ) we can also visualize the effect of inventory costs via increased safety stocks and the effect of pipeline inventory costs, adding to total costs/unit (*C/u* in the figure). When the value density is low, pipeline costs (the inventory costs during transport) will be negligible. When the value density becomes larger, the pipeline cost becomes significant. We provide an example of this trade-off below for the concrete case of laptop computers.

Example: laptops

In the case of a shipment of one container with 1000 laptops (20 pallets of 50 laptops) with a production value of $500 per laptop, each container will have a pipeline inventory cost of $5000 if the trip takes 36 days and the yearly interest rate is 10 per cent. The average transport rate of this container from Asia to Europe is $1500, so the pipeline cost for this shipment will exceed the sea transport cost by more than a factor of 3. The generalized logistics costs of $6500 would be roughly the same if these products had used the air mode. At 3 kilograms per laptop at $2 per kilogram, the transport charge would be $6000, whereas the pipeline costs would not be more than $500, owing to the short transit time.

The example shows that, although transport costs differ a lot per mode of transport, generalized costs show less variation, taking into account other logistic cost factors. When the volatility is high, retailers and distributors do need safety stocks in order to avoid empty shelves if the demand for a product is higher than the stock and the demand during the reorder period. Safety stocks can be avoided to a great extent if fast and reliable transport options exist that can guarantee the delivery of products within the customer service requirements. So trade-offs exist between inventory costs and transport costs, and the generalized cost concept should take these trade-offs into account. For further reading on how logistics costs can be used in explaining and forecasting the demand for freight transport interested readers are referred to Tavasszy et al. (2012), which also explains the way other drivers of change, described in the next chapter, will affect the volume and composition of freight transport demand.

4.4 Drivers of change in freight transport demand

4.4.1 Introduction

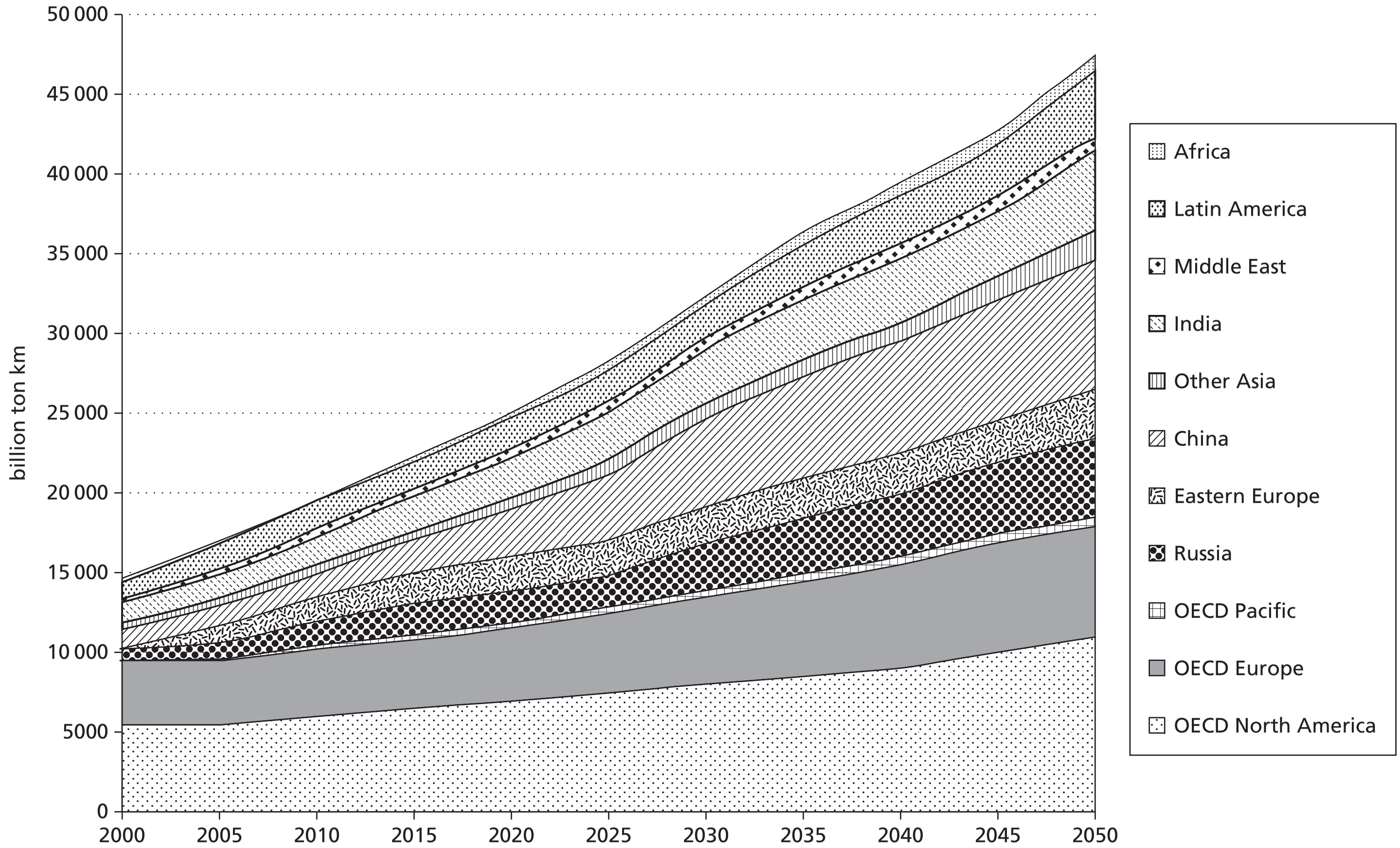
Over the longer term, continued growth of global freight flows is expected (see Figure 4.11). Although this growth will be most conspicuous in the emerging Asian economies (especially China and India), flows are expected at least to double in the first half of this century in all regions of the world. Some sources predict a doubling of present flows in half a century (WBCSD, 2004).

Within the EU, freight transport has doubled within a period of 30 years and forecasts are still equally strong. Land-based freight transport is expected to more than triple in developing countries. Apart from an increase in welfare due to economic growth and technological development, the expectation of further growth of freight transport is explained by a decrease in barriers to international trade and transport.

In the next sections we discuss three drivers of change in freight transport demand that have influenced production and consumption, trade, logistics and transport decisions: (1) economic growth, (2) globalization and (3) mass individualization. We discuss these main drivers in three separate subsections and conclude the section in 4.4.5.

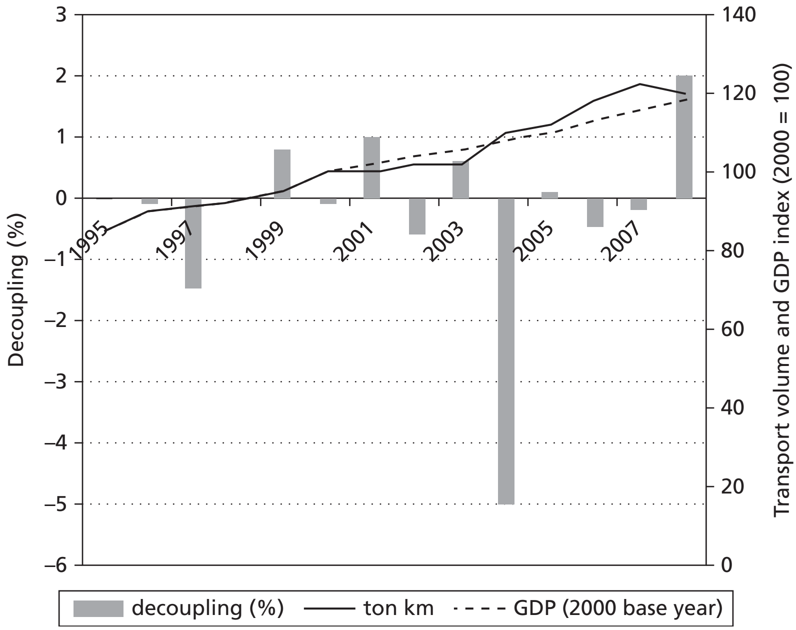
4.4.2 Economic growth

GDP appears to have a strong overall explanatory power for freight transport growth: earlier World Bank research suggests that it explains over 89



Source: WBCSD (2004).

Figure 4.11 Forecasts of growth in transport performance by land modes

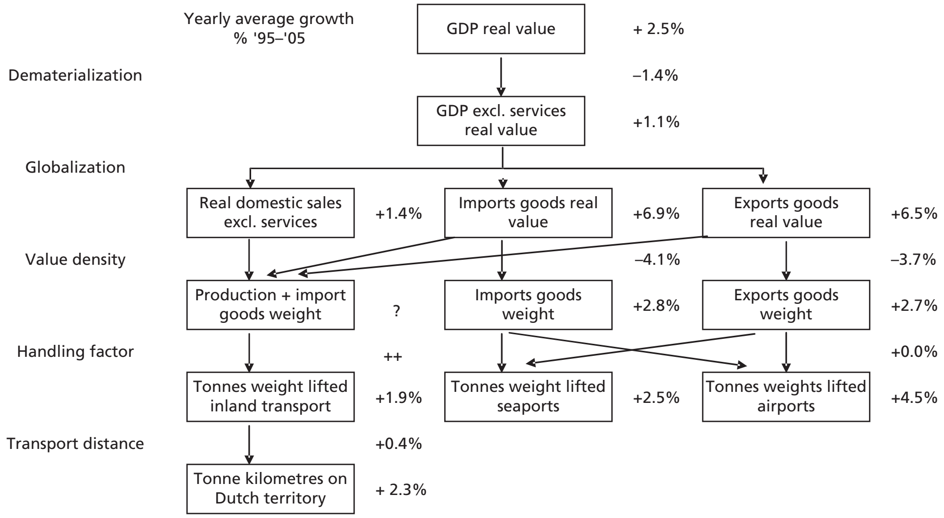


Source: Eurostat.

Figure 4.12 Relationship between GDP and transport growth within the EU

per cent of variation in observed freight volumes (Bennathan et al., 1992). In this study of about 17 countries, a fairly reliable indication of this relation was that every million US dollar in GDP would add 170 000 road ton kilometres in transport performance. In other words, economic performance seems to be a strong driver of transport performance, and we can expect economic growth to explain changes in freight transport flows as well. If we look at the relation between GDP and transport performance over a longer time period, we find that this relationship is quite stable, with more or less the same growth rates for a longer period (Figure 4.12).

Some recent studies indicate, however, that growth of GDP and transport has been decoupling, with the possible implication that GDP would be a less important driver in the future than it is now. A number of countries – Finland, Denmark and the UK – appeared to show a trend of decoupling (see Tapio, 2005; Verny, 2005; Kveiborg and Fosgerau, 2007; McKinnon, 2007).[3] Within the US, decoupling between GDP and freight activity had already taken place (Gilbert and Nadeau, 2002). Indeed it seems that domestic transport growth is decoupling from national GDP, owing to dematerialization of flows and changes in logistics. At the same time, the economic integration between countries results in a growth of international freight flows.



Source: KiM (2006).

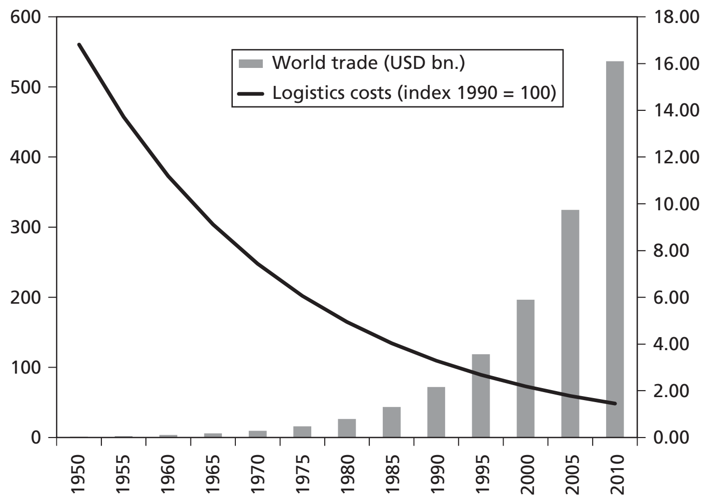
Figure 4.13 Causal analysis of changes in yearly transport performance

The net sum of these influences depends on the relative share of domestic and international transport and the rates of growth in each segment. An illustrative example is the analysis of the relationship between GDP and transport for the Netherlands (KiM, 2006). Figure 4.13 indicates the typical counteracting forces of an international economy such as the Netherlands. The net increase in transport performance that occurred during these years can largely be explained by a relatively strong growth in international flows on top of a modest growth in the domestic economy.

In conclusion, it seems that GDP is, and will remain, a strong determinant of freight growth, be it at domestic or international level.

4.4.3 Globalization and cost changes

Throughout the centuries, national economies have become more closely integrated owing to a decrease of trade barriers and improved technologies for international transportation and telecommunications. The growth of world trade is directly linked to the demand for international freight transport, maritime transport and air freight in particular. Although the share of transport costs in total logistic costs for expensive products is much less than for low-value bulk products, considerable cost savings have also been achieved for these products, because of both economies of scale and improved supply chain management. Advances in telecommunications and



Note: Based on: trade barriers – WTO; transport costs – Eurostat; logistics costs – ATKearney.

Source: Tavasszy (2010).

Figure 4.14 Development of logistics costs and global trade

information technology have given companies the means to manage the physical movement of products over long, often circuitous, routes (Anderson and Leinbach, 2007). Many carriers have invested heavily in 'track and trace' systems to be able to establish the location of any consignment at any time, improving the visibility of the global supply chain to shippers and their customers (see HIDC, 1998). In the coming decades, further changes are expected in the organization of global production networks owing to a further rationalization of supply chains and ongoing growth of global trade.

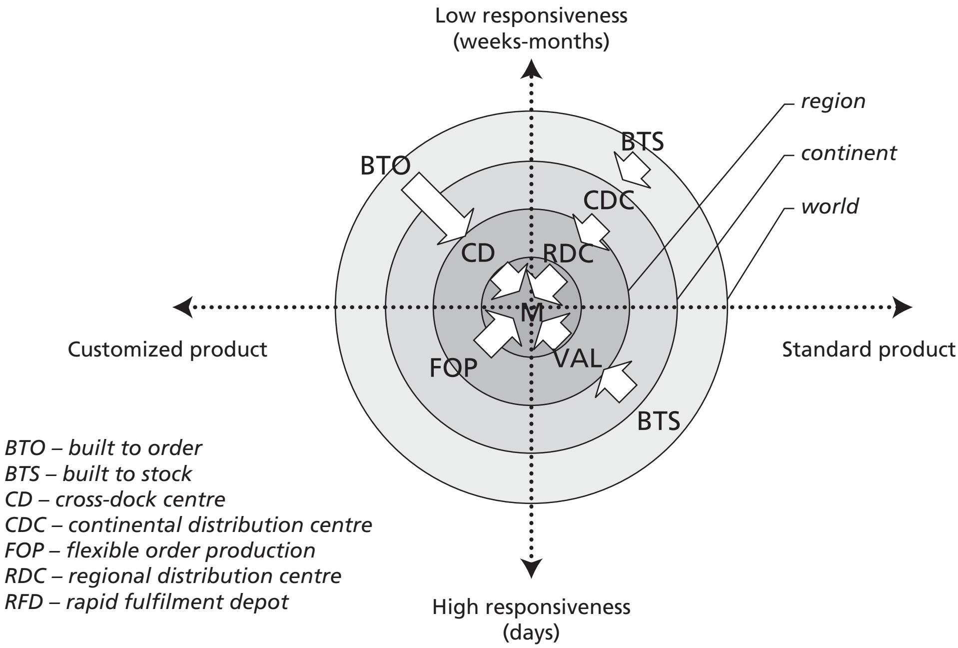
In the future, logistics and transport will be much more tightly integrated with production systems. Complex trading networks have evolved primarily to exploit labour cost differences, regional production specialization, global product differentiation opportunities and availability of raw materials in particular countries. Their development has also been facilitated by major regulatory and technological trends. Trade liberalization, particularly within trading blocs such as the EU and NAFTA, has removed constraints on cross-border movement and has reduced related 'barrier costs'. Along with the reduction in transport costs, the relaxation of trade barriers since the Second World War has given a great stimulus to the development of world trade (Hesse and Rodrigue, 2004; Rodrigue, 2006a, 2006b) (see Figure 4.14).

Transport systems all over the world, but especially in Europe, have benefited from increasing deregulation and cost decreases. By now, however, the possibilities of optimizing transport systems within a deregulated environment have largely been depleted. Road transport is facing new cost increases, owing to tightened safety and social regulations, congestion, road pricing and environmental pressure. Although we expect that this will have only a minor effect on the overall freight transport volumes, the underlying patterns of freight distribution will change because of a further rationalization of freight movements. As unit transport costs increase, we expect that distribution structures will be driven into more decentralized patterns. In addition, firms will shift part of their freight to alternative modes of transport. This trend will be reinforced by the internalization of external costs (e.g. through CO_2 emission rights), which will lead to a trend towards reduction of avoidable transport kilometres, towards the transport of higher volumes at lower speeds at a decreased frequency.

4.4.4 Mass individualization

Since the middle of the previous century there has been an increase in product variety, up to the level of single product or service units being individualized and unique. This is the result of the trend towards adjusting the supply of products to consumer preferences and using the principles of mass individualization. The reduction of logistics costs and globalization have allowed firms to develop mass-individualized products and services.

What is the implication of long-term changes in mass individualization on logistical structures and on transport flows at various spatial levels (local, regional, continental and global)? The effect of mass individualization upon production and distribution systems becomes visible through two main dimensions: product customization and increased responsiveness of services (see Vermunt and Binnekade, 2000; Lee, 2001). As services become more responsive and products more customized, new chain configurations will be deployed to satisfy product and service demand. The adaptation of chain configurations entails changes in production systems (building to order, flexible production, smaller batches) and relocation of production towards the consumer regions, if products are highly customized and response times are short. New chain configurations also need to allow a fast response towards changes in volume or the specification of orders – this implies that inventories move closer towards the end consumer. Value-added logistics and postponed manufacturing allow products to be tuned to individual consumer needs. Flexible logistics structures allow for adaption to variable service requirements, which means that different configurations will need to be maintained to serve different types of consumers (e.g. those with standard orders and those with urgent ones). As a result, supply chains will be



Source: Adapted from Vermunt and Binnekade (2000).

Figure 4.15 Logistics structures by demand segment

reconfigured into supply networks, where chains cross at strategic exchange locations such as warehouses, cross-docking centres or intermodal terminals.

These new chain configurations will vary according to the degree of customization and the degree of responsiveness required. Figure 4.15 shows examples of different supply chain configurations that result from the change from standard to customized products and conventional to responsive services. The figure includes the geographical dimensions from the global scale (outer ring) towards the local market (noted as M in the centre of the figure).

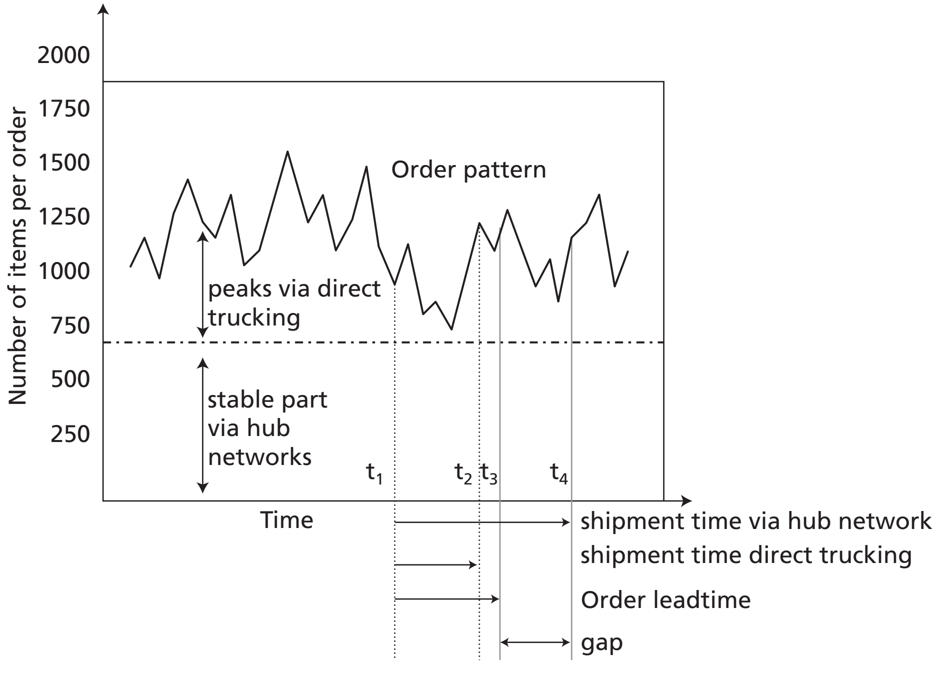
Typical changes in chain configuration concern the move from continental distribution centres, based on production to stock, towards production to order at a global scale, where delivery takes place directly or with an intermediate step of value-added logistics (e.g. packaging to prepare for the local market). New concepts like rapid fulfilment depots (for low-demand but urgent products like spare parts) and flexible order production (allowing fast switching in batch size and end-product specifications, close to the consumer market) are also being introduced to allow for better responsiveness. The more individualized products are, the more these activities will be located close to consumer markets. Centralized international distribution, introduced to reduce inventory and building on the decrease in trade barriers, is being supplemented by regional distribution centres.

4.4.5 Future logistics structures: differentiation and dynamics

The move towards more complex logistics structures will inevitably drive up logistics costs as well as the share of logistics costs in product costs. Only through a further rationalization of logistics processes will firms be able to control their logistics costs associated with increased quality of products and services.

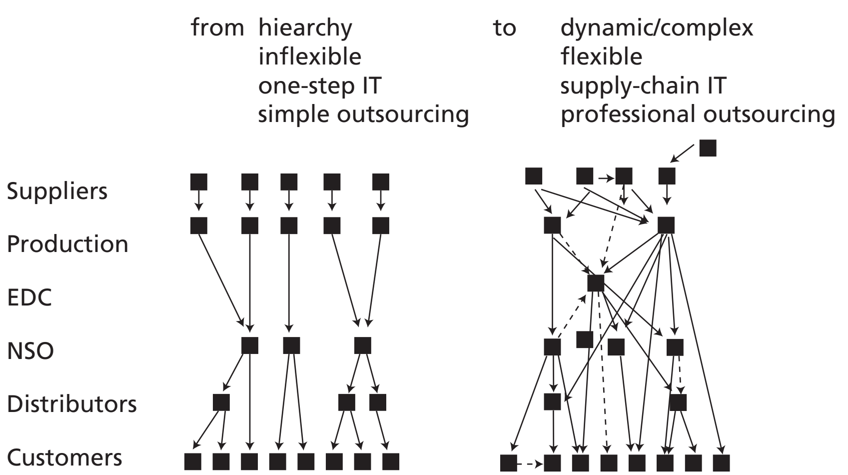
Consolidation and cooperation between chain partners are the most logical ways to generate lower cost per unit of freight. Through consolidation of flows, larger vehicles can be used and the loading efficiency is optimized. Through collaboration, the synchronization of logistic activities also becomes possible, which results in a much smoother, seamless flow of goods through the logistic system, and therefore in higher utilization of resources, but it also creates the possibility of using cheaper and slower modes of transport and avoids the need for safety stock (Groothedde, 2005). Note that the high level of responsiveness that is required could possibly conflict with the above-mentioned need for smoother flows of goods. Avoiding this possible conflict is one of the biggest challenges in the design of logistic networks. The set-up of hybrid supply chains (which create different possibilities for flows to reach their final destination), for production, warehousing and transportation, creates the flexibility required. Part of the production with a demand pattern that can be predicted well in advance is produced at far-away locations that have low labour costs. The rest of the production is postponed to the last possible moment at locations close to the customer (Groothedde et al., 2005). One of the ways this can be achieved is through the application of hybrid networks. Through the use of hybrid networks a flexible way of quickly adapting from one supply source to another can be created. How this principle works is clarified in Figure 4.16. The volatile part of the demand is supplied by a fast (and more expensive) network, while the stable part of the demand is delivered through the slow but cheap hub network that makes maximum use of economies of scale.

Hybrid networks can combine the advantages of both network alternatives and thus create a higher level of efficiency and flexibility. Some of these possibilities are clarified in Figure 4.17. This figure highlights some of the possibilities for creating hybrid structures through bypassing some of the echelons in a strict hierarchical network, which interconnects the prime suppliers with the ultimate customers.



Source: Groothedde (2005).

Figure 4.16 Being responsive by combining different networks



Note: EDC – European distribution centre; NSO – national sales organization.

Source: Van Goor et al. (2003).

Figure 4.17 Flexibility in hybrid networks

4.5 Conclusions

The most important conclusions of this chapter are as follows:

1. In this chapter we described a number of different dimensions of freight transport demand. We have examined three basic indicators: the total weight lifted (tons), transport performance (ton kilometres) and traffic performance (vehicle kilometres). The relationship between these three depends on the structure of the supply chains in terms of production, trade relations, inventories and transport modes used. We have looked at possible causes of changes through time. In brief, transport is growing across all these dimensions of measurement, and is increasing in efficiency. Freight transport growth is uneven across modes, with the road and sea modes taking the major share of increase in demand.
2. Explaining these changes and their interrelations is difficult, as demand for freight transport results from a large number of factors and is not homogeneous. It is the type of products that are being transported and the customer requirements for these products that determine which type of transport solutions will be relevant. We have introduced generalized logistics costs as an important determinant in the process of logistics optimization, given demand requirements. Product characteristics like shipment size, value density and demand uncertainty are extremely important for the evaluation of alternatives in transport. We have seen that through combinations of smart planning and network design it is possible to minimize logistics cost and maximize customer service at the same time.
3. The process of logistics cost minimization over time has been helped by deregulation and liberalization of trade regulations and improvements in transport and information technology. These developments have led to a constant decrease of logistics costs over time, which has facilitated the growth of international trade to a large extent. In addition, mass individualization leads to a change in the organization of supply chains, which places higher demand on the responsiveness of transport systems. For future developments in freight transport demand, the dominating factor is the growth in GDP. As the growth in world trade is expected to continue to grow after a period of decline or dampened growth due to the worldwide economic recession, it is expected that the growth of transport demand will go on and this will lead to a continued demand for transport infrastructure and also a continued challenge to find sustainable solutions.

NOTES

1 Nomenclature uniforme des marchandises pour les Statistiques de Transports, Revisée. This classification has 10 main commodity segments, broken down further into detailed categories.
2 We refer to CSCMP (2010) for a glossary of terms in supply chain management.
3 We note that this is partly a result of changes in freight operations that cause transport flows to fall outside the sampling system of statistics offices. These included changes in vehicle choices from trucks to vans and the entry of foreign transport companies into the trucking market (McKinnon, 2007).

REFERENCES

Anderson, W.P. and T.R. Leinbach (2007), 'E-commerce, logistics and the future of globalized freight', in T.R. Leinbach and C. Capineri (eds), *Globalized Freight Transport*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, pp. 167–188.

Bennathan, E., J. Fraser and L.S. Thompson (1992), 'What determines demand for freight transport', Policy Research Working Papers, Transport, Infrastructure and Urban Development Department, World Bank, Washington, DC.

Christopher, M. (1992), *Logistics and Supply Chain Management*, London: Financial Times Pitman Publishing.

CSCMP (2010), 'Supply chain management terms and glossary', accessed at http://cscmp.org/digital/glossary/glossary.asp.

Dicken, P. (2003), 'Global production networks in Europe and East Asia: the automobile components industries', GPN Working Paper 7, University of Manchester.

Eurostat (2007), 'Average loads, distances and empty running in road freight transport: 2005', *Statistics in Focus*, 117/2007, European Communities, Brussels.

FHWA (2010), *Freight Analysis Framework*, Washington, DC: US Department of Transportation.

Gilbert, R. and K. Nadeau (2002), 'Decoupling economic growth and transport demand: a requirement for sustainability', paper presented at Transportation and Economic Development Conference, organized by US Transportation Research Board, Portland, OR, 5–7 May.

Groothedde, B. (2005), 'Collaborative logistics and transportation networks: a modelling approach to hub network design', Ph.D. thesis, University of Delft.

Groothedde, B., C. Ruijgrok and L. Tavasszy (2005), 'Towards collaborative, intermodal hub networks: a case study in the fast moving consumer goods market', *Transportation Research Part E*, 41 (6), 567–583.

Hesse, M. and J.-P. Rodrigue (2004), 'The transport geography of logistics and freight distribution', *Journal of Transport Geography*, 12 (3), 171–184.

HIDC (1998), *Worldwide Logistics: The Future of Supply Chain Services*, The Hague: Holland International Distribution Council.

Jordans, M., B. Lammers, L.A. Tavasszy and C.J. Ruijgrok (2006), 'The basic potential for inland navigation, rail and short sea shipping' [in Dutch], report TNO 2006-D-R0839.

KiM (2006), *Mobiliteitsbalans*, The Hague: Netherlands Institute for Transport Policy Analysis, Ministry of Infrastructure and Environment.

Kuipers, B., H. van Rooden, A.D.M. van de Ven and A.A.C.M. Wierikx (1995), 'De logistieke wereld bezien door de bril van de overheid', *Tijdschrift voor Vervoerwetenschap*, 1, 55–76.

Kveiborg, O. and M. Fosgerau (2007), 'Decomposing the decoupling of Danish road freight traffic growth and economic growth', *Transport Policy*, 14 (1), 39–48.

Lee, H.L. (2001), 'Aligning supply chain strategies with product uncertainties', *California Management Review*, 44 (3), 105–119.

McKinnon, A.C. (2007), 'Decoupling of road freight transport and economic growth trends in the UK: an exploratory analysis', *Transport Reviews*, 27 (1), 37–64.

Mason, R., C. Lalwani and R. Boughton (2007), 'Combining vertical and horizontal collaboration for transport optimization', *Supply Chain Management*, 12 (3), 187–199.

Rodrigue, J-P. (2006a), 'Transportation and the geographical and functional integration of global production networks', *Growth and Change*, 37 (4), 510–525.

Rodrigue, J.-P. (2006b), 'Challenging the derived transport demand thesis: issues in freight distribution', *Environment and Planning A*, 38 (8), 1449–1462.

Rosstat, Statistics of Russia, Moscow, 2010.

Tapio, P. (2005), 'Towards a theory of decoupling: degrees of decoupling in the EU and the case of road traffic in Finland between 1970 and 2001', *Transport Policy*, 12 (2), 137–151.

Tavasszy, L.A. (2010), 'Goederenvervoer: tijd voor beleid!', Inaugural address, Delft University of Technology, Delft.

Tavasszy, L.A., I. Davydenko and C.J. Ruijgrok (2009), 'The extended generalized cost concept and its application in freight transport and general equilibrium modelling', paper prepared for integration of Spatial Computable General Equilibrium and Transport Modelling, Bilateral Joint Seminar under agreement between NWO and JSPS, Sanjo Hall, University of Tokyo, 19–20 August.

Tavasszy, L.A., K. Ruijgrok and I. Davydenko (2012), 'Incorporating logistics in freight transport demand models: state-of-the-art and research opportunities', *Transport Reviews*, 32 (2), 203–219.

Van Goor, A., M. Ploos van Amstel and W. Ploos van Amstel (2003), *Fysieke distributie: denken in toegevoegde waarde* [Physical distribution: thinking in value added], Leiden: Stenfert Kroese.

Vermunt, J. and F. Binnekade (2000), *European Logistics*, The Hague: Holland International Distribution Council.

Verny, J. (2005), 'The problem of decoupling between freight transport and economic growth', paper presented at the 8th NECTAR conference, Las Palmas, Gran Canaria, 2–4 June.

Vieira, L.F.M. (1992), 'The value of service in freight transportation', Ph.D. thesis, Massachusetts Institute of Technology, Boston, MA.

WBCSD (2004), *Mobility 2030*, Geneva: World Business Council for Sustainable Development.

5

Land use and transport

Bert van Wee

5.1 Introduction

As explained in Chapter 2 people travel because they want to carry out activities such as living, working, shopping and visiting friends and relatives at different locations. Land-use patterns therefore would seem to have a potentially large impact on transport. However, even after several decades of scientific and policy debate, there is still no consensus about the impact of land use on travel behaviour. Some researchers have concluded that land use may have a significant impact on travel behaviour or even that it can be a very fundamental way to influence travel behaviour (e.g. Banister, 1999). Others have found no (or hardly any) impact, and some are sceptical about the impact of land-use planning on travel behaviour (e.g. Martens, 2000).

Note that land use (and infrastructure) has a long-term impact on travel behaviour. Once a residential area is designed and built, it will be very hard to change, for example, densities (e.g. dwellings per square kilometre), or locations of shops and services. So, if 'wrong' choices are made, these will be difficult to 'repair', and the impacts on travel behaviour will be long-lasting. On the other hand, it will take several years before the impacts of land use on travel behaviour are fully materialized. A mismatch between preferences of households and the characteristics of their residential areas may be long-lasting, and may only disappear after households relocate.

The purpose of this chapter is:

1. to discuss the theory explaining the potential impacts of land use on travel behaviour;
2. to give an overview of research findings; and
3. to discuss the pros and cons of the effectiveness of land-use policies.

The chapter is limited to passenger transport, because much more is known about the impact of land use on passenger transport than on goods transport.

Section 5.2 examines the impact of land use on travel behaviour from the perspective of Chapter 2 (especially Figure 2.1). Section 5.3 explains from a theoretical perspective why land use, at least theoretically, affects travel behaviour. Section 5.4 gives examples of studies of land-use impacts on transport. Section 5.5 discusses reasons for the differences in conclusions between the studies. Section 5.6 discusses the evaluation of the impact of land use on travel behaviour in general, and is followed by section 5.7, which goes into the trade-off between environmental versus accessibility benefits of land-use concepts in more depth. Finally, section 5.8 summarizes the most important conclusions.[1]

5.2 A conceptual model for trends in passenger transport – the link with Chapter 2

As expressed in Chapter 2, given the overall population size and demographic characteristics, the total volume of passenger transport and the split between transport modes depend on the locations of human activities, the needs, wants, desires and preferences of people and the transport resistance (generalized transport costs). Locations include activities such as living, working, shopping, recreation and education. The needs and desires of people are related to socio-economic, cultural and personal factors. Income, age, education level and household characteristics are important socio-economic variables (see Chapter 3). Cultural factors, for example, (partly) explain why in some cultures cycling is 'un-cool', whereas in others it is very common. Personal factors relate to attitudes and people's preferences (regardless of variables such as age and income; see Chapter 3). Transport resistance depends on monetary factors, travel times, comfort and the reliability of all the options (see Chapter 6). The top of Figure 2.1 illustrates the relationships between these determinant categories. The current situation reflects a kind of continuously changing equilibrium (or maybe a better term is 'disequilibrium'). This is because new changes occur in advance of the long-term equilibrium actually coming about.

5.3 Why should and how can land use affect travel behaviour?

This section explains why theoretically land-use related determinants can affect travel behaviour. The general theoretical underpinnings are discussed, and this is followed by a summary of the most often mentioned land-use determinants: density, mixed land use, neighbourhood design, and distance of origins and destinations to public transport modes, such as railway stations. Subsection 5.3.3 explores the relationships between land-use variables, other variables and travel behaviour. An overview of the literature on empirical findings related to the theoretical underpinnings is given in section 5.4.

5.3.1 The potential impacts of land use on travel behaviour: the theory of utilitarian travel demand

The theoretical foundation for the impact of land use on travel behaviour can be found in the theory of utilitarian travel demand (see handbooks on transport economics, e.g. Button, 2010; see Chapter 3). This theory postulates that the demand for travel does not derive its utility from the trip itself, but originates rather from the need to reach the locations where activities take place, such as the dwelling, the workplace, and services and facilities. So, from this perspective of utility, travel is seen as 'derived demand'.[2] The demand for travel depends, on the one hand, on the utility of the activity and, on the other, on the (aggregate) costs to reach that destination. These aggregate costs, often labelled as the generalized transport costs (GTC) (the individual's valuation of the time, money and effort needed to cover the distance; see Chapter 6), are determined by characteristics of the transport system and by the spatial structure, such as the distribution of activities over space (see Chapter 2). So according to this theory the locations of activity, in combination with the utility of the activities at those locations and the characteristics of the transport system, have a strong impact on people's travel behaviour. The importance of location of origins and potential destinations for travel behaviour is reflected in the so-called traditional four-step model (see Chapter 15) and was recognized more than half a century ago (Mitchell and Rapkin, 1954).

5.3.2 Key land-use variables and their impact on travel behaviour

Density

Densities refer to the number of opportunities per square kilometre (or acre or any other surface indicator), such as dwellings, households, people and jobs. Gross densities relate to overall available space, net densities to the space that is (or can be) developed (excluding roads, open space, water, etc.). A fundamental question here is: why should density influence travel behaviour? Let us assume two equal land-use scenarios, but with densities in scenario A higher than those in scenario B. Other factors such as average travel speeds for all modes and the type of dwellings, households, jobs, shops and so on are equal. Travel for almost all modes costs money. Other factors being equal, people will prefer to travel less to save money. And travel also costs time. Many studies have shown that people value travel time negatively (see, for example, Wardman, 1998; Gunn, 2001; see Chapter 6). Other factors being constant, people in scenario A, where density is greater, will travel less than in scenario B: people can reach the same locations for their activities while travelling fewer kilometres and so save time and money. In addition, because of the shorter distances more destinations can be reached through slow modes, theoretically resulting in a shift from car to slow modes of transport.

Of course, other factors do not remain constant. For example, owing to the higher densities people can get to more activities in the same time. At least part of the potentially saved travel time can be compensated if people choose a more remote destination: the additional utility for reaching a more remote destination may be greater than the disutility of additional travel (see also section 5.7). Nevertheless, higher densities offer the possibility of travelling less. In order to obtain the benefits from higher densities, housing, offices and other locations should be built in higher densities.

It is important to notice that the spatial scale at which densities occur matters. If, for example, a new residential area of 1000 inhabitants with high densities is built in a region with low densities outside the new residential area, travel behaviour will hardly be affected because almost all destinations will be outside the neighbourhood. And people in that residential area will still have only a few opportunities within reach for a given time (or GTC) budget.

Mixed land use

This factor focuses on the level of mixing of several categories of land use, such as dwellings, workplaces (firms), shops, schools and medical services.

Let us again assume two scenarios for a town. In scenario A all shops, schools and other services are located in the centre of town. In scenario B some of the shops, schools and other services are spread across town throughout all the neighbourhoods. The average distance from all dwellings to the nearest services is much smaller in scenario B than in scenario A. Keeping mode choice constant, the smaller distances will result in fewer passenger and vehicle kilometres. There is also an effect on modal choice, as slow modes will be relatively more attractive because of the shorter distances.

Again, part of the initial effect might be lost. In the centre, for instance, there might be relatively fewer parking places, the distances from these parking places to the shops and services might be greater, and, unlike in the neighbourhoods, there might be paid parking. So, although distances from dwellings to the services are greater, the share of the car might be lower in scenario A than in scenario B. Nevertheless, potentially mixed land use can influence travel behaviour.

Neighbourhood design

Neighbourhood design is related to land use at the lowest scale, starting from the dwellings or buildings and linked to the direct vicinity of the dwellings. Design quality might be important for travel behaviour (Marshall and Banister, 2000; Boddy, 2007). For example, if at the dwelling the place where bicycles are stored is near the road, the bicycles' share might be larger compared to dwellings where the bicycle storage facilities are at the back. If the car can be parked near the dwelling, on one's own property or in a public parking place near the dwelling, car use will be more attractive than where there are central parking places further away from dwellings. An attractive environment might stimulate people to walk or cycle. This attractiveness relates to allocation of space for several land-use categories, the architecture of the built environment and the presence of features such as parks and trees. Road infrastructure design may also be of relevance. For example, if there are attractive pavements and cycling lanes the share of slow modes may be larger than if there are no or less attractive facilities for slow modes.

Distance to public transport connections

The distance to railway stations may have an impact on modal choice as (differences in) travel times by car and train are very important. For public transport, total travel time contains both the 'in vehicle' (in train) time and the access and egress time (e.g. from home to the station by bus or bike or walking, and from the station to the office by walking or taking local public transport). If more dwellings and opportunities (e.g. jobs) are located near railway stations, access and egress times are shorter for a larger number of trips, resulting in a higher share for the train as compared to the car. In this way, distance and connections to railway stations may result in a higher share for public transport.

Interactions between determinants

The determinants can interact. For example, the effect of building in higher densities may have an additional impact on mode choice if the areas are located near railway stations. Then there is an additional effect on modal choice which was not mentioned in the section above on densities. And if the area around a railway station is attractive it can encourage more people to take the train, increasing the modal shift potential of building in high densities.

5.3.3 Relationships between land-use variables, other variables and travel behaviour

Some studies have been published that simply investigated correlations between land-use variables and travel behaviour variables (see Figure 5.1). The best known of these was the study by Newman and Kenworthy (1988) on the impact of urban densities on energy use for travel per person (see section 5.5).

Since the late 1980s, studies into the impact of land use on travel behaviour have included socio-economic and demographic variables, as visualized in Figure 5.2 (see Chapter 3 for the general impact of socio-economic and demographic variables). Figure 5.2 shows that travel behaviour is influenced not only by land-use variables but also by socio-economic and demographic

Figure 5.1 Relationships between land-use variables and demographic variables and travel behaviour – the traditional approach

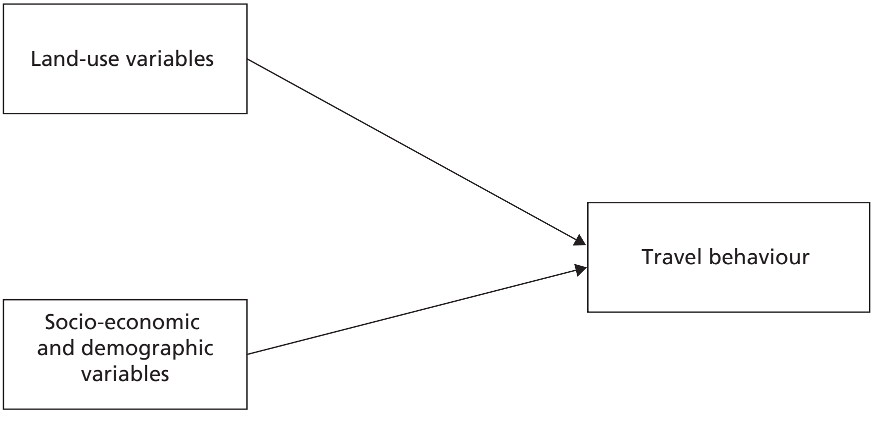


Figure 5.2 Relationships between land-use variables, socio-economic and demographic variables, and travel behaviour

variables, such as income, age, education level and household variables. The studies find lower impacts of land-use variables on travel behaviour variables, as part of the impact is accounted for by the socio-economic and demographic variables. For example, people with higher incomes can afford to live in relatively expensive houses in lower-density residential areas. The impact of densities on travel behaviour would be overestimated if the difference in income levels was ignored.

Most recently, three more (often interrelated) changes have been made that are not included in Figure 5.2. Firstly, researchers have added more people and household variables not covered by socio-economic and demographic variables. These variables relate to attitudes, lifestyles and preferences for modes. Such variables express the possibility that people with the same socio-economic and demographic variables can still differ. Such attitudes, lifestyles and preferences relate to the needs of people, as explained in Chapter 3. For example, some people may have a more pro-environmental attitude and lifestyle, or a more culture-oriented lifestyle as opposed to a more material lifestyle, or people might simply prefer, more than others, to travel by car ('car lovers'), public transport ('public transport lovers') or in some countries by slow modes (walk, bike). Studies that include such variables generally find a decrease in the impact of land-use variables compared to studies that do not include such variables. The second change is that researchers have recognized that people self-select themselves to locate in specific residential areas. Mokhtarian and Cao (2008), based on Litman (2005), state that self-selection refers to 'the tendency of people to choose locations based on their travel abilities, needs and preferences'. Self-selection can then result from 'traditional' variables such as income but also from attitudes and lifestyles or preferences for modes. Thirdly, and partly related to the phenomenon of residential self-selection, researchers have changed the model structure of variable categories by explicitly adding that socio-economic and demographic variables, and variables related to attitudes, lifestyles and preferences for modes have an impact on residential choice. The more complex relationships between the categories of variables that result can be estimated by so called structural equation models (SEM models). Figure 5.3 conceptualizes the three changes that result from including additional variables and the more complex model structure.

5.4 The impact of land use on transport – a short overview of the literature

The aim of this section is not to give an extensive review but to summarize some empirical examples of the theory as discussed in section 5.3. We

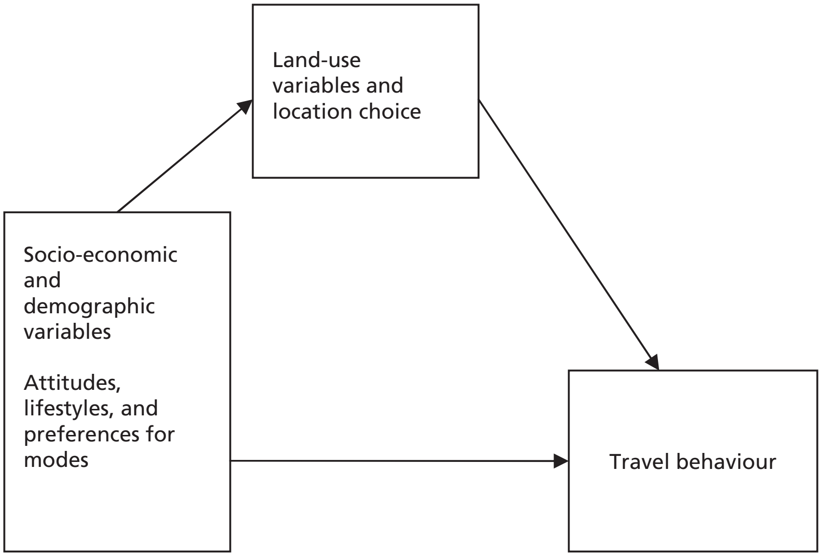


Figure 5.3 Relationships between land-use variables and location choice, socio-economic and demographic variables, attitude variables and travel behaviour – the current approach

first give examples of studies into the impact of densities, mixed use, neighbourhood design and distance to public transport connections, followed by studies that add attitude, lifestyle and preference of modes variables and the use of SEM models.

5.4.1 Densities

One of the most frequently cited studies on the impact of densities is the study by Newman and Kenworthy (1988), who concluded that energy use for travel per inhabitant is negatively related to urban density. The study has been criticized frequently. An important reason is that they did not correct for differences in incomes and fuel prices or transport costs. In other words, their research was based on the too simple variables structure of Figure 5.1 instead of the structure of Figure 5.2. Inspired by the critics, Newman and Kenworthy (1999; see also Kenworthy and Laube, 1999) updated their study, correcting for the most important intervening factors. They still found a strong relationship between densities and energy use for transport, but the impact of densities was only about half that of their original study. These findings are in line with those of Gordon (1997), who concludes that about one-third of variation in per capita transport energy consumption is attributable to land-use characteristics. Dargay and Hanly (2004) conclude that land-use characteristics (population density, settlement size, local access to shopping and other facilities and accessibility of public transport) play a significant role in car ownership and use of the car. Density has a greater impact than settlement size, and proximity to local facilities encourages walking instead of car travel. Other references supporting the positive impact of urban density on travel behaviour are Cervero (1996), Cervero and Kockelman (1997), Badoe and Miller (2000) and Giuliano and Dargay (2006). However, some researchers doubt the impact of densities on travel behaviour (see, for example, Gordon and Richardson, 1997; Snellen, 2001).

5.4.2 Mixed use

The impact of mixed use on travel behaviour is often studied together with other land-use characteristics. Stead (2001) carried out research into the impacts of land use on travel behaviour. He concluded that socio-economic characteristics explain about half the variation in travel distance per person across different wards in the UK, whereas land-use characteristics often explain up to one-third of the variation in travel distance per person. He concluded that the results do indicate that land-use characteristics, such as mixed-use development, settlement size and the provision of local facilities, have a role to play in promoting more sustainable development. Anderson et al. (1996) reviewed many studies on this subject and concluded that urban land-use patterns have a significant impact on travel behaviour in general and on car use more specifically. Compact building and mixed land use are of particular importance. Other references concluding that mixed land use has a positive impact on travel behaviour are Frank and Pivo (1994) and Cervero (1996).

5.4.3 Neighbourhood design

Manaugh and El-Geneidy (2011) show for Montreal that, if walkability of a neighbourhood is high, people make significantly more walking trips for most non-work purposes. Pan et al. (2009) found that, in the pedestrian and cyclist-friendly neighbourhoods in Shanghai, residents travel shorter distances than in other neighbourhoods, and that pedestrian- and cyclist-friendly urban form makes the non-motorized modes feasible options.

5.4.4 Distance to public transport connections

Distances to public transport connections relate to both the origin and the destination of trips. An important category of destinations is work locations, including offices. Brons et al. (2009) show that access to railway stations contributes to travel by train. They conclude that expanding access services to the railway station can increase the mode share of rail. Cervero and Duncan (2006: 53) emphasize the importance of both ends of a trip: concentrating 'housing near rail stops will do little to lure commuters to trains and buses unless the other end of the trip – the workplace – is similarly convenient to and conducive to using transit'.

5.4.5 Attitudes, lifestyles and preferences for modes, residential self-selection, and SEM models

One of the first studies that included attitude variables was the study by Kitamura et al. (1997). They studied travel behaviour of people living in five neighbourhoods in the San Francisco Bay area, including socio-economic and demographic variables, land-use variables, and 39 attitudinal variables. They found that attitudinal variables explained the highest proportion of the variation in the data. Examples of studies that include residential self-selection are Bagley and Mokhtarian (2002), Cervero and Duncan (2002), Cao et al. (2006) and Næss (2009). Note that self-selection is not limited to residential self-selection but may also occur with respect to work location choice or the choice of other destinations (van Wee, 2009). Studies that include land-use variables and socio-economic and demographic variables, together with variables related to attitudes, lifestyles and preferences for modes, residential self-selection and travel behaviour, generally use SEM models to disentangle the relationships between variables. Examples of such studies include Van Acker et al. (2007) and Bohte et al. (2009). See Cao et al. (2009) for an overview of empirical studies in the area of residential self-selection; see Mokhtarian and Cao (2008) for an overview of methodologies of studies in the area of residential self-selection.

An overall conclusion of such studies is that travel behaviour results from a complex interplay of all these variables, as conceptualized by Figure 5.3. The impact of land-use variables on travel behaviour can easily be overestimated if such complex relationships are overlooked. An important question then becomes what the impact of overestimation would be. Should the conclusion be that land-use policies that encourage people to travel by public transport or slow modes or over shorter distances do not make sense, or at least make less sense than often assumed? Some people do not agree with this conclusion. For example, Næss (2009) argues that, even if attitudes of people explain residential choice as well as travel preferences, people will only act according to their preferences if their residential area allows them to do so. Secondly, he argues that car ownership, and to some extent also transport attitudes, is influenced by the characteristics of the residential locations. A third reason is provided by Schwanen and Mokhtarian (2005), namely that some people face a lack of congruence between physical neighbourhood structure and their preferences. One could argue that, for quality of life reasons, reducing such a lack of congruence matters.

5.5 Why are the conclusions different?

Why do researchers find different results with respect to the possible impact of land use on transport? There are several reasons for this and some of them will be presented in this section.

5.5.1 The research method

A major source of differences as found between studies relates to research methods. Section 5.4 has already explained the importance of including other variables that have an impact on travel behaviour, including the importance of phenomena such as residential self-selection and ensuring an adequate structure between variables. Generally speaking, Handy (1996) concludes that more advanced research methods have generally found smaller and less significant effects of land use on travel behaviour. The studies on the impact of density by Newman and Kenworthy (1988) provide an obvious example: their initial study ignored intervening factors such as prices and incomes. After the inclusion of those factors the impact of densities on energy use per person halved.

Next we discuss more reasons why study results differ, some of them also being methodological reasons.

5.5.2 The level of difference in crucial factors

One of the causes for different results from research is that the study areas sometimes do not really differ with respect to the most important land-use factors that influence transport, such as densities and the level of mixed land use. Small differences between areas with respect to important land-use factors of course result in small differences in travel behaviour.

5.5.3 The geographical scale

An often neglected aspect in the discussion is the geographical scale. First, the definitions or indicators for densities differ between studies (see subsection 5.3.2, the remark on gross versus net densities). Secondly, the size of the area at which forms of mixed use or densities occur matters (see subsection 5.3.2). Thirdly, several scales can be distinguished, including the direct surroundings of the dwellings, the neighbourhood, the town or city, the region, the part of a country, the whole country and even the international scale. The scale of research may affect the results. For example, Newman and Kenworthy (1988, 1999) focus on large cities. In several countries such as Denmark or Belgium, such cities do not exist. Therefore it is not possible to conclude beforehand that building in high-density locations in such countries will result in lower energy use for transportation equivalent to the results found by Newman and Kenworthy.

5.5.4 The time horizon

Another reason may be the time horizon included. It is generally recognized in academic literature that, after changes in determinants for travel behaviour, people will not change their behaviour instantaneously. This comment relates not only to changes in land use but also to changes in infrastructure and prices. The relevance of how much time has passed since changes took place and when the empirical research is undertaken has implications for the findings. Suppose a new railway station is built in or near a residential area or office park. Many people might not immediately change their travel behaviour (in this example, their mode choice) but people might move from the residential area and new households that move to the dwellings that become available might be more inclined to travel by train. As explained above, people self-select themselves into residential areas based on their attitudes toward travel and preferences for modes. As a result the impact of the new railway station will probably increase over time.

5.5.5 Differences between countries

Differences between countries (and even, within countries, regions) complicate the transferability of results. Firstly, concepts play an important role in the discussion on the impact of land use on transport. Let us assume the compact city. What is considered to be a low-density residential area in many European regions and countries may be an example of compact building in the USA. In other words, what people consider as compact will differ between regions and countries.

There are other reasons that make the translation of land-use effects from one country to another risky. Examples relate to the role of different modes in the transport system but also to cultural differences. In Denmark and the Netherlands the bicycle plays an important role in the transport system. For short distances, the bike may compete with the car or local public transport. But in many other countries the role of the bike is limited or absent. The role of the public transport system also varies from country to country. Such differences may result in other effects of the same land-use concepts. For example, as rail transport is a minor form of passenger travel in the US, results from empirical research there on the impacts of distances to and from stations would have little relevance to Europe, where levels of rail use are much higher, and vice versa.

Cultural differences may also be of importance, not only between, but even within, countries. As already explained in section 5.2, in some cultures cycling is not at all 'sexy', whereas in others such cultural barriers do not exist.

5.5.6 Indirect effects

Figure 2.1 shows that land use may have a direct impact on passenger travel, but it also shows the existence of several indirect effects. These effects may be very complex – see all the arrows in Figure 2.1. It is very complicated to distinguish all the kinds of direct and indirect effects quantitatively.

5.5.7 The impact of policy

Policy can also affect land-use patterns. The travel behaviour of people living in a high-density, public-transport-oriented city such as Tokyo may differ strongly from the travel behaviour of people living in a low-density, car-oriented city such as Los Angeles. But this does not mean that it will be easy to implement Tokyo's urbanization patterns and infrastructure systems in the USA. So, even if land use has an impact on travel behaviour, it does not mean that it is easy to use land-use *policy* as an instrument to influence travel behaviour. The effect of land-use policies may therefore be limited compared to the effect of land use in general. Generally speaking, researchers who assume relatively strong impacts of policy on land use are more optimistic about possibilities to influence travel behaviour by land-use planning than those who have not made this assumption (Anderson et al., 1996).

To summarize, there are many reasons for different researchers and policy-makers to draw different conclusions with respect to the impact of land use on travel behaviour. Probably methodological reasons dominate. A systematic analysis of possible causes for differences will be needed to discover these reasons.

5.6 Evaluating the impact of land use on travel behaviour: indicators and evaluation methods

Several of the more advanced studies have found significant impacts of land-use variables on travel behaviour, even though there is no consensus on this impact (Cao et al., 2009). Here it is assumed that there is enough empirical evidence to conclude that land-use characteristics as discussed above have an impact on travel behaviour. And several options for policy-making exist, such as building in higher densities, mixing land use, building in high densities near railway stations, and design options for walking- and cycling-friendly neighbourhoods. Would the conclusion then be that policy-makers should be advised to base land-use planning only on impacts on travel behaviour, or even only on impacts on car use? Probably not, as choices with respect to land-use policies should be made considering all relevant aspects, including travel behaviour but also impacts. This section will first discuss these impacts, followed by a discussion on how to evaluate land-use alternatives using these criteria.

5.6.1 Indicators

Most research (and policy documents) on the impact of land use on travel behaviour places the topic in an environmental context: land use could improve the environment by reducing car use and its negative impacts. Research generally uses the following indicators to express this impact: (1) kilometres (vehicles, passengers), mostly by mode, often by trip purpose; and (2) the number of trips, mostly by mode, often by trip purpose. Furthermore, some studies also focus on trip distances. Only a small minority of studies provide environmental indicators, such as CO_2 and NO_x emissions. These indicators are relevant but very often are not produced on a comprehensive scale. Additional (categories of) indicators are now discussed. It should be noted that some of the suggestions are related or even overlap. The first three suggestions are related to the positive aspects of travel, aspects that are very often neglected in transport discussions in general and in discussions on the impact of land use on transport.

Accessibility

Many studies do not use indicators that express the quality of the land-use and infrastructure system: to what extent does this system enable us to travel between locations we want to visit and thereby participate in the activities desired? Therefore additional accessibility indicators could be used, including those used in geography (potential accessibility, time–space related accessibility indicators), utility-based indicators (see Geurs and van Wee, 2004, for a literature review of accessibility indicators) and travel time indicators (e.g. Schwanen et al., 2002). See section 5.7 for further discussion.

The option value

Current evaluations focus on user benefits only, but, if non-user benefits are relevant, then the option value may be important. The option value in the context of land use and transport can be described as an individual's valuation of the opportunity to be able to use a particular transport mode or piece of infrastructure in the future that is not being used in the present, or the option to have access to a specific destination that is currently not visited. For example, car-owners may value the ability to use a public transport service when they cannot make use of the car, for example owing to unavailability or a breakdown, bad weather, increases in fuel prices or other car costs, or the loss of the ability to operate a car. Or a person may value having access to shops she does not currently visit (see Geurs et al., 2006, for an empirical study of option values).

The consumer surplus

The consumer surplus plays an important role in evaluations (including cost–benefit analyses – CBAs) of infrastructure projects. The consumer surplus is the difference between the market price of a product or service and the value for a consumer. For example, if a consumer is willing to pay 50 euros for a book but the price is 20 euros, the consumer's surplus is 30 euros. Some trips probably have a very high consumer surplus. For example, if one wants to visit a relative in hospital, the visit may be worth much more than the costs of travel. Let us assume two scenarios: scenario A with the current pattern of hospitals and scenario B with many fewer hospitals (e.g. in order to profit from scale effects). In scenario A, visiting a relative in the hospital is possible after a 10-minute cycling trip, whereas in scenario B a car trip of half an hour is needed. Assuming the visiting frequency is the same in both scenarios, the consumer surplus of the A is much larger than that of scenario B. The difference is the GTC of the car trip in scenario B minus the GTC of the bicycle trip in scenario A (costs include both monetary and non-monetary costs such as travel time).

Safety

Safety impacts of land-use and transport alternatives may differ. If they do, such impacts should be included in the ex ante evaluation of these alternatives.

Health impacts due to exercise

If people travelled more by slow modes, this would have not only environmental and safety advantages but also health advantages (e.g. Saelensminde, 2004; Frank et al., 2007).

Environmental impacts

If environmental impacts are included, it is usually done by listing emissions levels. However, for local air pollution, exposure is very important. The same amount of kilograms of emissions may have different exposure impacts. It is highly relevant if pollutants are emitted on a road with many people (living, working, carrying out recreational activities) in its vicinity as opposed to a road in an agricultural area (see Chapter 10). Noise effects are also highly dependent on the direct vicinity of a road or railway.

Valuation by the people

Much is known about residential choice preferences (see, for example, Dieleman, 2001). Many people prefer living in spacious homes on spacious plots. On the other hand, building in low densities results in less accessibility to opportunities and to the public transport system. Building in low densities also results in a larger space claim on residential areas and thus in less green space between cities and towns. How do people value such items? We hardly know. But people's opinions are relevant for an overall view of the pros and cons of land-use scenarios.

We probably know even less about what people think of the job location. What do people prefer – a job location on the edge of town, near a motorway, or in the inner city near a railway station? Probably different groups of people have different preferences. Such valuations of people could be relevant for the ex ante evaluation of land-use concepts.

Financial aspects

The relevant factors here include the costs of construction, maintenance and exploitation of land-use and transport alternatives. We know more about the financial aspects of the transport system, in general, and more about infrastructure costs, in particular, than we know about costs of land-use alternatives. The impact on GDP (and unemployment) is also relevant. The (valuation of) indirect effects of land-use and transport alternatives (such as effects on the labour and housing market) are much more difficult to estimate than the direct effects (see SACTRA, 1999, or Banister and Berechman, 2000, for a discussion of the indirect effects of infrastructure). A distinction should be made between costs for society as a whole, for the government and for the users or consumers.

Robustness

Another issue is the robustness of the land-use and transport system. In other words, how vulnerable are we to, for example, an expected or unexpected limitation on energy availability for transport? Energy limitations may be the result of political instability in oil producing countries, much higher oil prices (for example, due to 'peak oil'; see Chapter 10) or stringent environmental (climate) policies. Preferences of consumers and firms might also differ in the future. In addition, what will happen if sustainably produced energy becomes available at reasonable prices? The question then will change from 'How can land use contribute to reducing transport problems?' to 'How can land use enable us to perform activities in different places under different conditions?' This changing role is important not only for land use but also for the role of public transport and slow modes, and for ICT. Probably land-use transport strategies that are positively valued with respect to travel behaviour impacts will be robust. Such strategies include compact building, mixed land use and availability of good-quality public transport.

5.6.2 Evaluation methods

The use of a multi-criteria analysis (MCA) or a CBA for land-use scenarios is generally recommended (see Chapter 14). A lot is known about the valuations of reductions in travel time and benefits of additional travel. Much is also known about people's willingness to pay for noise reductions or risk reductions. However, choosing price tags (CBA) or weight factors (MCA) for several other output indicators is partly a political choice. This includes aspects like the value of CO_2 emissions reduction, nature conservation and 'spatial quality'. Researchers can assist policy-makers who have to weight components, and they can advise on methods, carry out research and use the results of other studies. For this reason, combining a CBA and an MCA provides an interesting basis for policy-making. For a further discussion on evaluation methods, see Chapter 14.

5.7 Evaluating the impact of land use on travel behaviour: the environment versus accessibility

As explained above, research into the area of land use and transport has generally placed itself in the context of environmental gains: land use could contribute to lower levels of car use, and an increase in the use of slow modes and public transport. However, as explained above, several researchers have found only limited impacts of land use on travel behaviour. Some of them have concluded that related policies therefore make no sense. In line with insights from economics, it can be argued that land use potentially could decrease motorized mobility (including car use), but in practice people may not travel less using motorized modes. Then there must be accessibility gains that at least equal the potential gains in savings of GTC. Therefore the general way of evaluating land-use concepts on (only) travel behaviour and environmental gains is insufficient, and accessibility benefits should be added. Van Wee (2011) expands on this proposition: 'If the potential (theoretically possible) impact of land use on travel behaviour does not occur in practice, there must be accessibility benefits for travellers, that they value at least as highly as the benefits of the potential decreases in Generalized Transport Costs.'

I define accessibility benefits as all the benefits that provide utility to travellers, related to the activities they carry out at different locations (working, shopping, visiting friends, etc.). For the *net* benefits these benefits should be corrected for the disutility of travel (generalized transport costs; see section 5.3 and Chapter 6).

For economists and others with a background in utility-based discrete choice theory, the proposition is not at all surprising, as it formulates the economic approach to accessibility.

Below, the proposition is explained, using a simple example. Suppose we could 'shrink' a certain region to 25 per cent of its original size. I assume a closed region (no external trips). In addition I assume that all other determinants remain constant, in particular the locations of activities. In this case, trip distances would be reduced by 50 per cent (surface area expands quadratically with an increase in radius). I assume all trips are made by car. If there were no behavioural changes other than distance reductions, people could participate in the same activities by travelling only half of the kilometres and would need only half of the original travel time. More generally, GTC would be reduced by 50 per cent. But, according to the theory of constant travel time budgets (TTB), because people trade off the benefits of activities and the disutility (GTC) of travel, it can be expected that people will choose more remote destinations. They could, for example visit another, more remote supermarket because it is cheaper or offers more products. Or they could choose another job at a greater distance from their home because it pays better, is more challenging or offers more career prospects. Note that, according to the theory of constant TTB, not all people would change their behaviour, and certainly not everyone would travel as many kilometres as before the 'shrink'; the TTB is about averages over a large group. Some people might not change jobs at all, but others might choose a job at a distance four times the commuting distance after the shrink. Now let us assume that at the level of all people who travel in the region the potential decrease in travel times and distances completely disappears because of behavioural changes. In that case no effects of land use would be found even though there must be accessibility benefits with a value that at least equals the potential savings of GTC. If not, people would take advantage of the decrease of GTC by 50 per cent. It is also likely that there will be greater benefits from choosing the more remote destinations. Again, please note that this insight is not new for economists; it applies to all goods and services with elastic demand.

Of course reality is more complex than our example. People can adapt their behaviour in more respects, such as a mode change or a change in trip frequencies. A shrink, as in our example, might lead to a decrease in average travel speeds due to higher densities on the road network and parking capacity problems. These complexities change the potential reduction of GTC due to the 'shrink' but do not change the principle that the behavioural changes after the shrink, other than simply reducing travel distance and time, must have benefits that at least equal the value of the benefits of the potential reduction of GTC.

The same line of reasoning applies for mixed use or a reduction in distances to public transport nodal points: thanks to the land-use changes a potential reduction in GTC is possible owing to distance reduction and/or mode change, and that is valued positively. If, in practice, people did not travel less (but travelled to destinations further away) there must be accessibility benefits that the travellers value at least as much as the benefits that would be possible from the reduction of GTC.

To conclude, if research showed that land-use concepts that allow for a reduction in GTC would not result in people really reducing their GTC, it is not correct to reject such concepts because then significant accessibility effects occur.

5.8 Conclusions and discussion

These are the most important conclusions of this chapter:

1. The theoretical foundation for the impact of land use on travel behaviour can be found in the theory of utilitarian travel demand. This theory postulates that the demand for travel does not derive its utility from the trip itself, but rather it originates from the need to reach the locations of activities.
2. The land-use related determinants that have an impact on travel behaviour that are most often mentioned in the literature are density, mixed land use, neighbourhood design, and distance of origins and destinations to public transport nodes such as railway stations.
3. Travel behaviour results from a complex interplay of land-use variables, socio-economic and demographic variables, variables related to attitudes, lifestyles and preferences for modes, and residential self-selection. The impact of land-use variables on travel behaviour can easily be overestimated if such complex relationships are overlooked.
4. Empirical studies into the impact of land use on travel behaviour often find different and even contradicting results, methodological reasons probably being the most important cause of these differences.
5. Land-use concepts are often promoted because of policy reasons, environmental reasons being the most mentioned. However, most studies focus only on impacts of land use on travel behaviour. To make such studies more relevant for policy, additional output variables are relevant. These include environmental impacts, accessibility impacts, the option value, the consumers' surplus, safety and health impacts, the valuation of land-use concepts by people, financial aspects, and long-term robustness of the land-use and transport system.
6. If most effects of land-use concepts can be quantified and expressed in monetary terms, a CBA could be an attractive method to evaluate such concepts ex ante. If important effects are difficult to quantify or express in monetary terms, combining a CBA and an MCA provides an interesting basis for such evaluation.
7. If the potential (theoretically possible) impact of land use on travel behaviour does not occur in practice, there must be accessibility benefits for travellers that they value at least as highly as those of potential decreases in generalized transport costs.

NOTES

1 Sections 5.3 to 5.6 are partly based on van Wee (2002). Section 5.7 is based on van Wee (2011).

2 Note that not all travel is derived (see Mokhtarian and Salomon, 2001, and Chapters 2 and 3).

REFERENCES

Anderson, W.P., P.S. Kanaroglou and E.J. Miller (1996), 'Urban form, energy and the environment: a review of issues, evidence and policy', *Urban Studies*, 33 (1), 7–35.

Badoe, D.A. and E.J. Miller (2000), 'Transportation–land-use interaction: empirical findings in North America, and their implications for modeling', *Transportation Research Part D*, 5, 235–263.

Bagley, M.N. and P.L. Mokhtarian (2002), 'The impact of residential neighborhood type on travel behavior: a structural equations modeling approach', *Annals of Regional Science*, 36, 279–297.

Banister, D. (1999), 'Planning more to travel less: land use and transport', *Town Planning Review*, 70 (3), 313–338.

Banister, D. and J. Berechman (2000), *Transport Investment and Economic Development*, London: University College London Press.

Boddy, M. (2007), 'Designer neighbourhoods: new-build residential development in nonmetropolitan UK cities – the case of Bristol', *Environment and Planning A*, 39 (1), 86–105.

Bohte, W., K. Maat and B. van Wee (2009), 'Measuring attitudes in research on residential self-selection and travel behaviour: a review of theories and empirical research', *Transport Review*, 29 (3), 325–357.

Brons, M., M. Givoni and P. Rietveld (2009), 'Access to railway stations and its potential in increasing rail use', *Transportation Research Part A*, 43 (2), 136–149.

Button, K.J. (2010), *Transport Economics*, 3rd edition, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Cao, X., S. Handy and P.L. Mokhtarian (2006), 'The influences of the built environment and residential self-selection on pedestrian behavior', *Transportation*, 33 (1), 1–20.

Cao, X., P.L. Mokhtarian and S. Handy (2009), 'Examining the impact of residential self-selection on travel behavior: a focus on empirical findings', *Transport Reviews*, 29 (3), 359–395.

Cervero, R. (1996), 'Mixed land-uses and commuting: evidence from the American housing survey', *Transportation Research A*, 30 (5), 361–377.

Cervero, R. and M. Duncan (2002), *Residential Self-Selection and Rail Commuting: A Nested Logit Analysis*, Berkeley: Transportation Center, University of California.

Cervero, R. and M. Duncan (2006), 'Which reduces vehicle travel more: jobs–housing balance or retail–housing mixing?', *Journal of the American Planning Association*, 74 (4), 475–490.

Cervero, R. and K. Kockelman (1997), 'Travel demand and the 3Ds: density, diversity, and design', *Transportation Research D*, 2 (3), 199–219.

Dargay, J. and M. Hanly (2004), 'Land use and mobility', paper presented at the World Conference on Transport Research, Istanbul, July.

Dieleman, F.M. (2001), 'Modelling residential mobility: a review of recent trends in research', *Journal of Housing and the Built Environment*, 16, 249–265.

Frank, L. and G. Pivo (1994), 'Impacts of mixed use and density on utilization of three modes of travel: single occupant vehicle, transit, and walking', *Transportation Research Record*, 1466, 44–52.

Frank, L.D., B.E. Saelens, K.E. Powell and J.E. Chapman (2007), 'Stepping towards causation: do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity?', *Social Science and Medicine*, 65 (9), 1898–1914.

Geurs, K., R. Haaijer and B. van Wee (2006), 'Option value of public transport: methodology for measurement and case study for regional rail links in the Netherlands', *Transport Reviews*, 26 (5), 613–643.

Geurs, K.T. and B. van Wee (2004), 'Accessibility evaluation of land-use and transport strategies: review and research directions', *Transport Geography*, 12, 127–140.

Giuliano, G. and J. Dargay (2006), 'Car ownership, travel and land use: a comparison of the US and Great Britain', *Transportation Research Part A*, 40 (2), 106–124.

Gordon, I. (1997), 'Densities, urban form and travel behaviour', *Town and Country Planning*, 66 (9), 239–241.

Gordon, P. and H. Richardson (1997), 'Are compact cities a desirable planning goal?', *Journal of the American Planning Association*, 63 (1), 95–106.

Gunn, H. (2001), 'Spatial and temporal transferability of relationships between travel demand, trip cost and travel time', *Transportation Research E*, 37, 163–189.

Handy, S. (1996), 'Methodologies for exploring the link between urban form and travel behaviour', *Transportation Research D*, 1 (2), 151–165.

Kenworthy, J. and F. Laube (1999), 'A global review of energy use in urban transport systems and its implications for urban transport and land-use policy', *Transportation Quarterly*, 53 (4), 23–48.

Kitamura, R., P.L. Mokhtarian and L. Laidet (1997), 'A micro-analysis of land use and travel in five neighbourhoods in the San Francisco Bay area', *Transportation*, 24, 125–158.

Litman, T.A. (2005), 'Land use impacts on transport: how land use factors affect travel behavior', accessed 13 January 2006 at http://www.vtpi.org/landtravel.pdf.

Manaugh, K. and A. El-Geneidy (2011), 'Validating walkability indices: how do different households respond to the walkability of their neighbourhood?', *Transportation Research Part D*, 16, 309–315.

Marshall, S. and D. Banister (2000), 'Travel reduction strategies: intentions and outcomes', *Transportation Research Part A: Policy and Practice*, 34 (5), 321–338.

Martens, K. (2000), *Debatteren over mobiliteit: over de rationaliteit van het ruimtelijk mobiliteitsbeleid* [Debating mobility: on the rationality of transport related land-use policy], Nijmegen: Catholic University of Nijmegen.

Mitchell, R.B. and C. Rapkin (1954), *Urban Traffic: A Function of Land Use*, New York: Columbia University Press.

Mokhtarian, P.L. and X. Cao (2008), 'Examining the impacts of residential self-selection on travel behavior: a focus on methodologies', *Transportation Research Part B*, 42 (3), 204–228.

Mokhtarian, P.L. and I. Salomon (2001), 'How derived is the demand for travel? Some conceptual and measurement considerations', *Transportation Research Part A*, 35 (8), 695–719.

Næss, P. (2009), 'Residential self-selection and appropriate control variables in land use: travel studies', *Transport Reviews*, 29 (3), 293–324.

Newman, P.W.G. and J.R. Kenworthy (1988), 'The transport energy trade-off: fuel-efficient traffic versus fuel-efficient cities', *Transportation Research A*, 22A (3), 163–174.

Newman, P. and J. Kenworthy (1999), *Sustainability and Cities: Overcoming Automobile Dependence*, Washington, DC: Island Press.

Pan, H., Q. Shen and M. Zhang (2009), 'Influence of urban form on travel behaviour in four neighbourhoods of Shanghai', *Urban Studies*, 46 (2), 275–294.

SACTRA (Standing Advisory Committee on Trunk Road Assessment) (1999), *Transport and the Economy*, London: Department of the Environment, Transport and the Regions.

Saelensminde, K. (2004), 'Cost–benefit analyses of walking and cycling track networks taking into account insecurity, health effects and external costs of motorized traffic', *Transportation Research Part A: Policy and Practice*, 38 (8), 593–606.

Schwanen, T., M. Dijst and F.M. Dieleman (2002), 'A microlevel analysis of residential context and travel time', *Environment and Planning A*, 34 (8), 1487–1507.

Schwanen, T. and P.L. Mokhtarian (2005), 'What if you live in the wrong neighborhood? The impact of residential neighborhood type dissonance on distance travelled', *Transportation Research Part D*, 10 (2), 127–151.

Snellen, D. (2001), *Urban Form and Activity-Travel Patterns: An Activity-Based Approach to Travel in a Spatial Context*, Eindhoven: University of Technology.

Stead, D. (2001), 'Relationships between land use, socioeconomic factors, and travel patterns in Britain', *Environment and Planning B*, 28 (4), 499–529.

Van Acker, V., F. Witlox and B. van Wee (2007), 'The effects of the land use system on travel behavior: a structural equation modeling approach', *Transportation Planning and Technology*, 30 (4), 331–354.

Van Wee, B. (2002), 'Land use and transport: research and policy challenges', *Transport Geography*, 10, 259–271.

Van Wee, B. (2009), 'Self-selection: a key to a better understanding of location choices, travel behaviour and transport externalities?', *Transport Reviews*, 29 (3), 279–292.

Van, Wee, B. (2011), 'Evaluating the impact of land use on travel behaviour: the environment versus accessibility', *Journal of Transport Geography*, 19 (6), 1530–1533.

Wardman, M. (1998), 'The value of travel time: a review of British evidence', *Journal of Transport Economics and Policy*, 32 (3), 285–316.

6

Transport resistance factors: time, money and effort

Jan Anne Annema

6.1 Introduction

As explained in Chapter 2, passenger and freight transport volume is determined by the locations of activities, needs and resistance factors. This chapter aims to describe the transport resistance factors and their impact on passenger and freight transport demand.

From a physical perspective, resistance is everything that stops or obstructs a force. In this image the locations of activities and the need for trips are 'forces', resulting in transport. Resistance factors such as travel time, money costs and effort obstruct these 'forces'. The lower the resistance factors, the higher the amount of transport. For, with low resistance factors, even the not so very important trips will still be made. The opposite is true also: the higher the resistance factors, the lower the amount of transport.

Economists use other jargon but the same basic idea as used in the physical analogy to explain transport. In the economic utility theory the idea is that a trip results in benefits (see also Chapter 2), for if there are no benefits the trip would not be made. At the same time the trip comes with costs – the resistance factors. The trip takes time, perhaps a fare or petrol has to be paid and perhaps the trip has to be made in a very busy train compartment that is too warm. The utility theory states that only if the benefits outweigh the costs will the trip be made. Thus, in countries or regions with poor road infrastructure and consequently high costs, relatively few long-distance road trips – only the highly beneficial ones – will be made. Here, it is important to realize that resistance factors influence not only the amount of transport but also modal choice. For example, if carriers succeed in improving rail freight services (e.g. lowering costs, increasing frequencies) some freight will probably be shifted from road to rail.

It is clear from the previous paragraphs that we define transport resistance in this chapter in broad terms. Resistance is related not only to travel time and monetary costs but also to the more vague but sometimes important concepts of 'effort' in passenger transport and 'transport services' in freight transport. Effort and transport services are terms for a broad class of factors influencing the decision to make a trip, such as discomfort, worries about reliability and so on. Economists sometimes use the term 'generalized travel or transport costs'. With this term they mean the whole of transport resistance factors. In most cases they add up all the different transport resistance factors into one generalized travel cost unit, mostly a monetary unit (dollar, pound sterling, euro, yuan), sometimes a time unit.

Transport experts often speak of demand and supply factors to explain transport volumes. In these terms forces and resistance factors can also be recognized. For example, a demand factor for freight transport is the amount of goods produced and consumed at different locations. A supply factor for freight transport is the infrastructure quality which determines freight transport transit times and tariffs. The final freight transport volume is a result of interaction between demand and supply or, in other words, between forces and resistance.

The importance of resistance factors such as travel time and travel monetary costs will be explained in this chapter using so-called elasticities (see Chapter 3). In economics, elasticity is the ratio of the percentage change in one variable to the percentage change in another variable. Elasticity in this chapter is a tool for explaining the responsiveness of transport volumes to changes in resistance factors such as time and money. For example, the fuel price elasticity for car use explains the responsiveness of car use to changes in fuel price. If the fuel price elasticity is –0.2 it means that a fuel price increase of 1 per cent results in a car use decrease of 0.2 per cent, all other things being equal.

Sections 6.2 to 6.4 discuss the resistance factors of time, money and effort for passenger transport, respectively. The impact of resistance factors on freight transport is explained in section 6.5. Section 6.6 summarizes the main conclusions.

6.2 The role of travel time in passenger transport

Travel time components

The time required to make a trip is an important resistance factor. Table 6.1 shows that travel time can be split into different components. When compar-

Table 6.1 Possible travel time components from origin to destination (from top to bottom) for four modes of transport

Time – passenger transport			
Car	Public transport	Bicycle	Walking
	Hidden waiting time [a]		
Walking time to car park	Time to get to bus stop, bus, train or metro station	Time to get to bicycle storage location	
In-vehicle travel time: – free flow time – congestion time	In-vehicle travel time	Cycling time	Walking time
Time to find car park	Walking time transfer	Time to store bicycle	
Walking time from car park to final destination	Waiting time transfer	Time to get from bicycle storage facility to final destination	
	Time to get from bus, train or metro station to final destination		

Note: [a] Public transport travellers are dependent on the departure schedule as decided by the transport companies. Therefore travellers sometimes have to wait at their location of origin before it makes sense for them to depart.

ing total travel time from origin to destination between different transport modes, it could be an idea just to add up all the different time components (Table 6.1). However, by doing so the comparer forgets that people value time components differently. For example, public transport waiting time can be perceived as especially burdensome when travellers have to wait in difficult environments, such as in cold, hot or rainy weather, or in a seemingly unsafe or insecure condition (Iseki et al., 2006). Iseki et al. (2006) cite American modelling studies which showed that the value of walk time, compared to transit in-vehicle time, ranges between 2.0 and 4.5 for cities such as Houston, Cleveland, Minneapolis–St Paul and Chicago. In a large UK study on values of travel time (Albrantes and Wardman, 2011) it was found that walk and wait times are valued in the UK at somewhat less than twice in-vehicle time, with some evidence that valuations depend upon the levels the variables take. An interesting result from the study is that car time spent in congested traffic conditions is, on average, valued 34 per cent more highly than time spent in free-flow traffic. In other words, people are willing to pay more money to avoid congestion time than to have lower free-flow in-vehicle time. To state the obvious, people feel more resistance to congestion time than to free-flow time.

Table 6.2 UK meta-model implied monetary values of in-vehicle time for different users, motives and distances travelled

	Miles	Absolute values				Relative car users' value of car IVT		
User		Bus	Rail	Car	Car	Bus	Rail	Car
Valued		Bus	Rail	Bus	Car	Bus	Rail	Bus
Commute	2	4.4	8.0	7.7	6.7	0.65	1.20	1.15
	10	5.6	10.4	10.0	9.3	0.61	1.12	1.07
	50	n/a	17.0	n/a	13.0	n/a	1.31	n/a
	100	n/a	19.0	n/a	14.9	n/a	1.27	n/a
Other	2	3.9	7.2	6.9	6.0	0.65	1.20	1.15
	10	5.1	9.3	8.9	8.4	0.61	1.12	1.07
	50	n/a	15.2	n/a	11.6	n/a	1.31	n/a
	100	n/a	17.0	n/a	13.4	n/a	1.27	n/a
	200	n/a	19.0	n/a	15.4	n/a	1.23	n/a

Notes:
Values are for outside London and the South-East, RP data, nine comparisons and other terms at their base level; n/a denotes this distance is not applicable for the mode in question.
Monetary values in pence/minute in quarter 4 2008 prices; 100 pence exchange rate in 2008 was around 120 euro cents and 160 dollar cents.

Source: Albrantes and Wardman (2011).

In some transport studies researchers apply a weighted summation for the different time components (Table 6.1). In such summations, for example, one minute's walk time in a public transport transfer weighs heavier than one minute's in-vehicle time.

Value of time

As already remarked in the introduction to this chapter, the basic economic idea is that people and shippers choose transport modes with the lowest resistance. If only travel time determined resistance, people and shippers would choose the fastest transport mode, for with faster modes people get more time to carry out their preferred activities such as shopping, visiting family and friends and doing fun activities on their holidays. Shippers also tend to prefer low freight travel times because in that case they can transport the same amount of goods with fewer vehicles and personnel and thus save money. Therefore in transport economics the concepts of value of time (often abbreviated to VOT) or value of travel time savings (VTTS) play an important role (see also Chapters 3 and 12). Table 6.2 gives, as an illustration, UK monetary values for in-vehicle time for different modes, users and distances travelled. Albrantes and Wardman (2011) also found the so-called GDP elasticity for values of times, estimated over a 45-year period, to be 0.9 with a relatively narrow confidence interval. This means that with increasing incomes people are willing to pay more for travel time savings. In other words, the resistance factor for travel time increases in importance when people become richer (all other things being equal). This result justifies the widespread practice of uplifting the VOT in line with income, according to Albrantes and Wardman (2011).

The VTTS and VOT refer to the amount of money consumers or shippers are willing to pay to save a certain amount of travel time. In cost–benefit analysis (CBA) for new road infrastructure travel time savings (in money terms) are in most cases the most important societal benefit category.

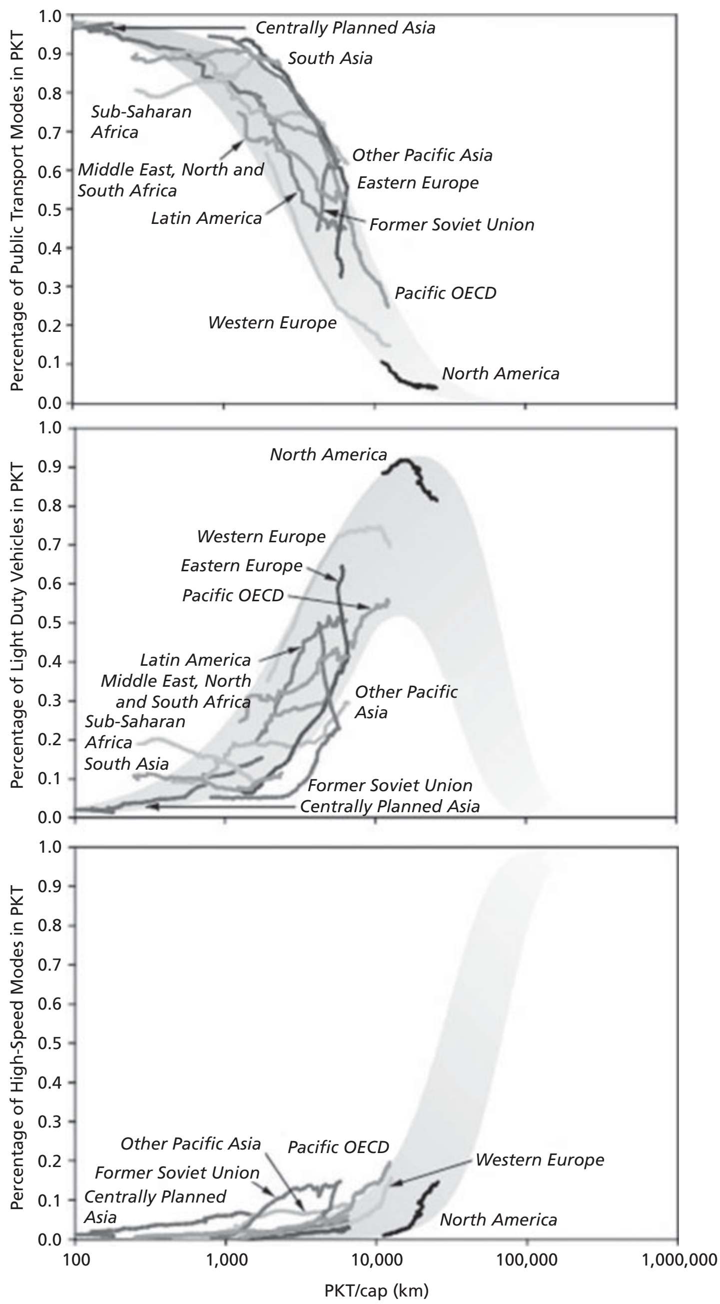
Different countries have estimated total annual congestion costs: for example, \$14 billion to \$200 billion for the US (VTPI, 2010), £20 billion to £30 billion for the UK (Goodwin, 2004) and €2 billion to €4 billion for the Netherlands (KiM, 2010). In these studies, using VOT estimates, the direct travel time losses compared to the 'ideal' free-flow situations are valued in monetary terms. These direct travel time losses are the main cost item in these estimates. In these estimates so-called indirect travel time costs are also considered. The reason is that in often congested areas some people and hauliers will involuntarily choose to avoid the traffic jams. For example, some people will choose to go by public transport and some hauliers will decide to change their planning by transporting goods outside rush hour. In both cases these involuntary choices are considered to be benefit losses (or costs) because, in the ideal, more free-flow situation, these people and hauliers would choose differently.

Constant time budgets

Consumers devote a limited amount of their time to travel. This is not extraordinary. However, the remarkable aspect is that this limited travel time budget has remained relatively constant on average on a country level for the past decade. Szalai (1972) and others carried out travel time research at the beginning of the 1970s in Eastern Europe, Western Europe and the United Stated and concluded that the average travel time per person is similar in all three regions despite the large differences in transport means and infrastructure. Schafer and Victor (2000) concluded that on average people spend 1.1 hour per day on travelling. They do that in the US, in Europe, in Africa, in South America and so forth. Of course, on an individual level large differences in travel time exist between, for example, a person living in a small African village and an inhabitant of Shanghai, but on the most aggregate level travel time budgets seem to be similar and fairly constant (Mokhtarian and Chen, 2004).

The theory of constant travel time budgets on the most aggregate level has important implications, as illustrated by Figure 6.1 (Schafer, 2006). If people keep their time budgets constant, they will travel longer distances when the transport resistance factor 'time' decreases, all other things being equal. The period 1950–2000 can be characterized as a period with a decrease in the transport time resistance factor: expansion of car ownership and of the road and air network took place all over the world. Additionally, in Japan and in parts of Europe (especially in France, Spain and Germany) a high-speed rail network was built. Consequently, all over the world there was a tendency to increase market share in kilometres people travelled by fast modes at the cost of the slower modes (Figure 6.1). Schafer (2006) expects this trend to continue (see the shaded parts of Figure 6.1). His expectation is based on two main assumptions. Firstly, people or households will need a higher absolute transport monetary budget in the future in order to be able to afford the faster but more expensive travel modes. As the world has shown, on average, a continuing economic growth (despite some occasional major dips), this first assumption of an increasing absolute household transport budget doesn't seem to be too wild. Secondly, faster transport modes will have to be available in the future. This assumption also doesn't seem to be too far off reality, as, for example, many governments at the time of writing this book are investing, or have plans to invest, in new airport capacity and high-speed rail all over the world.

Still, if the transport resistance factor of time decreases and consequently people travel longer distances, the question remains: why? Is happiness not to be found very close to one's own door? To be truthful, the answer to the 'why' question is not quite known. There are different explanations proposed. Marchetti (1994) has an anthropological explanation. According to him history shows that humans live just like animals defending and expanding their territory. Trying to find and explore new territories located further away is a basic instinct, in his view. Thus, if travel time resistance is increasingly lower, people tend to expand their territory. Economists argue that the probability of finding a new partner, a new job or a new house which satisfies people's preferences is perhaps higher in a larger searching area compared to looking in one's own village or town. Thus the chances of higher benefits drive the need for travelling longer distances, according to them. Others point at novelty-seeking or variety-seeking behaviour as explanations for a drive to travel (see, for example, Lal, 2006, who explains long-distance migration partly through genetic causes), so some humans perhaps have a genetic



Source: Schafer (2006).

Figure 6.1 Three phases in worldwide passenger mobility development (1950–2000): the decreasing share of relatively slow public transport (top), the growth and decline of the passenger car (middle) and the start of the high-speed era (high-speed trains and aircraft) urge to see the entire world or to change a holiday destination regularly, even if the previous destination met all of their expectations. Finally, information technology progress may also be an explanation for an increase in long-distance travel. Information technology such as e-mail, websites, voice-over IP programmes (for example, Skype), Facebook, tele-conferencing and so forth has made contact between people all over the world increasingly easier. If at the same time faster and not too expensive long-distance transport options are available, these increasing digital contacts may result also in increasing long-distance and fast physical travel, because people still love to have face-to-face contact once in a while.

Travel time elasticities and induced traffic

As mentioned in the introduction, the responsiveness of people to travel time changes is often expressed in so-called elasticities. A travel time elasticity of −1.0 means that, if travel time increases by 1 per cent, the amount of kilometres travelled decreases by 1 per cent. In scientific and applied transport research papers and reports, the reader can find an abundance of studies on travel time elasticities (e.g. Goodwin, 1996; de Jong and Gunn, 2001; Paulley et al., 2006; Hensher, 2008).

We give just one example. In a large European study, de Jong and Gunn (2001) compared modelled car travel time elasticities for commuting between different European regions: the Brussels region (Belgium), Italy and the Netherlands. They found short-term (less than one year) elasticities of −0.31 (Brussels), −0.87 (Italy) and −0.64 (the Netherlands), and long-term elasticities of −0.49 (Brussels), −1.38 (Italy) and −2 (the Netherlands). This example illustrates five important aspects:

1. In interpreting elasticities it is always important to be aware of the specifications. In this case, the travel time elasticity is only related to car commuting travel volumes.
2. It is always important to realize that the difference between the elasticities could be explained because of methodological reasons. For the purpose of this book, there would be too much detail to discuss methodological issues related to transport elasticities (see Hensher, 2008, for an assessment of systematic sources of variation in public transport elasticities). Here, the main message is that, whenever elasticities from other studies are used, the user should be aware of the influence of methodological choices on the outcome.
3. The same elasticity may differ significantly between regions and/or countries, as this example shows, because of different transport cir-

cumstances. For example, the regions may differ significantly in public transport quality and availability, which makes substitution to public transport when car travel time increases in one region far easier compared to another region. So this example shows that it is quite risky in policy studies to transfer uncritical elasticities established for one specific region or country to another. If these elasticities are then used for that other region to estimate, for example, the policy impact of a measure to increase public transport travel times, the policy conclusions could be completely wrong.

4. The long-term elasticities are higher compared to the short-term elasticities. This is an often occurring phenomenon related to the fact that in the long term people have more choice options when travel times decrease, such as moving or looking for another job, compared to the short term.
5. Car commuting is relatively sensitive to travel time changes. The long-term elasticities are often higher than −1, while the average travel time is −1 based on constant travel time budgets (Goodwin, 1996).

From a policy planning perspective, the notion that people are sensitive to travel time changes leads to an important consequence. When governments invest in new roads to relieve present or expected future congestion – as they often do all over the world – so-called induced traffic will arise. Induced traffic refers to all the traffic which would be present if an expansion of road capacity occurred which would not be there without the expansion (Goodwin and Noland, 2003). These authors also state that the induced traffic hypothesis is, in essence, only that there exists a demand curve for travel – the cheaper the travel, the more will be demanded. So one effect of making transport 'cheaper' (i.e. faster) by building new roads or extra road lanes is that it results in new traffic. Induced traffic is also related to the indirect travel time losses (see 'Value of time' above). Consequently, congestion relief will be less than anticipated or shorter in duration than if there is no such extra traffic. This will influence the cost–benefit appraisal (see Chapter 14) of the road project, as well as environmental impacts. From about 1970 to the mid-1990s, UK official practice, as in other countries, was mostly to assume that an induced traffic effect was negligible, until mounting evidence in a now famous report by the SACTRA (1994) changed this notion. Hymel et al. (2010) have estimated an induced demand effect in the US of about 0.16 in the long run. This figure means that expanding total road mileage in the US by 1 per cent will result in a 0.16 per cent increase in car use. Hymel et al. (2010) conclude that their result for the US is relatively low compared to that of prior literature that they reviewed, which best supports long-run road mileage elasticities for car use to the order of 0.4.

Cross-travel time elasticities

In many countries policy-makers hope to achieve policy goals such as fewer traffic jams and environmental benefits by investing in public transport. For example, in 2001 the European Commission published a policy strategy document: the White Paper *Time to Decide* (EC, 2001). In this document the EC recognized 'the railways as still being a means of transport with major potential, and it is renewal of the railways which is the key to achieving modal rebalance' (EC, 2001: 26). One of the means governments apply to achieve rebalance is investing in new or extra rail tracks in order to increase rail transport's quality and speed. However, despite all kinds of efforts, the share of car transport increased in Europe in the period 1990–2009 (van Essen et al., 2009). This points at a low responsiveness of car users to shift to public transport when improvements in public transport are carried out. A way of expressing this responsiveness is by using cross-travel time elasticities. The term 'cross' means that these kinds of elasticities reflect the responsiveness of a percentage change in a characteristic in one mode (e.g. public transport travel time changes) to a percentage change in the use of another mode (e.g. car use). Paulley et al. (2006) cite UK rail time cross-elasticities for car use of 0.057 and for coach use of 0.20. These figures imply that improving rail travel time will have a relatively low impact on car use. The impact on the competing coach market is higher. However, it is important to note that in absolute numbers the car use decrease may still be significant in the UK when train travel time increases by 1 per cent. The reason is that car use expressed in car kilometres travelled has a high market share in the UK, implying that a 0.057 per cent decrease is still a relatively high amount of absolute car kilometres which are substituted to train kilometres. With cross-elasticities it is always important to be aware of the market shares of the modes considered. Transferring a cross-elasticity estimated for a certain region to another region can give completely wrong policy information if the two regions differ significantly in transport mode market shares (Balcombe et al., 2004). Nevertheless, the rather low car responsiveness is still disappointing for many politicians. One reason for the low responsiveness is that consumers take other factors into account as well as travel time when deciding to take a certain mode of transport. These include monetary costs (see section 6.3) and effort (see section 6.4). Habit also plays a role in the low responsiveness (see, for example, Aarts, 1996, and Chapter 3). The role of habit in decision-making is that people do not tend to consider all the pros and cons of a choice all the time. This practice saves time and energy. Taking the car for commuting could become a kind of habit, and the people with this habit are unaware of or not interested in – perhaps better put, they are less open to – information on positive changes in competing transport modes.

Changes in choices often happen only when large new events take place in people's life such as obtaining a new job or moving house. Large changes in transport mode characteristics could also trigger reconsideration of transport mode choice, such as the opening of a completely new train line to a village or suburb or sudden large increases in road congestion.

6.3 The role of travel monetary costs in passenger transport

The second important resistance factor is the money people have to spend on the trip. Like travel time, monetary costs can be split into sub-components (Table 6.3).

Monetary costs can be classified in many ways. An often used classification is fixed costs versus variable costs. Fixed costs are the amount of money to be paid independent of the amount of kilometres travelled. Depreciation costs and yearly annual taxes are examples. In contrast, variable costs are dependent on the number of kilometres travelled. Examples are car fuel costs and public transport fares.

Constant money cost budgets

There seems to be a constant money budget for people's mobility as a percentage of people's income. Schafer (1998) researched worldwide mobility expenditures and confirmed a result previously concluded by Zahavi (1979):

Table 6.3 Possible travel money cost components for four modes of transport

Money – passenger transport			
Car	**Public transport**	**Bicycle**	**Walking**
Depreciation costs	Fares	Depreciation costs	Depreciation costs (shoes)
Car maintenance costs	Costs for trip to and from station (e.g. taxi)	Maintenance costs	Repair costs (shoes)
Parking costs		Parking costs (in commercial storing facilities)	
Tolls		Insurance costs	
Fuel costs			
Taxes			
Insurance costs			

on average people per class of income tend to spend a constant share of their income on transportation. For countries within the Organisation for Economic Co-operation and Development (OECD) this share is estimated at roughly 10 to 15 per cent (Schafer, 1998; see also Chapter 3). One implication of this constancy is that, if the resistance factor of monetary costs increases, travel decreases and vice versa, all other things being equal.

Price and monetary cost elasticities

The responsiveness to monetary changes can be expressed in elasticities, just as in the travel time case (section 6.2). A large amount of transport price, fare and cost elasticities can be found from all over the world (e.g. Graham and Glaister, 2002, 2004; Hanly et al., 2002; Gillen et al., 2004; Goodwin et al., 2004; Brons, 2006; Paulley et al., 2006; Small and Van Dender, 2007; Brons et al., 2008; Barla et al., 2009; Dahl, 2011; VTPI, 2011).

To avoid confusion, it is important to realize that one can find price and fare elasticities and travel cost elasticities in this literature. Price elasticities relate to the responsiveness to changes in prices such as fuel price. Travel or transport cost elasticities relate to behavioural changes dependent on actual cost changes. To be clear, for example, in a certain region the price of petrol is 1.50 euros. All car users who fill their petrol tank in that region have to pay this fuel price. However, the actual petrol cost per kilometre driven in that region could be 0.10 euros for a relatively fuel-efficient car user and 0.15 euros for a less fuel-efficient car user. So it seems plausible that fuel price increases will result not only in fewer car kilometres to avoid the increased travelling costs but probably also in the purchase of more fuel-efficient cars in order to avoid the cost increase. In other words, it seems plausible that fuel price elasticities for fuel use are higher compared to fuel cost elasticities for fuel use. To put it differently, fuel price elasticities for car use will probably be lower compared to fuel price elasticities for fuel use (see further below when it will be shown that this phenomenon is indeed true).

It is impossible to summarize fully the huge amount of scientific research. Very broadly, one could conclude that empirical research worldwide shows that:

1. transport consumers are indeed price and cost sensitive;
2. the extent of their responsiveness is dependent on many factors (culture, income, time and so forth);
3. the responsiveness to price changes is fairly modest in most cases (elasticities are in most cases between 0 and 1).

Table 6.4a Indicative fuel price elasticities

	Short term (1 year)	Long term (5 to 10 years)
Car ownership	–0.05 to –0.2	–0.1 to –0.65
Car use	–0.1 to –0.2	–0.25 to –0.5
Car fuel efficiency	0.1 to 0.15	0.3 to 0.4
Car fuel use	–0.25 to –0.35	–0.6 to –0.8

Source: Geilenkirchen et al. (2009).

Table 6.4b Indicative fare elasticities

	Short term (1 year)	Long term (5 to 10 years)
Bus	–0.2 to –0.5	–0.6 to –1.0
Train	–0.3 to –0.7	–0.6 to –1.1
Metro	–0.1 to –0.3	–0.3 to –0.7

Source: Geilenkirchen et al. (2009).

As an example, Geilenkirchen et al. (2009) have summarized fuel price elasticities for car use and car ownership (Table 6.4a) and fare elasticities for public transport use (Table 6.4b), as found in numerous studies and reviews. The numbers presented are applicable – more or less – for Western European countries. However, as in the case of the time travel elasticities (see before), it is important to stress that the specific elasticities for a region can differ considerably from the values presented in Tables 6.4a and 6.4b dependent on the specific geographic, cultural and technical circumstances. The elasticities presented in Tables 6.4a and 6.4b have to be considered as indicative.

Tables 6.4a and 6.4b show two interesting aspects. Firstly, regardless of the exact values, it is clear that the long-term fuel price elasticity for car use and car fuel usage is higher compared to the short-term elasticities. Also, the price responsiveness for public transport usage is higher in the long term. The explanation is simple: in the short term it is relatively difficult for people to make changes. In the longer term, this is different. Then people can choose a different car or change their dwelling or job location. Secondly, fuel price elasticities for car fuel use are indeed higher compared to the fuel price elasticities for car use, by roughly two times (Table 6.4a). Especially in the long run, the elasticities (–0.6 to –0.8) show that, as a response to a fuel price increase, people not only use their car somewhat less but, even more, try to avoid higher fuel costs by purchasing more fuel-efficient cars or, eventually, by driving in a way that is more economical on fuel.

Dargay and Gately (1997) concluded that consumers show a stronger response to price increases compared to price decreases. This implies that a fuel price increase followed by a price decrease of the same magnitude does not result in restoring the transport and fuel demand which occurred just before the price increase. Dargay (2007) also found that income increase has a greater impact on car ownership compared to an income decrease of the same magnitude.

6.4 Effort resistance factors

Next to time and money, there are more resistance factors which determine the amount of passenger transport. We summarize these factors in this section as 'effort' factors. Effort may seem to be a relatively unimportant resistance factor. However, Chapter 3 shows the existence of a large amount of social and psychological factors which influence travel behaviour.

Effort, as a transport resistance factor, consists of different aspects (Table 6.5).

All effort resistance aspects given can influence the decision to travel or the decision to choose a certain mode. In their introduction to a special issue on the role of cycling and walking in advancing healthy and sustainable urban

Table 6.5 Possible effort aspects for four modes of transport which can be regarded as resistance factors

Effort – passenger transport			
Car	**Public transport**	**Bicycle**	**Walking**
Discomfort or physical and mental effort of car driving	Discomfort	Discomfort or physical effort[a]	Discomfort or physical effort[a]
(Mental) strain, stress	(Mental) strain, stress		
Reliability	Reliability	Accident risk	Accident risk
Accident risk	Physical effort (stairs, walks during transfer, carrying luggage)	Travellers' feelings of safety	Travellers' feelings of safety
Availability of information	Availability of information	Availability of information	Availability of information
	Travellers' feelings of safety		

Note: [a] Physical effort due to cycling and walking is perhaps the only resistance factor that is also a kind of attractiveness factor, as some people may assess the physical effort of cycling and walking as additional reasons to choose these modes, for example for commuting. These modes give them an 'easy' daily opportunity for exercise.

areas, Tight and Givoni (2010: 386) give a summary of possible different resistance factors for these slow modes: 'both types of user are exposed to the weather and the environment, are very vulnerable if involved in a collision with a motor vehicle, are unlicensed, and both are often poorly represented in the provision of facilities and infrastructure compared to other modes'.

We will now discuss the effort factors in more detail.

Discomfort and physical effort

The resistance factor of discomfort contains a large number of different issues. Discomfort is considered an important resistance factor especially when related to public transport. In Balcombe et al. (2004) and Paulley et al. (2006) different aspects are mentioned which may influence the comfort of public transport travelling: the quality of the waiting environment, the quality of the vehicle and rolling stock, the quality of the front-line staff to customers, crowding, seat-place availability, the quality of on-board facilities, cleanliness and the quality of interchange between modes. We refer to Chapter 8 in Balcombe et al. (2004) or paragraph 3 in Paulley et al. (2006) for more detailed information and empirical studies on these comfort aspects.

For cycling, discomfort factors related to the natural environment have a large influence on both the decision to cycle and the frequency (Heinen et al., 2010). Heinen et al. (2010), based on an overview of the literature on commuting by bicycle, found that hilliness has a negative effect on cycling and weather has a large influence on the cycling frequency. Commuters are less influenced by temperature than other cyclists, implying that many people only choose to cycle for leisure purposes when the weather is pleasant (Heinen et al., 2010).

Reliability

A reliable transportation system means that travellers and hauliers can make trips according to their expectations, especially related to expected travel time. According to Bates et al. (2001), travel time reliability is a potentially critical influence on any mode or route choice. Although reliability is related to travel time, it is important to stress that reliability is a separate resistance factor. Travel time reliability is a measure of the expected range in travel time. For example, a car commuter who faces the 'same' traffic jam every working day, with a 30- to 40-minute delay, loses travel time compared to free flow but still has a rather reliable trip. This commuter knows beforehand more or less exactly what time (within a 10-minute spread) he or she will be at work or at home and can make arrangements accordingly. However, if the traffic jam is unpredictable – sometimes it is only a few minutes, another day suddenly more than an hour – this commuting trip becomes highly unreliable. The consequence of the unpredictability of travel times means that people will have to adapt their behaviour but because of the unpredictability they are uncertain about the best course of action. This uncertainty comes with a cost. So, like a VOT, one would expect also that a so-called value of reliability (VOR) exists: the willingness to pay to make trips more reliable or, to put it differently, more predictable. Lam and Small (2001) have measured the value of time and reliability from 1998 data on the actual behaviour of commuters on State Route 91 in Orange County, California, US, where the commuters could choose between a free and a variably tolled route. They indeed found VORs. In their best model (according to Lam and Small, 2001), VOT for these commuters amounted to \$22.87 per hour (2001 prices), while VOR was \$15.12 per hour for men and \$31.91 per hour for women. So, in this case, women especially seem averse to travel time variability.

Travel information

Another effort resistance factor is information or, perhaps better put, the insufficient availability of travel information or good-quality travel information. There are numerous studies on traveller behaviour with limited knowledge (for an overview, see Chorus et al., 2006). These studies have uncovered travellers' dislike of knowledge limitations and their inclination to reduce these knowledge limitations by acquiring information. Broadly speaking, traveller information relates to route information (en route and beforehand) and mode information (e.g. fares and travel times to be expected, waiting and bicycle storing facilities and so forth). The availability of high-quality traveller information can improve travelling comfort and trip reliability and decrease travel stress. Chorus et al. (2010) used search theory to evaluate the value of travel information. Their results indicate, amongst other things, that travellers prefer information that adds previously unknown alternatives to their choice-set over information that provides estimates for uncertain attributes of known alternatives. As can be expected, Chorus et al. (2010) found a substantial heterogeneity with respect to travellers' valuation of the costs and benefits of travel information.

Travellers' feelings of safety

Travellers' feelings of safety are a resistance factor for public transport and the slow modes of cycling and walking. The factor relates to feelings of uneasiness when people have to wait or transfer at remote stations or bus stops or when they have to travel in almost empty trains, buses or subways later at night. Using dark or remote roads for cycling and walking can also be an unattractive endeavour. Heinen et al. (2010) mention darkness as a factor which results in people choosing to cycle less. In all cases, the resistance is caused by the fact that people may just feel uneasy in an undefined way or they may actually fear being harassed or robbed. Lack of travellers' safety measures, especially for females, seems to be an important resistance factor for choosing these modes for travel, especially in the evenings.

Accident risk

This factor is somewhat related to the previous resistance factor but here we mean that people might fear that they, or their children, will be involved in traffic accidents if they choose a certain transport mode. This fear might influence their mode choice. An interesting example is the study by Adams (1999) (also mentioned in Chapter 12) which showed that in 1971 around 80 per cent of British children went to school by themselves. In 1990 this share had decreased to only 9 per cent. The most important reason for this was that parents were said to be afraid of traffic accidents. In their overview of the literature on commuting by bicycle, Heinen et al. (2010) found that safety (low accident risk) is often mentioned in the literature as a reason not to cycle. Interestingly, they cite one piece of research which showed that all respondents thought that cycling was less safe than using other modes, but cyclists gave the highest rating for bicycle safety.

Mental strain, stress

Stress can be seen as an indicator for the importance of effort resistance factors. Wener and Evans (2011) compared the stress of car and train commuters in metropolitan New York City. In their paper they mention several studies that relate travel effort factors to stress such as crowded trains, discomfort from poor design, feelings of no control (a car driver often finds that he or she has more control over a particular trip when driving a car compared to using public transport) and unpredictability. Wener and Evans (2011) found that car commuters showed significantly higher levels of reported stress compared to train commuters. Driving effort and predictability largely accounted for the elevated stress associated with car commuting, according to this study.

Specific constants

As the resistance factors of time and money are often objectively measurable, this is far more difficult for most of the effort resistance factors. An often used method to include effort components in the total resistance factor is by using specific weighting factors for the different parts of the trip such as the penalties already mentioned for waiting time or arriving unexpectedly too early at the destination. In transport models, often so-called specific constants are used in the resistance functions per transport mode in order to include all kinds of effort components (Chapter 15).

6.5 Goods transport and resistance factors

In sections 6.2 to 6.4, we have focused on passenger transport. In this section we focus specifically on the role of transport resistance factors in freight transport. Broadly speaking, shippers and hauliers take transport resistance factors into account in a similar way to that of consumers. Monetary costs and transit time are also important factors shippers consider when deciding on a specific transport mode. Additionally, 'transport service quality' (reliability, frequency and so forth) plays a role in the decision-making process for transport freight and by what mode. Like the 'effort' factor this may seem a vague resistance factor, but increasing empirical evidence shows that it is an important factor.

Transit time

It seems obvious that in freight transport actors are willing to pay for transit time savings. For example, when, owing to infrastructural improvements, freight transport time is saved, hauliers can deliver the same amount of goods in less time compared to the situation without the infrastructural improvement and, by so doing, can save personnel and vehicle operating costs. Feo et al. (2011) present different VOTs obtained by themselves and other researchers for road haulage. The values presented differ significantly, which is not surprising, because empirical evidence (reviewed by Punakivi and Hinkka, 2006; Beuthe and Bouffioux, 2008) showed that freight transport costs and transit times are valued differently in relative terms depending on the nature and the peculiarities of the freight being shipped. Feo et al. (2011) present VOT values for road haulage of between 0.01 and 340.73 (expressed in 2005 euros per hour per ton). The extreme high value is from a Norwegian study cited in Feo et al. (2011) and applies to edible refrigerated goods. It makes sense that, especially for these perishable goods, shippers are willing to pay relatively high amounts of money to save transit time. Most VOT values presented by Feo et al. (2011) are in the range of 0.5 to 3 euros per ton per hour (at 2005 prices).

Table 6.6 Indicative price elasticities for road freight transport demand

Price elasticity, total[a]	–0.6 to –0.9
Whereof substitution (fewer tonne kilometres road transport, more tonne kilometres rail and/or barge, similar tonnes production and consumption)	–0.4 to –0.5
Whereof less transport (fewer tonne kilometres road transport, similar tonne kilometres rail and/or barge, similar tonnes production and consumption)	–0.2 to –0.4
Whereof less production and consumption (fewer tonne kilometres road transport, similar tonne kilometres rail and/or barge, fewer tonnes production and consumption)	Low (<–0.1)

Notes:
Road freight transport demand in tonne kilometres.
[a] The top row is a summation of the three rows underneath. The elasticities are valid for limited price increases.

Source: Geilenkirchen et al. (2009).

Monetary costs

As in the case of passenger transport, the monetary cost resistance factor for freight transport can be explained using elasticities. For example, Table 6.6 shows the responsiveness of road freight transport in the Netherlands to price increases. The figures are based on international literature but are highly indicative as empirical data are relatively scarce. It should be noted that freight demand elasticity studies vary significantly in terms of the demand measure, data type, estimation method, commodity type and so forth (Li et al., 2011). According to Li et al. (2011), this wide variation makes it difficult to compare empirical estimates when the differences may arise partly from the methods and data used.

Many people argue that freight transport is non-responsive to price increases as transport costs have a modest share in final product prices. To give an example, in 2002 in the US total transportation costs amounted to roughly only 5 to 6 per cent of US domestic product in the same year (MacroSys Research and Technology, 2005). However, as Table 6.6 illustrates for road transport, this does not mean that price increases will not affect road transport volumes (in tonne kilometres), as shippers and hauliers have different possible behavioural reactions to deal with a road price increase as well as just passing on the price increase to the consumers of goods.

Firstly, transport mode substitution may take place (Table 6.6, second row). Increase of road freight costs by new taxes, tolls or oil price rises, for example, may entice shippers to switch to other modes, such as rail. Secondly, hauliers may react by trying to implement increased efficiency in the road freight sector (Table 6.6, third row) in terms of using larger vehicles (more tonnes per kilometre driven), reduced number of empty runs (more tonnes per kilometre driven), improved loading (more tonnes per kilometre driven), buying more fuel-efficient trucks (lower fuel costs per kilometre driven) and concentration in the haulier business (cost reductions because of less overhead), for example. Thirdly, less production and consumption of freight may take place (Table 6.6, fourth row). Shippers may include (a part of the) higher transport costs in the final product price. It is imaginable that for some products the resulting higher prices will lead to lower demand and, subsequently, lower production. Nevertheless, this third possible response to a transport price increase is not very strong, as shown in Table 6.6. The reason is twofold. Firstly – as remarked on before – freight transport costs have only a modest share in final product prices. Secondly, by mode substitution and/or increased transport efficiency (more tonnes per kilometre lifted) hauliers will absorb some or all of the price increase.

Transport service

Shippers also take transport service quality into account when deciding to use a particular transport freight mode. Arunotayanun and Polak (2011) found that service quality aspects are such things as reliability, safety, truck condition, travel route, rail terminal access and train information and service flexibility (i.e. time of departure, flexibility of service and responsiveness to problems). These authors investigated shippers' mode choice behaviour in Java, Indonesia. Their results indicate the presence of significant levels of taste heterogeneity among shippers, only some of which can be accounted for by conventional commodity-type based segmentations. One policy implication of this is that it is often far more difficult to change freight modal share than policy-makers may think. The reason is that shippers not only take a relatively large amount of resistance factors into account (not only money costs and transit times but also all kinds of transport service aspects) when deciding to use a certain transport mode, but also each choose very differently, even within one type of commodity.

Feo et al. (2011) found empirical evidence for the willingness to pay for transport service improvements. They investigated freight forwarders' modal choice on the so-called south-west Europe motorway of the sea. By creating motorways of the sea, the European Union hopes to provide door-to-door maritime services that can compete with the road alternative. Feo et al. (2011) found, amongst other things, a value of reliability of 9 euros (at 2005 prices) for a 1 per cent reduction in the number of road shipments that fail to comply with delivery times and conditions. As regards maritime frequency, their results indicated that the freight forwarders would be willing to pay 24 euros per shipment if the number of weekly maritime dispatches were increased by one additional departure. The reason is probably that, by increasing the maritime frequency, freight forwarders can operate more flexibly because they have more maritime transport options.

6.6 Conclusions

The most important conclusions of this chapter are summarized as follows:

1. Important transport resistance factors for passenger transport are travel time, monetary costs and effort. Passengers are responsive to changes in these resistance factors.
2. Important resistance factors for freight transport are transit time, monetary costs and transport services. Shippers and freight forwarders are responsive to changes in these resistance factors.
3. On an aggregated scale (large region or country level) it turns out that people, on average, tend to keep their travel time budget constant. Thus, on an aggregate level, people tend to exchange lower transport travel times for more kilometres travelled, and not to spend the saved travel time on other activities.
4. Travel time and price elasticities for passenger transport in the long run are higher compared to short-term elasticities. This phenomenon is related to the fact that in the long term people have more choice options when travel times or prices change compared to the short term, such as moving or finding another job.
5. A rise in transportation monetary costs for a certain mode results in fewer transport and/or mode shifts. However, higher fuel prices result not only in less road transport and to a small extent a mode shift, but also in the purchase of more fuel-efficient cars. By buying more fuel-efficient cars people can avoid (a part of) the price increase and thus can continue to travel by car.

REFERENCES

Aarts, H. (1996), *Habit and Decision Making: The Case of Travel Mode Choice*, Nijmegen: Katholieke Universiteit.

Adams, J. (1999), 'The social implications of hypermobility: speculations about the social consequences of the OECD scenarios for environmentally sustainable transport and business-as-usual trend projections', in Project on Environmentally Sustainable Transport (EST), *The Economic Implications of Sustainable Transportation*, Ottawa Workshop, 20–21 October 1998, Paris: OECD.

Albrantes, P.A.L. and M.R. Wardman (2011), 'Meta-analysis of U.K. values of travel time: an update', *Transportation Research Part A*, 45, 1–17.

Arunotayanun, K. and J.W. Polak (2011), 'Taste heterogeneity and market segmentation in freight shippers' mode choice behaviour', *Transportation Research Part E*, 47, 138–148.

Balcombe, R., R. Mackett, N. Paulley, J. Preston, J. Shires, H. Titheridge, M. Wardman and P. White (2004), *The Demand for Public Transport: A Practical Guide*, TRL report TRL593, accessed March 2011 at http://www.demandforpublictransport.co.uk/TRL593.pdf.

Barla, P., B. Lamonde, L.F. Miranda Morena and N. Boucher (2009), 'Travelled distance, stock and fuel efficiency of private vehicles in Canada: price elasticities and rebound effect', *Transportation*, 36, 389–402.

Bates, J., J. Polak, P. Jones and A. Cook (2001), 'The valuation of reliability', *Transportation Research Part E*, 37 (2–3), 191–229.

Beuthe, M. and C. Bouffioux (2008), 'Analysing qualitative attributes of freight transport from stated orders of preference experiment', *Journal of Transport Economics and Policy*, 42, 105–128.

Brons, M.R.E. (2006), *Meta-Analytical Studies in Transport Economics: Methodology and Applications*, Amsterdam: Vrije Universiteit.

Brons, M., P. Nijkamp, E. Pels and P. Rietveld (2008), 'A meta-analysis of the price elasticity of gasoline demand: a SUR approach', *Energy Economics*, 30, 2105–2122.

Chorus, C.G., E.J.E. Molin and G.P. van Wee (2006), 'Use and effects of advanced traveller information services (ATIS): a review of the literature', *Transport Reviews*, 26, pp. 127–149.

Chorus, C.G., J.L. Walker and M.E. Ben-Akiva (2010), 'The value of travel information: a search-theoretic approach', *Journal of Intelligent Transportation Systems*, 14 (3), 154–165.

Dahl, C.A. (2011), 'Measuring global gasoline and diesel price elasticities', *Energy Policy*, 41, 2–13.

Dargay, J.M. (2007), 'The effect of prices and income on car travel in the UK', *Transportation Research Part A*, 41, 949–960.

Dargay, J.M. and D. Gately (1997), 'The demand for transportation fuels: imperfect price-reversibility?', *Transportation Research*, 31, 71–82.

De Jong, G. and H. Gunn (2001), 'Recent evidence on car cost and time elasticities of travel demand in Europe', *Journal of Transport Economics and Policy*, 35 (2), 137–160.

EC (European Commission) (2001), *White Paper: European Transport Policy for 2010: Time to Decide*, COM(2001)370, Brussels: EC.

Feo, M., R. Espino and L. Garcia (2011), 'A stated preference analysis of Spanish freight forwarders modal choice on the south-west Europe motorway of the sea', *Transport Policy*, 18, 60–67.

Geilenkirchen, G.P., K.T. Geurs, H.P. van Essen, A. Schroten and B. Boon (2009), *Effecten van prijsbeleid in verkeer en vervoer*, Bilthoven, Planbureau voor de Leefomgeving/Delft: CE Delft.

Gillen, D.W., W.G. Morrison and C. Stewart (2004), 'Air travel demand elasticities: concepts, issues, and measurement', accessed at http://www.fin.gc.ca/consultresp/Airtravel/airtravStdye.html.

Goodwin, P.B. (1996), 'Empirical evidence on induced traffic: a review and synthesis', *Transportation*, 23, 35–54.

Goodwin, P. (2004), *The Economic Costs of Road Traffic Congestion*, London: ESRC Transport Studies Unit, University College London.

Goodwin, P., Dargay, J. and M. Hanly (2004), 'Elasticities of road traffic and fuel consumption with respect to price and income: a review', *Transport Reviews*, 24, 275–292.

Goodwin, P. and R. Noland (2003), 'Building new roads really does create extra traffic: a response to Prakash et al.', *Applied Economics*, 35, 1451–1457.

Graham, D.J. and S. Glaister (2002), 'The demand for automobile fuel: a survey of elasticities', *Journal of Transport Economics and Policy*, 36 (1), 1–26.

Graham, D.J. and S. Glaister (2004), 'Road traffic demand elasticity estimates: a review', *Transport Reviews*, 24 (3), 261–274.

Hanly, M., J. Dargay and P. Goodwin (2002), 'Review of income and price elasticities in the demand for road traffic', ESRC TSU publication 2002/13, ESRC Transport Studies Unit, Centre for Transport Studies, University College London.

Heinen, E., B. van Wee and K. Maat (2010), 'Commuting by bicycle: an overview of the literature', *Transport Reviews*, 30 (1), 59–96.

Hensher, D.A. (2008), 'Assessing systematic sources of variation in public transport elasticities: some comparative warnings', *Transportation Research Part A*, 42, 1031–1042.

Hymel, K.M., K.A. Small and K. Van Dender (2010), 'Induced demand and rebound effects in road transport', *Transportation Research Part B*, 44 (10), 1220–1241.

Iseki, H., B.D. Taylor and M. Miller (2006), *The Effects of Out-of-Vehicle Time or Travel Behavior: Implications for Transit Transfers*, Los Angeles: Institute of Transportation Studies, University of California, Los Angeles, accessed at http://www.its.ucla.edu/research/EPIC/Appendix%20A.pdf.

KiM (Kennisinstituut voor Mobiliteitsbeleid) (2010), *Mobiliteitsbalans 2010*, The Hague: KiM.

Lal, D. (2006), *Reviving the Invisible Hand: The Case for Classical Liberalism in the Twenty-First Century*, Princeton, NJ: Princeton University Press.

Lam, T.C. and K.A. Small (2001), 'The value of time and reliability: measurement from a value pricing experiment', *Transportation Research Part E*, 37, 231–251.

Li, Z., D.A. Hensher and J.M. Rose (2011), 'Identifying sources of systematic variation in direct price elasticities from revealed preference studies of inter-city freight demand', *Transport Policy*, 18 (5), 727–734.

MacroSys Research and Technology (2005), 'Logistics costs and U.S. gross domestic product', accessed at http://www.ops.fhwa.dot.gov/freight/freight_analysis/econ_methods/lcdp_rep/index.htm.

Marchetti, C. (1994), 'Anthropological invariants in travel behavior', *Technological Forecasting and Social Change*, 47, 75–88.

Mokhtarian, P.L. and C. Chen (2004), 'TTB or not TTB, that is the question: a review and analysis of the empirical literature on travel, time (and money) budgets', *Transportation Research Part A*, 38 (9/10), 643–675.

Paulley, N., R. Balcombe, R. Mackett, H. Titheridge, J. Preston, M. Wardman, J. Shires and P. White (2006), 'The demand for public transport: the effects of fares, quality of service, income and car ownership', *Transport Policy*, 12, 295–306.

Punakivi, M.M. and V. Hinkka (2006), 'Selection criteria of transportation mode: a case study in four Finnish industry sectors', *Transport Reviews*, 26 (2), 207–219.

SACTRA (Standing Advisory Committee for Trunk Road Assessment) (1994), *Trunk Roads and the Generation of Traffic*, London: HMSO.

Schafer, A. (1998), 'The global demand for motorized mobility', *Transportation Research A*, 32 (6), 455–477.

Schafer, A. (2006), 'Long-term trends in global passenger mobility', *The Bridge*, 36 (4), 24–32.

Schafer, A. and D.G. Victor (2000), 'The future mobility of the world population', *Transportation Research A*, 34 (3), 171–205.

Small, K.A. and K. Van Dender (2007), 'Fuel efficiency and motor vehicle travel: the declining rebound effect', *Energy Journal*, 28 (1), 25–51.

Szalai, A. (ed.) (1972), *The Use of Time: Daily Activities of Urban and Suburban Populations in Twelve Countries*, The Hague: Mouton.

Tight, M. and M. Givoni (2010), 'The role of walking and cycling in advancing healthy and sustainable urban areas', *Built Environment*, 36 (4), 385–390.

Van Essen, H., X. Rijkee, G. Verbraak, H. Quak and I. Wilmink (2009), 'Modal split and decoupling options', CE Delft, accessed at www.eutransportghg2050.eu.

VTPI (Victoria Transport Policy Institute) (2011), 'Transportation cost and benefit analysis II: congestion costs', accessed January 2011 at www.vtpi.org.

Wener, R.E. and G.W. Evans (2011), 'Comparing stress of car and train commuters', *Transportation Research Part F*, 14, 111–116.

Zahavi, Y. (1979), *The UMOT Project*, Washington, DC: US Ministry of Transport.

7

Traffic flow theory and modelling

Serge Hoogendoorn and Victor Knoop

7.1 Introduction

When do traffic jams emerge? Can we predict, given certain demand levels, when queuing will occur, how long the queues will be, how they will propagate in space and time and how long it will take for the congestion to resolve? Why does an overloaded traffic network underperform? This chapter gives a basic introduction to traffic flow theory which can help to answer these kinds of questions.

We start this chapter by explaining how it connects with the other chapters in this book (see Figure 7.1). In the top left of the figure, the reader will recognize the conceptual model used in Chapters 2 to 6, in a highly simplified form, to explain transport and traffic volumes.

One of the results of the interplay between people's and shippers' needs and desires, the locations of activities and the transport resistance factors (Figure 7.1, top left) is a certain volume of road traffic (Figure 7.1, middle left). Road traffic, and this is where this chapter starts, can be described by using flow variables such as speed and density (Figure 7.1, middle right). The density of traffic is the number of vehicles that are present on a roadway per unit distance. Road traffic flows on certain road stretches during certain time periods can be either free or congested and/or the flows can be unreliable. In the last two cases the transport resistance on these road stretches will be relatively high, as explained in Chapter 6 using concepts such as value of time and value of reliability, among other things. Consequently, high transport resistance implies negative repercussions on road traffic volumes (see the arrow from flow variables to transport resistance, Figure 7.1, top).

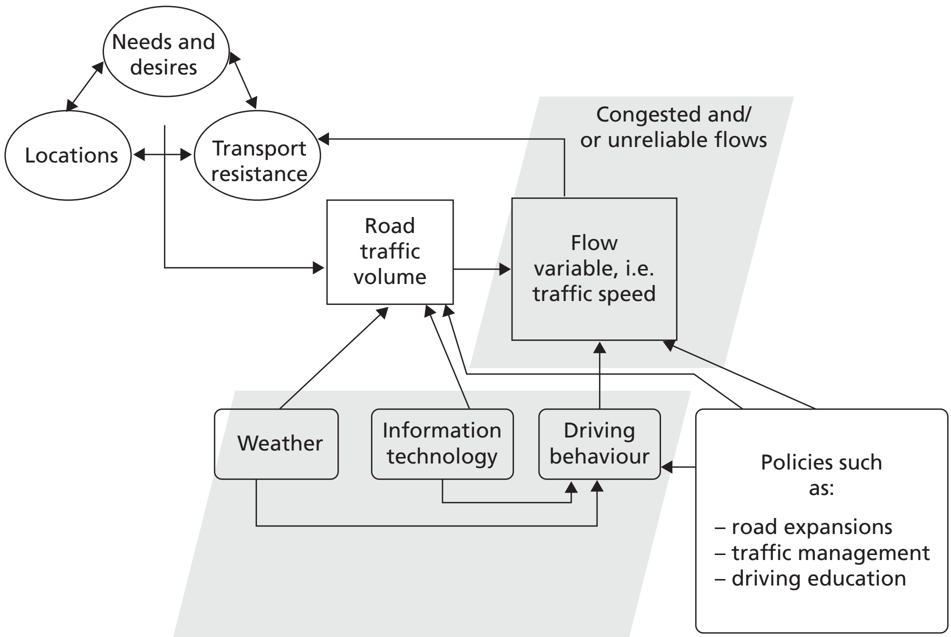


Figure 7.1 The connection between this chapter (grey area) and the simple conceptual framework (top left) as described in Chapter 2

To be clear, this chapter focuses on the road traffic flow variables and the interactions with aspects such as driving behaviour, weather, information technology and so forth (the grey areas in Figure 7.1). Thus traffic flow operations on a road facility are explained for a given traffic demand profile. Factors such as weather and information technology (e.g. navigation systems) can influence traffic flow characteristics through driving behaviour. Additionally, policies such as road expansions and traffic management measures can have an impact on traffic flow operations, either directly or indirectly, by influencing driving behaviour. Transport policies are discussed in Chapter 12.

Traffic flow theory entails the knowledge of the fundamental characteristics of traffic flows and the associated analytical methods. Examples of such characteristics are the road capacities, the relation between flow and density, and headway distributions. Examples of analytical methods are shockwave theory and microscopic simulation models.

Using the presented material, the reader will be able to interpret, analyse and – for simple situations – predict the main characteristics of traffic flows. For the greater part, the chapter considers traffic flow operations in simple infrastructure elements (uninterrupted traffic flow operations, simple discontinuities), although an important side step is made to look at network dynamics. In doing so, the chapter takes both a microscopic and a macroscopic perspective. The microscopic perspective reflects the behaviour of individual drivers interacting with surrounding vehicles, while the macroscopic perspective considers the average state of traffic. We discuss empirical facts, and some well-known analytical tools, such as shockwave theory, kinematic wave models and microscopic simulation models.

Section 7.2 introduces the basic variables on the microscopic level (the vehicle level), and section 7.3 the macroscopic variables, that is, the flow level. Section 7.4 discusses flow characteristics. Then, in section 7.5, traffic flow dynamics and the (self-)organization of traffic are discussed. Section 7.6 presents several theories on multi-lane traffic (i.e. motorways). Section 7.7 discusses car-following models, that is, microscopic flow models, and section 7.8 discusses the macroscopic flow models. Section 7.9 adds the dynamics of networks to this. Finally, in section 7.10 the conclusions are presented.

7.2 Vehicle trajectories and microscopic flow variables

The vehicle trajectory (often denoted as $x_i(t)$) of a vehicle (i) describes the position of the vehicle over time (t) along the roadway. The trajectory is the core variable in traffic flow theory which allows us to determine all relevant microscopic and macroscopic traffic flow quantities. Note that, for the sake of simplicity, the lateral component of the trajectory is not considered here.

To illustrate the versatility of trajectories, Figure 7.2 shows several vehicle trajectories. From the figure, it is easy to determine the distance headway S_i, and the time headway h_i, overtaking events (crossing trajectories), the speed $v_i = dx_i/dt$, the size of the acceleration (see top left where one vehicle accelerates to overtake another vehicle), the travel time TT_i and so forth.

However, although the situation is rapidly changing owing to so-called floating car data becoming more common, trajectory information is seldom available. Floating car data is information from mobile phones in vehicles that are being driven. In most cases, vehicle trajectory measurements only contain information about average characteristics of the traffic flow, provide only local information, or aggregate information in some other way (e.g. travel times from automatic vehicle identification or licence plate cameras).

Most commonly, traffic is measured by (inductive) loops measuring local (or time-mean) traffic flow quantities, such as (local) traffic flow q and local

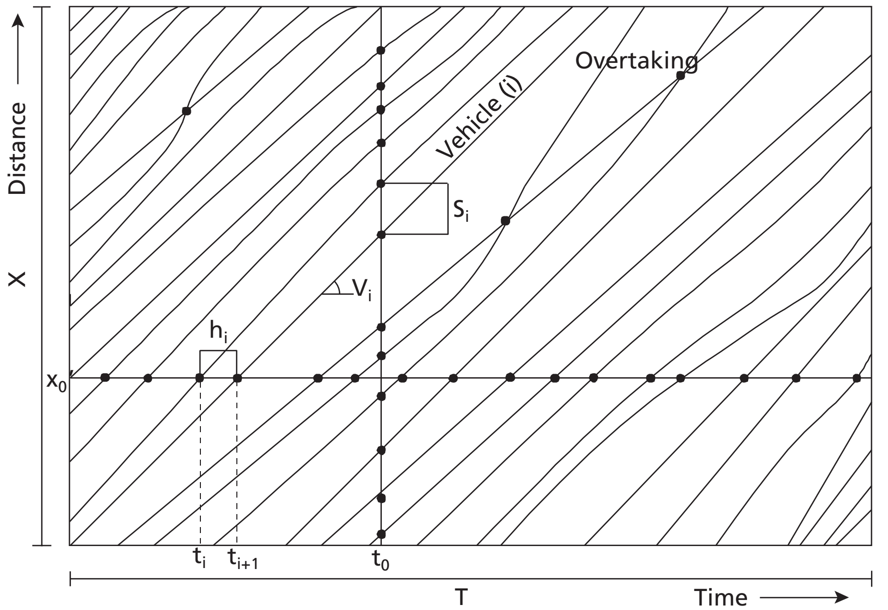


Figure 7.2 Vehicle trajectories and key microscopic flow characteristics

mean speed *u*. First, we will discuss the main microscopic traffic flow variables in detail. This type of flow variable reflects the behaviour of individual drivers interacting with surrounding vehicles.

Gross and net headways

The (gross) time headway (h) is one of the most important microscopic flow variables. It describes the difference between passage times t_i at a cross-section x of the rear bumpers of two successive vehicles:

$$h_i(x) = t_i(x) - t_{i-1}(x) \tag{1}$$

The time headway, or simply headway, is directly determined by the behaviour of the driver, vehicle characteristics, flow conditions and so on. Its importance stems from the fact that the (minimal) headways directly determine the capacity of a road, a roundabout and so forth. Typically, these minimal headways are around 1.5 seconds in dry conditions. Time headways, combined with the speeds, lead to the distance headways (see 'Gross and net distance headways' below).

The net time headway or gap is defined by the difference in passage times between the rear bumper of the lead vehicle and the front bumper of the following vehicle. This value is particularly important for driving behaviour analysis, for instance when analysing and modelling the amount of space drivers need to perform an overtaking manoeuvre (critical gap analysis).

Gross and net distance headways

We have seen in the preceding sub-section that time headways are *local* microscopic variables: they relate to the behaviour of an individual driver and are measured at a cross-section. On the contrary, distance headways (often denoted by the symbol s) are *instantaneous* (measured at one moment in time) microscopic variables, measuring the distance between the rear bumper of the leader and the rear bumper of the follower at time instant t:

$$s_i(t) = x_{i-1}(t) - x_i(t) \tag{2}$$

In congested conditions, distance headways are determined by the behaviour of drivers, which in turn depends on the traffic conditions, driver abilities, vehicle characteristics, weather conditions and so forth. In free flow with no interaction between the drivers, the headways are determined largely by the demand (that is, they are determined by the moments when drivers enter the freeway).

Net distance headways are defined, similarly to the net time headways, as the distance between the position of the rear bumper of the leader and the front bumper of the follower.

It should be clear that the time headways and the distance headways are strongly correlated. If v_{i-1} denotes the speed of the leading vehicle, it is easy to see that:

$$s_i = v_{i-1}h_i \tag{3}$$

7.3 Macroscopic flow variables

So far, we have mainly looked at microscopic traffic flow variables. Macroscopic flow variables, such as flow, density, speed and speed variance, reflect the average state of the traffic flow in contrast to the microscopic traffic flow variables, which focus on individual drivers. Let us take a closer look at the most important variables.

Traditional definitions of flow, density and speed

In general, the flow q (also referred to as intensity or volume) is traditionally defined by the 'average number of vehicles (n) that pass a cross-section during a unit of time (T)'. According to this definition, flow is a local variable (since it is defined at a cross-section). We have:

$$q = \frac{n}{T} = \frac{n}{\sum_{i=1}^{n} h_i} = \frac{1}{\bar{h}} \tag{4}$$

This expression shows that the flow can be computed easily by taking the number of vehicles n that have passed the measurement location during a period of length T. The expression also shows how the flow q relates to the average headway $\bar{h}$, thereby relating the macroscopic flow variable to average microscopic behaviour (i.e. time headways).

In a similar way, the density k (or concentration) is defined by the 'number of vehicles per distance unit'. Density is, therefore, a so-called instantaneous variable (i.e. it is computed at a time instance), defined as follows:

$$k = \frac{m}{X} = \frac{m}{\sum_{i=1}^{m} s_i} = \frac{1}{\bar{s}} \tag{5}$$

This expression shows that the density can be computed by taking a snapshot of a roadway segment of length X, and counting the number of vehicles m that occupy the road at that time instant. The expression also shows how density relates to average microscopic behaviour (i.e. distance headways, s). Note that, contrary to the flow, which can generally be easily determined in practice by using cross-sectional measurement equipment (such as inductive loops), the density is not so easily determined, since it requires observations of the entire road at a time instant (e.g. via an aerial photograph).

Similarly to the definitions above, average speeds u can be computed in two ways: at a cross-section (local mean speed or time-mean speed u_L), or at a time instant (instantaneous mean speed or space-mean speed u_M). As will be shown in the following sub-section, the difference between these definitions can be very large. Surprisingly, in practice the difference is seldom determined. For instance, the Dutch motorway monitoring systems collect time-mean speeds, while for most applications (e.g. average travel time) the space-mean speeds are more suitable.

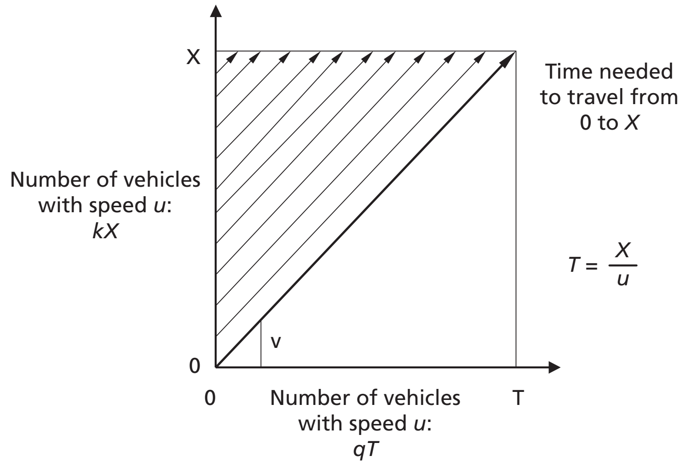


Figure 7.3 Derivation of the continuity equation

Continuity equation

An important relation in traffic flow theory is the continuity equation: $q = ku$ (flow equals density times the speed). This equation is used to relate the instantaneous characteristic density to the local characteristic flow. The derivation of this equation is actually quite straightforward (Figure 7.3).

Consider a road of length X. All vehicles on this road drive at an equal speed u. Let us define the period T by $T = X/u$. Under this assumption, it is easy to see that the number of vehicles that are on the road at time $t = 0$ – which is equal to the density k times the length X of the roadway segment – is equal to the number of vehicles that will pass the exit at $x = X$ during period $[0,T]$, which is in turn equal to the flow q times the duration of the period T. That is:

$$kX = qT \Leftrightarrow q = k\frac{X}{T} = ku \tag{6}$$

Clearly, the continuity equation holds when the speeds are constant. The question is whether the equation $q = ku$ can also be applied when the speeds are not constant (e.g. u represents an average speed) and, if so, which average speed (time-mean or space-mean speed) is to be used. It turns out that $q = ku$ can indeed be applied, but only if $u = u_M$, that is, if we take the space-mean speed.

Intuitively, one can understand this as follows (mathematical proof can be found in May, 1990). A detector lies at location x_{det}. Now we reconstruct which vehicles will pass in the time of one aggregation period. For this,

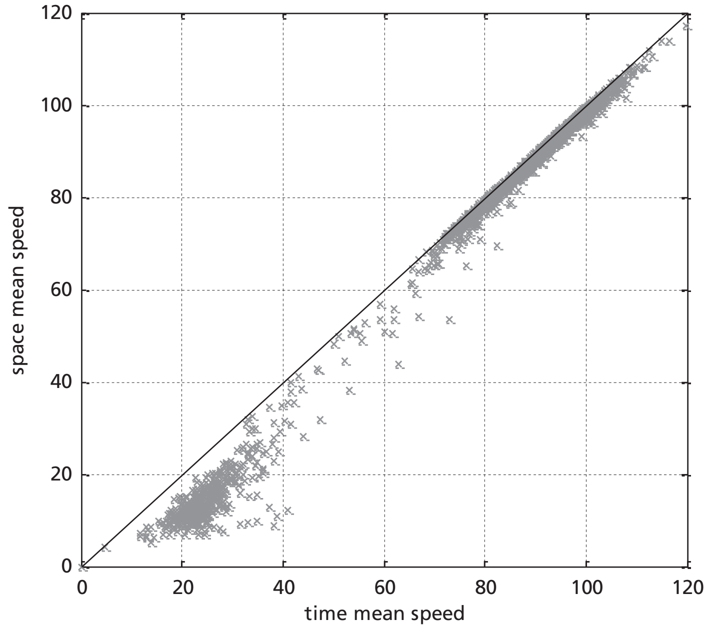


Figure 7.4 Differences between time-mean speed and space-mean speed for the A9 motorway

the vehicle must be closer to the detector than the distance it travels in the aggregation time t_{agg}:

$$x_{\mathrm{det}} - x_j \leq t_{\mathrm{agg}} v_i \tag{7}$$

In this formula, x is the position on the road. For faster vehicles, this distance is larger. Therefore, if one takes the local arithmetic mean, one overestimates the influence of the faster vehicles. If the influence of the faster vehicles on speeds is overestimated, the average speed u is overestimated (compared to the space-mean speed u_m).

The discussion above might be conceived as academic. However, if we look at empirical data, then the differences between the time-mean speeds and space-mean speeds become apparent. Figure 7.4 shows an example where the time-mean speed and space-mean speed have been computed from motorway individual vehicle data collected on the A9 motorway near Amsterdam, the Netherlands. Figure 7.4 clearly shows that the differences between the speeds can be as high as 100 per cent. Also note that the space-mean speeds are always lower than the time-mean speeds. Since, in most countries where inductive loops are used to monitor traffic flow operations, arithmetic mean speeds are computed and stored, average speeds are generally overestimated, affecting travel time estimations. Furthermore, since $q = ku$ can only be used for space-mean speeds, we cannot determine the density k from the local speed and flow measurements, complicating the use of the collected data, for example for traffic information and traffic management purposes.

Generalized traffic flow variables

Alternative measurement methods, such as automatic vehicle identification (AVI), radar and floating car measurements, provide new ways to determine the flow variables described above. One of the benefits of these new methods is that they provide information about the temporal and spatial aspects of traffic flow. For instance, using video we can observe the density in a region directly, rather than by determining the density from local observations.

For the relation between instantaneous and local variables, the work of Edie (1965) is very relevant. Edie (1965) introduces generalized definitions of flow, density and speed. These apply to regions in time and space, and will turn out to be increasingly important with the advent of new measurement techniques.

Consider a rectangular region in time and space with dimensions T and X respectively (see Figure 7.5). Let d_i denote the total distance travelled by vehicle i during period T and let r_i denote the total time spent in region X. Let us define the total distance travelled by all vehicles by:

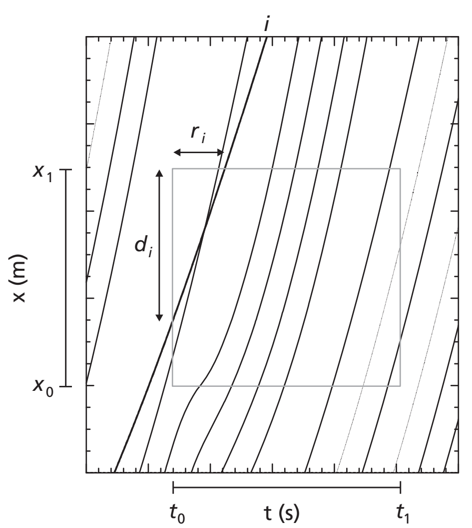


Figure 7.5 Generalization according to Edie (1965)

$$P = \sum_i d_i \tag{8}$$

Based on this quantity P, which is referred to as the *performance*, Edie defined the generalized flow as follows:

$$q = \frac{P}{XT} \tag{9}$$

Note that we can rewrite this equation as follows:

$$q = \frac{\sum_i d_i / X}{T} \tag{10}$$

Let us now define the total travel time R as follows:

$$R = \sum_i r_i \tag{11}$$

Edie defines the generalized density by:

$$k = \frac{R}{XT} = \frac{\sum_i r_i / T}{X} \tag{12}$$

For the generalized speed, the following intuitive definition is used:

$$u = \frac{q}{k} = \frac{P}{R} = \frac{\text{total distance travelled}}{\text{total time spent}} \tag{13}$$

These definitions can be used for any regions in space–time, even non-rectangular ones.

7.4 Microscopic and macroscopic flow characteristics

The preceding sections introduced the different microscopic and macroscopic variables. This section shows the most common flow characteristics, entailing both relations between the flow variables, or typical distribution, and so on. These flow characteristics, in a sense, drive the traffic flow dynamics that will be discussed below. As well as providing a short description of the characteristics and their definition, this section will discuss empirical examples as well as key issues in identifying these parameters.

Headway distributions

If we were to collect headways at a specific location x, then we would observe that these headways are not constant but rather follow some probability distribution function. This is also the case when the flow is stationary during the data collection period. The causes are manifold: there are large differences in driving behaviour between different drivers and differences in the vehicle characteristics, but there is also variation within the behaviour of one driver. A direct and important consequence of this is that the capacity of the road, which is by and large determined by the driving behaviour, is not constant either, but a stochastic variable.

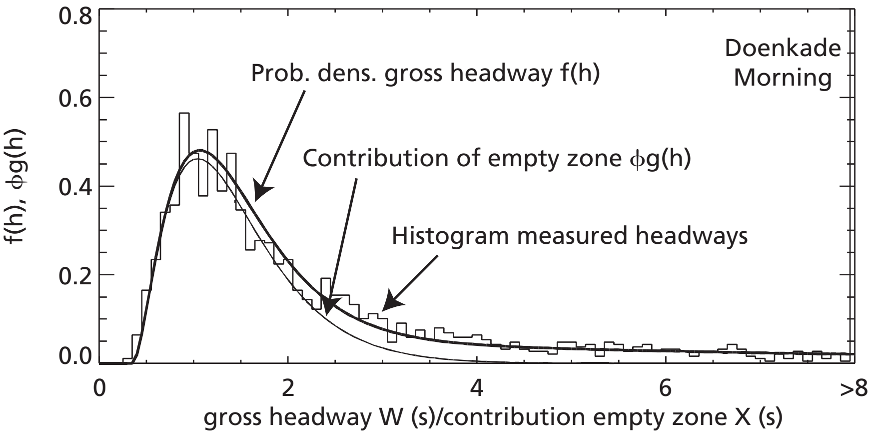
The headway distribution can be described by a probability density function $f(h)$. In the literature, many different kinds of distribution functions have been proposed, with varying success. It can be shown that if the flows are small – there are few vehicle interactions – the exponential distribution will be an adequate model. When the flows become larger, there are more interactions amongst the vehicles, and other distributions are more suitable. A good candidate in many situations is the log-normal distribution; we refer to Cowan (1975) for more details. In Hoogendoorn (2005), an overview is given of estimation techniques for the log-normal distributions in specific situations.

The main problem with these relatively simple models is that they are only able to represent available measurements but cannot be extrapolated to other situations. If, for instance, we are interested in a headway distribution for another flow level than the one observed, we need to collect new data and re-estimate the model.

To overcome this, so-called composite headway models have been proposed. The main characteristic of these models is that they distinguish between vehicles that are flowing freely and those that are constrained by the vehicle in front. Buckley (1968) was one of the first proposing these models, assuming that the headways of the free driving vehicles are exponentially distributed. He showed that the probability density function $f(h)$ of the observed headways h can be described by the following function:

$$f(h) = \phi g(h) + (1 - \phi) w(h) \tag{14}$$

In this equation, g describes the probability density function of the headways of vehicles that are following (also referred to as the distribution of the empty zones), while w denotes the probability density function of those vehicles



Note: W denotes the gross headway, which is composed of the empty zone X and the free headway W-X.

Source: Hoogendoorn and Botma (1997).

Figure 7.6 Composite headway probability density function (p.d.f.) applied to data on Doenkade site

that are driving freely. For the latter, an exponential distribution is assumed, and ϕ denotes the fraction of vehicles which are following.

There are different ways to estimate these probability density functions from available headway observations. Wasielewski (1974), later improved by Hoogendoorn (2005), proposed an approach in which one does not need to choose a prior form of the constrained headway distribution. To illustrate this, Figure 7.6 shows an example of the application of this estimation method on a two-lane motorway in the Netherlands in the morning (the location is the Doenkade).

This example nicely illustrates how the approach can be applied for estimating capacities, even if no capacity observations are available. We find the maximum flow (or capacity flow, C) when all drivers are following (as opposed to driving freely). We directly observe from the Buckley model (Buckley, 1968; see above) that the observed headways in that case follow g. The number of vehicles per hour equals 3600 seconds per hour divided by the average headway (H, following distribution g) in seconds (or the expectation value thereof, indicated by E). We therefore get:

$$C = 3600/E(H) \text{ where } H \sim g \tag{15}$$

In other words, the capacity flow equals one over the mean (minimum) headway value, on the condition that all vehicles are following. Using this approach, we can find estimates for the capacity even if there are no direct capacity observations available. For the example above, we can compute the mean empty zone value by looking at the p.d.f. $g(h)$, which turns out to be equal to 1.69. Based on this value, we find a capacity estimate of $3600/1.69 = 2134$ vehicles per hour.

Desired speed distributions

Generally, the free speed or desired speed of a driver–vehicle combination (hereafter, simply called vehicle or driver) is defined by the speed driven when other road users do not influence the driver. Knowledge of free speeds on a road under given conditions are relevant for a number of reasons. For instance, the concept of free speed is an important element in many traffic flow models. As an illustration, the free speed distribution is an important input for many microscopic simulation models. Insights into free speeds and their distributions are also important from the viewpoint of road design and for determining suitable traffic rules for a certain facility. For instance, elements of the network should be designed so that drivers using the facility can traverse the road safely and comfortably. It is also of interest to see how desired speed distributions change under varying road, weather and ambient conditions and how these distributions vary for different types of travellers. So speed distribution is also an important characteristic amongst drivers for design issues.

The free speed will be influenced by the characteristics of the vehicle, the driver, the road and (road) conditions such as weather and traffic rules (speed limits). Botma (1999) describes how individual drivers choose their free speed, discussing a behavioural model relating the free speed of a driver to a number of counteracting mental stresses a driver is subjected to. A similar model can be found in Jepsen (1998). However, these models have not been successful in their practical application. The problem of determining free speed distributions from available data is not trivial. In Botma (1999), an overview of alternative free speed estimation approaches is presented. Botma (1999) concluded that all the methods he reviewed have severe disadvantages, which is the reason why another estimation approach is proposed. This approach is based on the concept of censored observations (Nelson, 1982) using a parametric estimation approach to estimate the parameters of the free speed distribution. Speed observations are marked as either censored (constrained) or uncensored (free flowing) using subjective criteria (headway and relative speed). Hoogendoorn (2005) presents a new approach to estimating the distribution of free speeds based on the method of censored observations.

Gap acceptance and critical gaps

Gap acceptance is a process that occurs in different traffic situations, such as crossing a road, entering a roundabout or performing an overtaking manoeuvre on a bi-directional road. The minimum gap that a driver will accept is generally called the critical gap. Mathematical representations of the gap acceptance process are an important part of traffic simulation models, for instance.

In general terms the gap acceptance process can be described as follows: traffic participants who want to make a manoeuvre estimate the space they need and estimate the available space. Based on the comparison between required and available space, they decide to start the manoeuvre or to postpone it. The term 'space' is deliberately somewhat vague; it can be expressed either in time or in distance. The required space is dependent on characteristics of the traffic participant, the vehicle and the road. The available space is dependent on the characteristics of, for instance, the on-coming vehicles and the vehicle to be overtaken (the passive vehicle). Traffic participants have to perceive all these characteristics, process them and come to a decision. Humans differ highly in perception capabilities; for example, the ability to estimate distances can vary substantially between persons, and they differ in the acceptance of risk. The total acceptance process is dependent on many factors, of which only a subset is observable. This has led to the introduction of stochastic models.

Many different methods to estimate the distribution of critical gaps by observing the gap acceptance process in reality can be found in the literature (Brilon et al., 1999). Let us consider the problem of estimating the critical gap distribution. Suppose, as an example, a driver successively rejects gaps of 3, 9, 12 and 7 *s* and accepts a gap of 19 *s*. The only thing one can conclude from these observations is that this driver has a critical gap between 12 and 19 *s*. Stated in other words, the critical gap cannot be observed directly. The observations are, thus, censored. Note that it can also be concluded that only the maximum of the rejected gaps is informative for the critical gap (assuming that the driver's behaviour is consistent); the smaller gaps are rejected by definition.

Capacity and capacity estimation

Capacity is usually defined as follows: 'The maximum hourly rate at which people or vehicles can reasonably be expected to traverse a point or uniform section of a lane or roadway during a given time period (usually 15 minutes) under prevailing roadway, traffic and control conditions.'

Maximum flows (maximum free flows of queue discharge rates) are not constant values and vary under the influence of several factors. Factors influencing the capacity are, among other things, the composition of the vehicle fleet, the composition of traffic with respect to trip purpose, weather, road, ambient conditions and so on. These factors affect the behaviour of driver–vehicle combinations and thus the maximum number of vehicles that can pass a cross-section during a given time period. Some of these factors can be observed and their effect can be quantified. Some factors, however, cannot be observed directly. Furthermore, differences exist between drivers, implying that some drivers will need a larger minimum time headway than other drivers, even if drivers belong to the same class of users. As a result, the minimum headways will not be constant values but follow a distribution function (see the discussion on headway distribution modelling in 'Headway distributions' above). Observed maximum flows thus appear to follow a distribution. The shape of this distribution depends, among other things, on the capacity definition and measurement method or period. In most cases, a normal distribution can be used to describe the capacity.

Several researchers have pointed out the existence of two different maximum flow rates, namely pre-queue and queue discharge. Each of these has its own maximum flow distribution. We define the pre-queue maximum flow as the maximum flow rate observed at the downstream location just before the onset of congestion (a queue or traffic jam) upstream. These maximum flows are characterized by the absence of queues or congestion upstream of the bottleneck, high speeds and instability leading to congestion onset within a short period and maximum flows showing a large variance. The queue discharge flow is the maximum flow rate observed at the downstream location as long as congestion exists. These maximum flow rates are characterized by the presence of a queue upstream of the bottleneck, lower speeds and densities, and a constant outflow with a small variance which can be sustained for a long period, but with lower flow rates than in the pre-queue flow state. Both capacities can only be measured downstream of the bottleneck location. Average capacity drop changes are in the range of −1 to −15 per cent, but −30 per cent changes are also reported (see section 7.5 for more on capacity drop changes).

There are many approaches that can be applied to compute the capacity of a specific piece of infrastructure. The suitability of the approach depends on a number of factors, such as:

1. type of infrastructure (e.g. motorway without on- or off-ramps, on-ramp, roundabout, unsignalized intersection, etc.);
2. type of data (individual vehicle data, aggregate data) and time aggregation;
3. location of data collection (upstream of, in or downstream of the bottleneck);
4. traffic conditions for which data are available (congestion, no congestion).

We refer to Minderhoud et al. (1996) for a critical review of approaches that are available to estimate road capacity.

Fundamental diagrams

The fundamental diagram describes a statistical relation between the macroscopic traffic flow variables of flow, density and speed. There are different ways to represent this relation, but the most often used is the relation $q = Q(k)$ between the flow and the density. Using the continuity equation, the other relations $u = U(k)$ and $u = U(q)$ can be easily derived.

To understand the origin of the fundamental diagram, we can interpret the relation from a driving behaviour perspective. To this end, recall that the flow and the density relate to the (average) time headway and distance headway according to Equations (5) and (6) respectively. Based on this, we can clearly see which premise underlies the existence of the fundamental diagram: *under similar traffic conditions, drivers will behave in a similar way*. That is, when traffic operates at a certain speed u, then it is plausible that (on average) drivers will maintain (on average) the same distance headway $s = 1/k$. This behaviour – and therewith the relation between speed and density – is obviously dependent on factors like weather, road characteristics, composition of traffic, traffic regulations and so forth.

Figure 7.7 shows typical examples of the relation between flow, density and speed. The figure shows the most important points in the fundamental diagram, which are the roadway capacity C, the critical density k_c and the critical speed u_c (the density and speed occurring at capacity operations), the jam density k_{jam} (density occurring at zero speed) and the free speed u_0. In

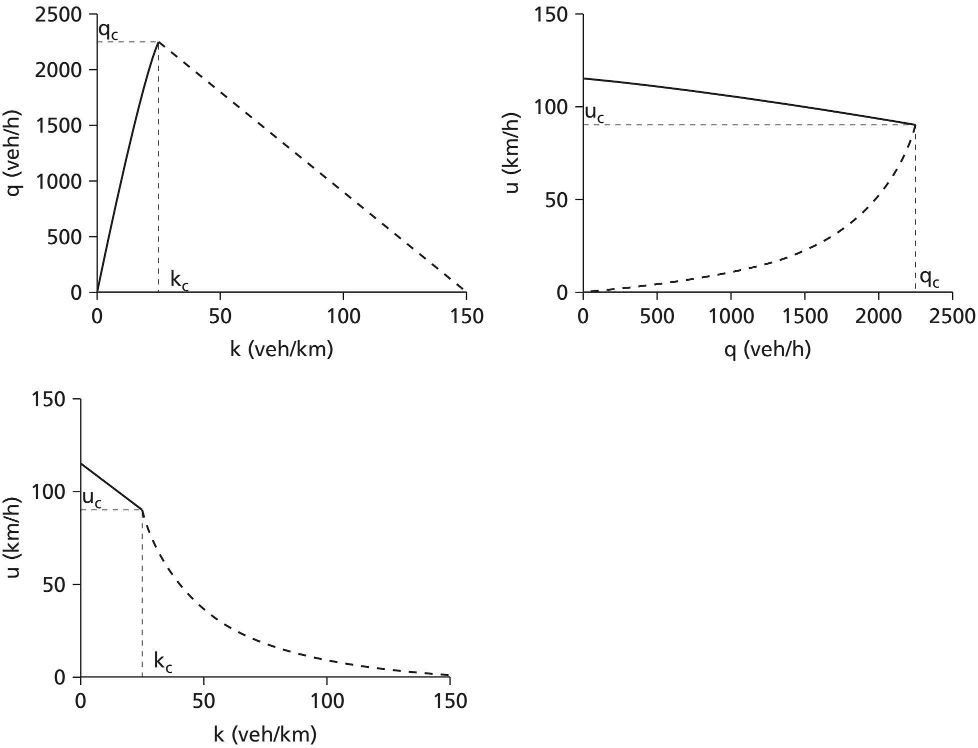


Figure 7.7 Example of the fundamental diagram

the figure, we clearly see the difference between the free conditions $(k < k_c)$ and the congested conditions $(k > k_c)$.

It is tempting to infer causality from the fundamental diagram: it is often stated that the relation $u = U(k)$ describes the fact that, with increasing density (e.g. reduced spacing between vehicles), the speed is reducing. It is, however, more the other way around. If we take a driving behaviour perspective, then it seems more reasonable to assume that, with reduced speed of the leader, drivers need smaller distance headways to drive safely and comfortably.

Fundamental diagrams are often determined from real-life traffic data. This is usually done by assuming that stationary periods can be identified during data measurements. To obtain meaningful fundamental diagrams, the data collection must be performed at the correct location during a selected time period.

7.5 Traffic flow dynamics and self-organization

So far, we have discussed the main microscopic and macroscopic characteristics of traffic flow. In doing so, we have focused on static characteristics of traffic flow. However, there are different characteristics, which are dynamic in nature or, rather, have to do with the dynamic properties of traffic flow.

Capacity drop

The first phenomenon that we discuss is the so-called capacity drop. The capacity drop describes the fact that, once congestion has formed, drivers are not maintaining a headway that is as close as it was before the speed breakdown. Therefore the road capacity is lower. This effect is considerable, and values of a reduction up to 30 per cent are quoted (Hall and Agyemang-Duah, 1991; Cassidy and Bertini, 1999; Chung et al., 2007). The effect of the capacity drop is illustrated in Figure 7.8.

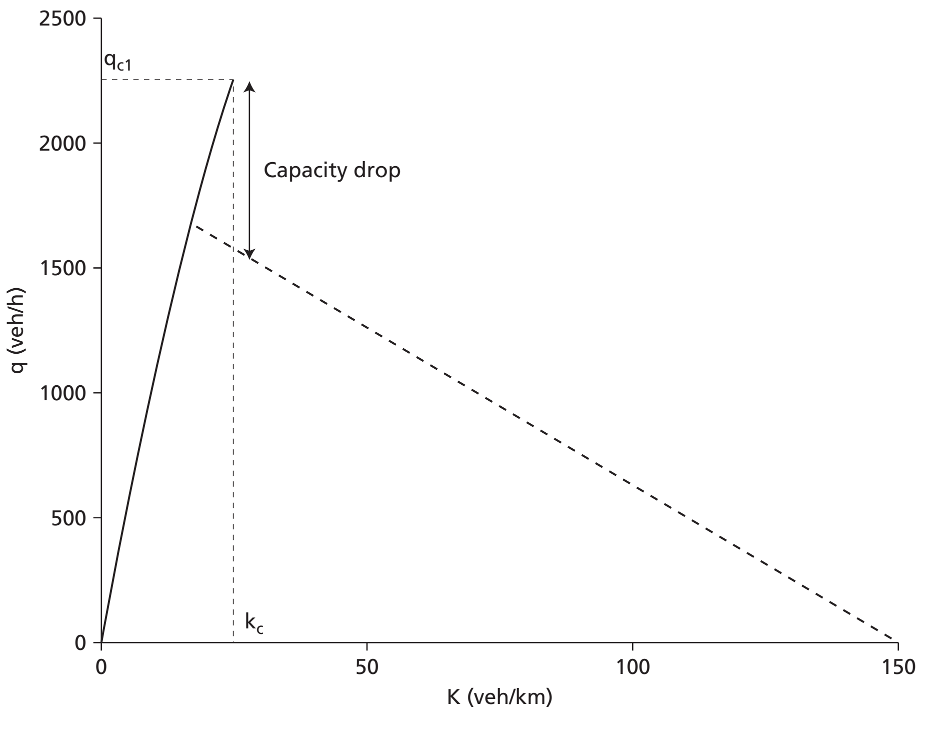


Figure 7.8 The capacity drop in the flow-density diagram

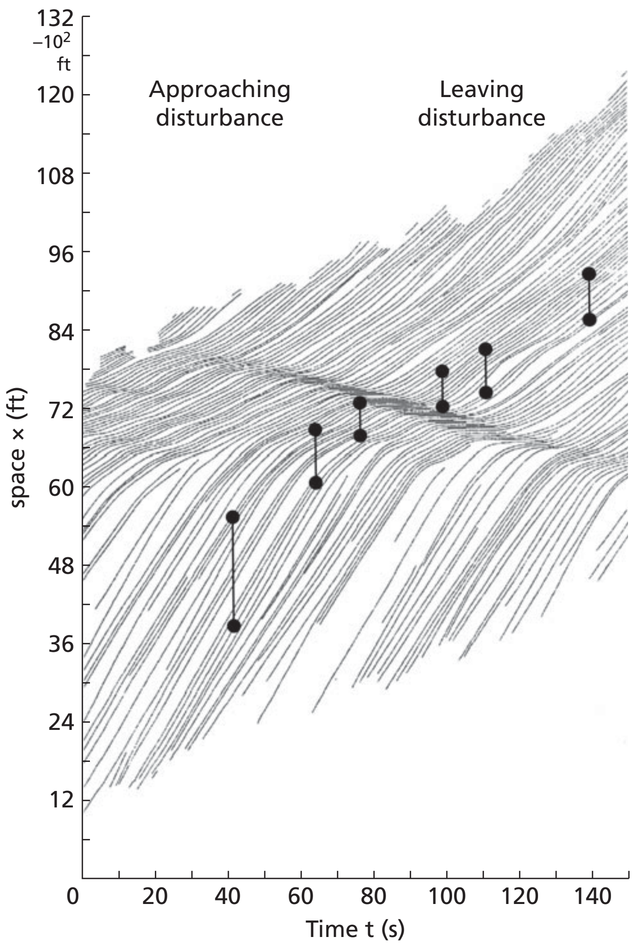
Traffic hysteresis

The different microscopic processes that constitute the characteristics of a traffic flow take time: a driver needs time to accelerate when the vehicle in front drives away when the traffic signal turns green. When traffic conditions in a certain location change, for instance when the head of a queue moves upstream, it will generally take time for the flow to adapt to these changing conditions.

Generally, however, we may assume that given that the conditions remain unchanged for a sufficient period of time – say, five minutes – traffic conditions will converge to an average state. This state is often referred to as the equilibrium traffic state. When considering a traffic flow, this equilibrium state is generally expressed in terms of the fundamental diagram. That is, when considering traffic flow under stationary conditions, the flow operations can – on average – be described by some relation between speed, density and flow. This is why the speed–density relation is often referred to as the equilibrium speed.

From real-life observations of traffic flow, it can be observed that many of the data points collected are not on the fundamental diagram. While some of these points can be explained by stochastic fluctuations (e.g. vehicles have different sizes, drivers have different desired speeds and following distances), some can be structural, and stem from the dynamic properties of traffic flow. That is, they reflect so-called transient states, that is, changes from congestion to free flow (acceleration phase) or from free flow to congestion (deceleration phase) in traffic flow. It turns out that generally these changes in the traffic state are not on the fundamental diagram. In other words, if we consider the average behaviour of drivers (assuming stationary traffic conditions), observed mean speeds will generally not be equal to the 'equilibrium' speed. The term 'equilibrium' reflects the fact that the observed speeds in time will converge to the equilibrium speed, assuming that the average conditions remain the same. That is, the average speed does not adapt instantaneously to the average or equilibrium speed.

This introduces traffic hysteresis, which means that for the same distance headway drivers choose a different speed during acceleration from that chosen during deceleration. Figure 7.9 shows the first empirical observation thereof by Treiterer and Myers (1974). The figure shows the time it takes for a platoon to pass a point along the roadway. The longer the arrow is, the longer that time is and hence the lower the flow (vehicles/hour). The arrow is long at the beginning, since some drivers are not car-following yet.



Source: Treiterer and Myers (1974).

Figure 7.9 Vehicle trajectories collected from an airborne platform clearly showing differences in average platoon length before and after the disturbance

At the second arrow, all vehicles are car-following and the flow is high (short arrow). In the disturbance, the flow is very low and we find a long arrow. After the disturbance, the flow increases but the headways are longer than before the vehicles entered the disturbance. Note also that, in exiting the traffic jam, all vehicles will be in car-following mode.

Three-phase traffic flows, phase transitions and self-organization

Amongst the many issues raised by Kerner (2004) is the fact that there are three phases (free flow, synchronized flows and jams) rather than two (free

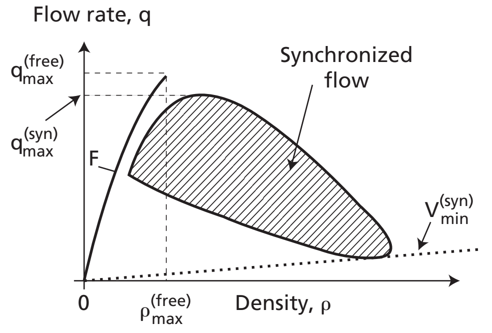


Figure 7.10 The fundamental diagram according to Kerner (2004)

flow and congestion). As in other theories, the first phase is the free flow phase. The second phase, in three-phase traffic flow theory, is the synchronized flow. In this phase, the speed between the lanes is more or less the same, rather than in free flow where the overtaking lane has a higher speed than the slow lane. Furthermore, Kerner claims that there are different equilibrium spacings (densities) for the same flow value, such that the phase should not be described by a line but by an area in the flow-density plane (Figure 7.10). The third phase, the wide moving jam, is identified by a (near) standstill of the vehicles. Owing to the very low flow, the queue will grow at the tail. At the same time, vehicles at the head of the queue can accelerate. This means that the queue moves in the opposite direction to the traffic; the wave speed is approximately 18 km/h propagating backwards from the driving direction.

As well as the distinguishing of the three phases, Kerner discusses transitions between the different traffic phases. Some of these transitions are induced ('forced'). An example is a phase transition induced by a bottleneck, such as an on-ramp. In this situation, the simple fact that traffic demand is at some point in time larger than the rest capacity (being the motorway capacity minus the inflow from the on-ramp) causes a transition from the free flow phase to the synchronized flow phase. Note that these kinds of phase transitions can be described by basic flow theories and models (shockwave theory, kinematic wave models) adequately. As an additional remark, note that these transitions are, although induced, still random events, since both the free flow capacity and the supply are random variables.

However, not all phase transitions are induced (directly); some are caused by intrinsic ('spontaneous') properties of traffic flow. An example is the spontaneous transition from synchronized flow to jammed flow (referred to by Kerner as wide moving jams). Owing to the unstable nature of specific

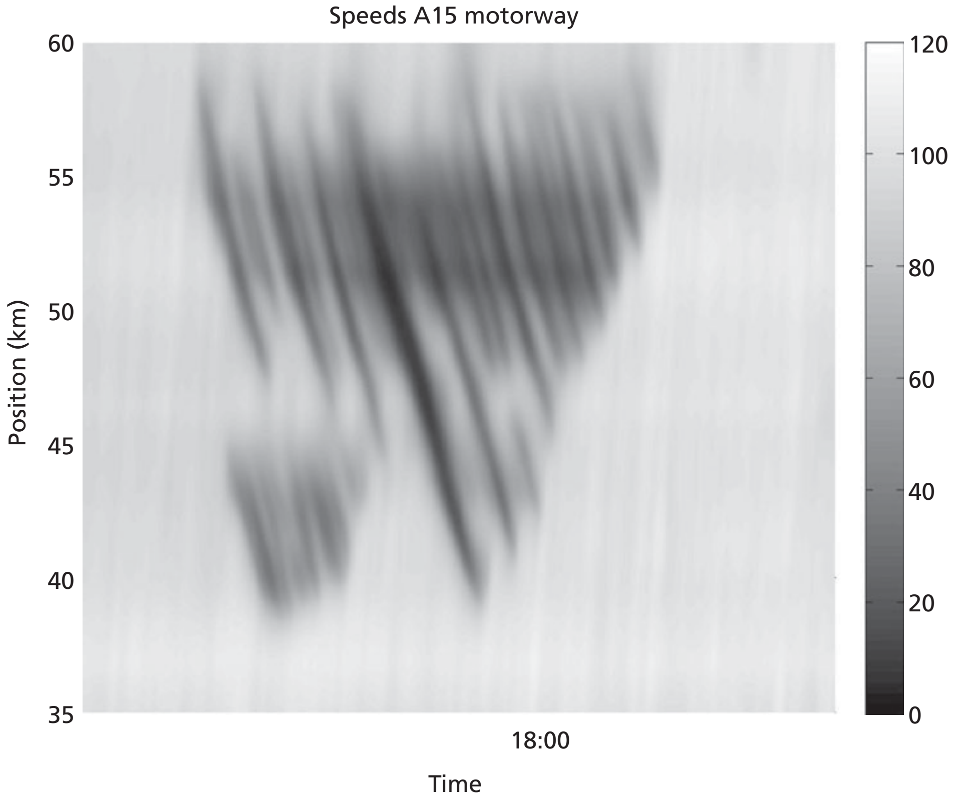


Figure 7.11 Typical traffic pattern on the A15 motorway in the Netherlands. A bottleneck can be determined at 55 km, and one can find wide moving jams propagating backwards at approximately 18 km/h

synchronized flow regimes, small disturbances in the congested flow will grow over time. For instance, a small localized high density cluster caused by a vehicle braking a bit too hard because the driver was temporarily distracted may grow because of vehicles moving at the back of the localized cluster subsequently needing to brake as well (and often doing so because of the finite reaction times of drivers). As a result, this upstream moving disturbance will gain in amplitude and will, in the end, become a wide moving jam.

This phenomenon is quite common in day-to-day motorway traffic operations. Figure 7.11 shows an example of the A15 motorway in the Netherlands. The picture clearly shows the frequent occurrence of these spontaneous transitions from synchronized to jammed flow, resulting in numerous upstream moving, wide moving jams. Note that, as wide moving jams have an outflow rate which is about 30 per cent lower than the free flow capacity, these jams are actually quite undesirable from a traffic efficiency perspective. Furthermore, they imply additional braking and acceleration, yielding increased fuel consumption and emission levels.

7.6 Multi-lane traffic flow facilities

Up to now, the chapter has considered each lane of the freeway to be equal. However, there are considerable differences between them. This section introduces just the basic concept. For a deeper insight, we refer to the literature mentioned in this chapter. For the sake of simplicity, we here assume driving on the right. For countries where a left-hand driving rule applies, like Japan, the United Kingdom or Australia, the lanes are exactly opposite. Daganzo (2002a, 2002b) poses a theory classifying traffic as slugs, defined by a low desired speed, and rabbits, defined by a high free flow speed. He states that, as soon as the speed in the right lane goes under a certain threshold, rabbits will move to faster lanes to the left. Furthermore, the theory states that, even if the density in the right lane is lower than in the left lane, the rabbits will not change to the right lane as long as the speeds in the left lane are higher. This traffic state, with two different speeds, is called a *two pipe regime*, since traffic is flowing as it were in two different, unrelated pipes. In this state, there is no equal density in both lanes. Only when the density in the left lane increases so much that the speed decreases to a value lower than the speed in the right lane will the rabbits move towards the right lane. Then the rabbits will redistribute themselves in such a way that the traffic in both lanes flows at the same speed. This is called a *one pipe regime*.

Note that the speeds in different lanes at the same densities can be different, owing to these effects or, basically, owing to the driver population in that lane. This leads to different fundamental diagrams in the left and right lanes. Usually, the free flow speed in the left lane is higher than in the right lane, owing to the higher fraction of rabbits in that lane. Kerner (2004) poses a similar theory on multi-lane traffic flow facilities.

7.7 Traffic flow models

Traffic flow models can be used to simulate traffic, for instance to evaluate ex ante the use of a new part of the infrastructure. The models can be helpful tools in answering the questions posed in the introduction to this chapter, such as: When do traffic jams emerge? How will they propagate in space and time? And how long does it takes for the congestion to resolve? Additionally, the models can be used to improve road safety.

Table 7.1 Overview of traffic flow model classification

Representation	Behavioural rules	
	Microscopic	Macroscopic
Vehicle based	Microscopic flow models	Particle models
Flow based	Gas-kinetic models	Macroscopic models

Traffic flow models may be categorized using various dimensions (deterministic or stochastic, continuous or discrete, analytical or simulation, and so forth). The most common classification is the distinction between microscopic and macroscopic traffic flow modelling approaches. However, this distinction is not unambiguous, owing to the existence of hybrid models. This is why models are categorized here based on the following aspects:

1. **representation** of the traffic flow in terms of flows (macroscopic), groups of drivers (macroscopic) or individual drivers (microscopic);
2. **underlying behavioural theory**, which can be based on characteristics of the flow (macroscopic) or individual drivers (microscopic behaviour).

The remainder of this section uses this classification to discuss some important flow models. Table 7.1 depicts an overview of these models.

The observed behaviour of drivers, that is, headways, driving speeds and driving lane, is influenced by different factors, which can be related to the driver–vehicle combination (vehicle characteristics, driver experience, age, gender and so forth), the traffic conditions (average speeds, densities), infrastructure conditions (road conditions) and external situational influences (weather, driving regulations). Over the years, different theories have been proposed to (dynamically) relate the observed driving behaviour to the parameters describing these conditions.

In the process, different driver sub-tasks are often distinguished. Table 7.2 provides a rough but useful classification of these tasks (Minderhoud, 1999). In general, two types of driver tasks are distinguished: longitudinal tasks (acceleration, maintaining speed, maintaining distance relative to the leading vehicle) and lateral tasks (lane changing, overtaking). In particular the longitudinal and (to a lesser extent) the lateral interaction sub-tasks have received quite a lot of attention in traffic flow theory research.

Table 7.2 Driving sub-tasks overview

	Longitudinal	Lateral
Infrastructure	Free speed	Course keeping
Interaction	Car-following	Mandatory and discretionary lane changing

Source: Minderhoud (1999).

A microscopic model provides a description of the movements of individual vehicles that are considered to be a result of the characteristics of drivers and vehicles, the interactions between driver–vehicle elements, the interactions between driver–vehicle elements and the road characteristics, external conditions and the traffic regulations and control. Most microscopic simulation models assume that a driver will only respond to the one vehicle that is driving in the same lane directly in front of him (the leader).

When the number of driver–vehicle units on the road is very small, the driver can freely choose his speed given his preferences and abilities, the roadway conditions, curvature, prevailing speed limits and so forth. In any case, there will be little reason for the driver to adapt his speed to the other road users. The target speed of the driver is the so-called free speed. In real life, the free speed will vary from one driver to another, but the free speed of a single driver will also change over time. Most microscopic models assume however that the free speeds have a constant value that is driver-specific. When traffic conditions deteriorate, drivers will no longer be able to choose the speed freely, since they will not always be able to overtake or pass a slower vehicle. The driver will need to adapt his speed to the prevailing traffic conditions, that is, the driver is following. In the rest of this section, we will discuss some of these car-following models. Models for the lateral tasks, such as deciding to perform a lane change and gap acceptance, will not be discussed in this section in detail. Ahmed et al. (1996) provide a concise framework of lane changing modelling.

Safe-distance models

The first car-following models were developed by Pipes (1953) and were based on the assumption that drivers maintain a safe distance. A good rule for following vehicle $i-1$ at a safe distance s_i is to allow at least the length S_0 of a car between vehicle i and a part which is linear with the speed v_i at which i is travelling:

$$s_i = S(v_i) = S_0 + T_r v_i \tag{16}$$

Here, S_0 is the effective length of a stopped vehicle (including additional distance in front), and T_r denotes a parameter (comparable to the reaction time). A similar approach was proposed by Forbes et al. (1958). Both Pipes's and Forbes's theories were compared to field measurements. It was concluded that, according to Pipes's theory, the minimum headways are slightly less at low and high velocities than observed in empirical data. However, considering the models' simplicity, the way they were in line with real-life observations was amazing (see Pignataro, 1973).

Stimulus response models

However, safe-distance models do not seem to capture much of the phenomena observed in real-life traffic flows, such as hysteresis, traffic instabilities and so on. Stimulus response models are dynamic models that describe more realistically the reaction of drivers to things like changes in distance, speeds and so on relative to the vehicle in front, by considering a finite reaction time, for example. These models are applicable to relatively busy traffic flows where the overtaking possibilities are small and drivers are obliged to follow the vehicle in front of them. Drivers do not want the gap in front of them to become too large so that other drivers can enter it. At the same time, the drivers will generally be inclined to keep a safe distance.

Stimulus response models assume that drivers control their acceleration (a). The well-known model of Chandler et al. (1958) is based on the intuitive hypothesis that a driver's acceleration is proportional to the relative speed $v_{i-1} - v_i$:

$$a_i(t) = \frac{d}{dt}v_i(t) = \alpha(v_{i-1}(t - T_r) - v_i(t - T_r)) \tag{17}$$

where T_r again denotes the overall reaction time, and α denotes the sensitivity. Based on field experiments, conducted to quantify the parameter values for the reaction time T_r and the sensitivity α, it was concluded that α depended on the distance between the vehicles: when the vehicles were close together, the sensitivity was high, and vice versa.

Stimulus response models have been applied mainly to single lane traffic (e.g. tunnels; see Newell, 1961) and traffic stability analysis (Herman, 1959; May, 1990). It should be noted that no generally applicable set of parameter estimates has been found so far, that is, estimates are site-specific. An overview of parameter estimates can be found in Brackstone and McDonald (1999).

Psycho-spacing models

The two car-following models discussed so far have a mechanistic character. The only human element is the presence of a finite reaction time T_r. However, in reality a driver is not able to:

1. observe a stimulus lower than a given value (perception threshold);
2. evaluate a situation and determine the required response precisely, for instance because of observation errors resulting from radial motion observation;
3. manipulate the acceleration and brake pedals precisely.

Furthermore, owing to the need to distribute his attention between different tasks, a driver will generally not be permanently occupied with the car-following task. This type of consideration has inspired a different class of car-following models, namely the psycho-spacing models. Michaels (1963) provided the basis for the first psycho-spacing, based on theories borrowed from perceptual psychology (see Leutzbach and Wiedemann, 1986).

The so-called action point models (an important psycho-spacing model) form the basis for a large number of contemporary microscopic traffic flow models. Brackstone and McDonald (1999) conclude that it is hard to come to a definitive conclusion on the validity of these models, mainly because the calibration of its elements has not been successful.

7.8 Macroscopic traffic flow models

In the previous section we have discussed different microscopic traffic flow modelling approaches. In this section, we will discuss the main approaches that have been proposed in the literature taking a macroscopic perspective.

Deterministic and stochastic queuing theory

The most straightforward approach to model traffic dynamics is probably the use of queuing theory. In queuing theory we keep track of the number of vehicles in a queue (n). A queue starts whenever the flow to a bottleneck is larger than the bottleneck capacity, where the cars form a virtual queue. The outflow of the queue is given by the infrastructure (it is the outflow capacity

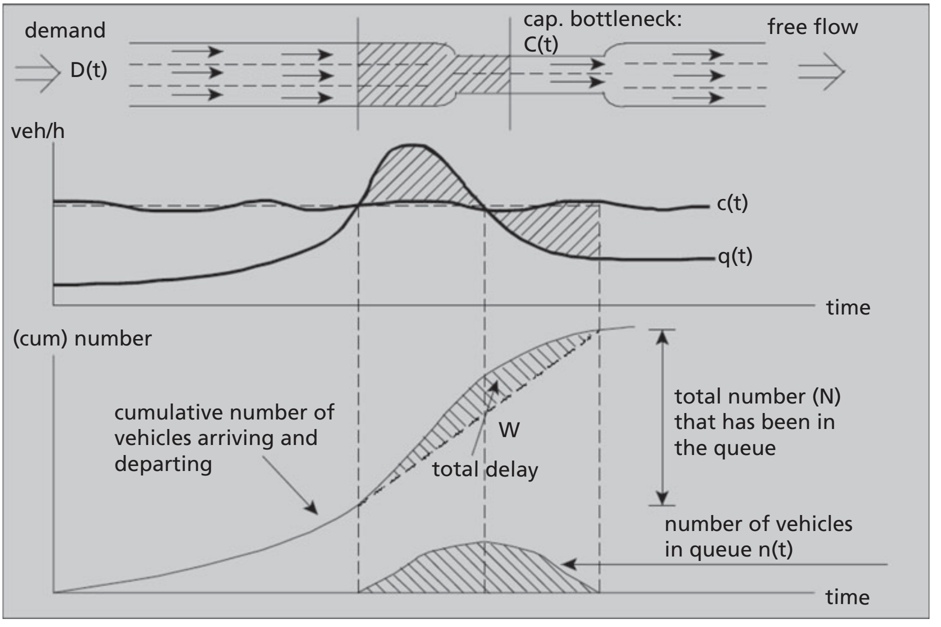


Figure 7.12 Functioning of queuing theory

of the bottleneck, given by C), whereas the inflow is the flow towards the bottleneck (q) as given by the traffic model. In an equation, this is written as:

$$dn = q(t)\,dt - C(t)\,dt \tag{18}$$

The number of vehicles in the queue (n; dn stands for the change in the number of vehicles in the queue) will evolve in this way until the queue has completely disappeared. Note that both the inflow and the capacity are time dependent in the description. For the inflow, this is due to the random distribution pattern of the arrival of the vehicles. Vehicles can arrive in platoons or there can be large gaps in between two vehicles. The capacity is also fluctuating. On the one hand, there are vehicle-to-vehicle fluctuations. For instance, some drivers have a shorter reaction time, hence a shorter headway leading to a higher capacity. On the other hand, on a larger scale, the capacities will also depend on road or weather conditions (e.g. wet roads, night-time).

Figure 7.12 shows how the number of vehicles in the queue, n, fluctuates with time for a given inflow and outflow curve.

The disadvantage of the queuing theory is that the queues have no spatial dimension, and they do not have a proper length either (they do not occupy space). Other models, which overcome these problems, are discussed below.

Shockwave theory

Queuing theory provides some of the simplest models that can be used to model traffic flow conditions. However, the spatial dimension of traffic congestion in particular is not well described or – in the case of vertical queuing models – not described at all. Shockwave theory is able to describe the spatio-temporal properties of queues more accurately. This sub-section briefly introduces shockwave theory.

A shockwave describes the boundary between two traffic states that are characterized by different densities, speeds and/or flow rates. Shockwave theory describes the dynamics of shockwaves, in other words how the boundary between two traffic states moves in time and space.

Suppose that we have two traffic states: states 1 and 2. Let *S* denote the wave that separates these states. The speed of this shockwave *S* can be computed by:

$$\omega_{12} = \frac{q_2 - q_1}{k_2 - k_1} \tag{19}$$

In other words, the speed of the shockwave equals the jump in the flow over the wave divided by the jump in the density. This yields a nice graphical interpretation (Figure 7.13): if we consider the line that connects the two traffic states 1 and 2 in the fundamental diagram, then the slope of this line is exactly the same as the speed of the shock in the time–space plane.

Shockwave theory provides a simple means to predict traffic conditions in time and space. These predictions are largely in line with what can be observed in practice, but they have their limitations:

1. Traffic driving away from congestion does not accelerate smoothly towards the free speed but continues driving at the critical speed.
2. Transition from one state to the other always occurs in jumps, not taking into account the bound acceleration characteristics of real traffic.
3. There is no consideration of hysteresis.
4. There are no spontaneous transitions from one state to the other.
5. Location of congestion occurrence is not in line with reality.

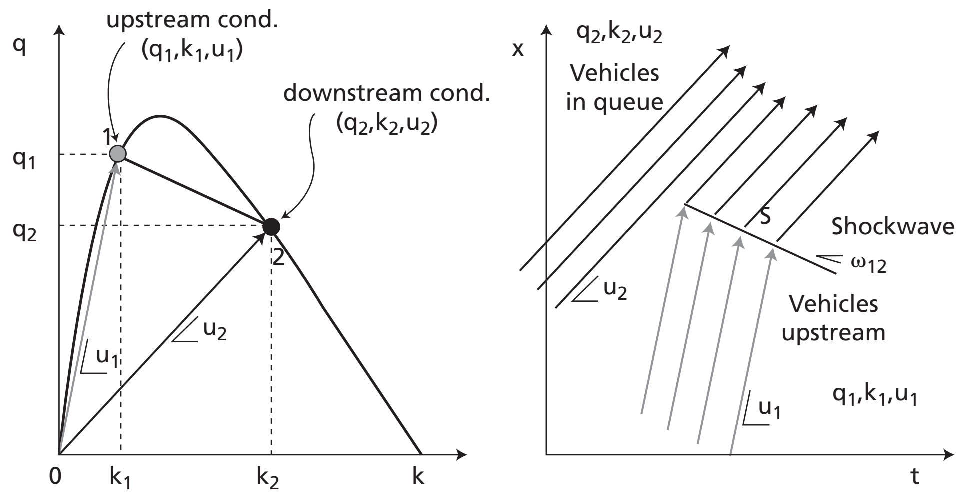


Figure 7.13 Graphical interpretation of shockwave speed

As a result, more advanced approaches have been proposed. Let us now consider the most important ones.

Continuum traffic flow models

Continuum traffic flow deals with traffic flow in terms of aggregate variables, such as flow, densities and mean speeds. Usually, the models are derived from the analogy between vehicular flow and the flow of continuous media (e.g. fluids or gases), complemented by specific relations describing the average macroscopic properties of traffic flow (e.g. the relation between density and speed). Continuum flow models generally have a limited number of equations that are relatively easy to handle.

Most continuum models describe the dynamics of density $k = k(x,t)$, mean instantaneous speed $u = u(x,t)$ and the flow $q = q(x,t)$. The density $k(x,t)$ describes the expected number of vehicles per unit length at instant t. The flow $q(x,t)$ equals the expected number of vehicles flowing past cross-section x during the time unit. The speed $u(x,t)$ equals the mean speed of the vehicle defined according to $q = ku$. Some macroscopic traffic flow models also contain partial differential equations of the speed variance $q = q(x,t)$, or the traffic pressure $P = P(x,t) = rq$. For an overview of continuum flow models, we refer to Hoogendoorn and Bovy (2001).

7.9 Network dynamics

In the preceding sections, we have presented some of the main traffic flow characteristics. Using the microscopic and macroscopic models discussed, flow operations on simple infrastructure elements can be explained and predicted. Predicting flow operations in a network is, obviously, more involved, since it also requires predicting the route traffic demand profiles, which in turn means modelling route choice, departure time choice, mode choice and so on.

Interestingly, it turns out that the overall dynamics of a traffic network can be described using a remarkably simple relation, referred to as the *macroscopic* or *network fundamental diagram (NFD)*. This diagram relates the vehicle accumulation – or average vehicle density – to the network performance. The network performance is defined by the flow, weighted by the number of lanes, and the length of the roadway segment for which the measured flow is representative.

This relation, which will be discussed in the following sections, shows one of the most important properties of network traffic operations, namely that its performance decreases when the number of vehicles becomes larger. In other words, when it is very busy in the network, performance goes down and fewer vehicles are able to complete their trip per unit of time. As a consequence, problems become even bigger.

Macroscopic fundamental diagram

Vehicular traffic network dynamics are atypical. Contrary to many other networks, network production (average rate at which travellers complete their trip) deteriorates once the number of vehicles in the network has surpassed the critical accumulation. Pioneering work by Daganzo and Geroliminis (2008) shows the existence of the NFD, clearly revealing this fundamental property. Figure 7.14 shows an example of the NFD. Knowledge of this fundamental property and its underlying mechanisms is pivotal in the design of effective traffic management.

Developing a macroscopic description of traffic flow is not a new idea. Thomson (1967) found the relationship between average speed and flow using data collected from central London streets. Wardrop (1968) stated that this relation between average speed and flow decreased monotonically, and Zahavi (1972) enriched Wardrop's theory by analysing real traffic data collected from various cities in the United Kingdom and United States.

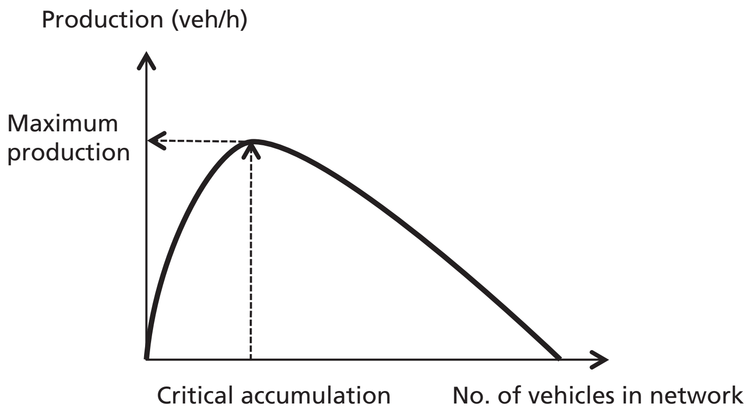


Figure 7.14 Example of the network fundamental diagram

Geroliminis and Daganzo (2008) have proven that NFDs exist in small networks, revealing the relation between the outflow and accumulation in the network. The accumulation is the number of vehicles in the network. The outflow is also called trip completion rate, reflecting the rate at which travellers reach their destinations. Similarly to a conventional link fundamental diagram, relating the local flow and density, three states are demonstrated on an NFD. When only a few vehicles use the network, the network is in the free flow condition and the outflow is low. With an increase in the number of vehicles, the outflow rises to the maximum. Like the critical density in a link fundamental diagram, the value of the corresponding accumulation when maximum outflow is reached is also an important parameter, called 'sweet spot'.

As the number of vehicles further increases, travellers will experience delay. If vehicles continue to enter the network, it will result in a congested state where vehicles block each other and the outflow declines (congested conditions). Furthermore, macroscopic feedback control strategies were introduced with the aim of keeping accumulation at a level at which outflow is maximized for areas with a high density of destination.

Causes of network degeneration

The two main causes of the production deterioration of overloaded networks are spill-back of queues possibly resulting in grid-lock effects, and the capacity drop. Spill-back occurs because of the simple fact that queues occupy space: a queue occurring at a bottleneck may propagate so far upstream that it will affect traffic flows that do not have to pass the bottleneck, for example when the queue passes a fork or an intersection upstream of the active bottleneck. As a result, congestion will propagate over other links of the network, potentially causing grid-lock phenomena. The capacity drop describes the fact that the free flow freeway capacity is considerably larger than the queue discharge rate.

7.10 Conclusions

The most important conclusions of this chapter are as follows:

1. Traffic flow theory and modelling are important in order to design comfortable and safe roads, to solve road congestion problems and to design adequate traffic management measures, amongst other things.
2. Traffic flow theory entails knowledge of the fundamental characteristics of traffic flows.
3. In traffic flow theory a basic distinction is made between microscopic and macroscopic traffic flow variables. Microscopic traffic flow variables focus on individual drivers. Macroscopic traffic flow variables reflect the average state of the traffic flow.
4. The fundamental diagram in traffic flow theory describes a statistical relation between the macroscopic flow variables of flow, density and speed. The basic premise underlying the fundamental diagram is that under similar traffic conditions drivers will behave in a similar way.
5. Traffic flow models can be used to simulate traffic, for instance to evaluate ex ante the use of a new part of the infrastructure. Models can be categorized based on, firstly, representation of the traffic flow in terms of flows (macroscopic), groups of drivers (macroscopic) or individual drivers (microscopic) and, secondly, underlying behavioural theory, which can be based on characteristics of the flow (macroscopic) or individual drivers (microscopic behaviour).
6. The overall dynamics of a traffic network can be described using a remarkably simple relation, referred to as the macroscopic or network fundamental diagram (NFD). This relation shows one of the most important properties of network traffic operations, namely that their performance decreases when the number of vehicles becomes greater.

REFERENCES

Ahmed, K., M. Ben-Akiva, H. Koutsopoulos and R. Mishalani (1996), 'Models of freeway lane-changing and gap acceptance behavior', *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Oxford: Pergamon, pp. 501–515.

Botma, H. (1999), 'The free speed distribution of drivers: estimation approaches', in P. Bovy (ed.), *Five Years Crossroads of Theory and Practice*, Delft: Delft University Press, pp. 1–22.

Brackstone, M. and M. McDonald (1999), 'Car-following: a historical review', *Transportation Research F*, 2, 181–186.

Brilon, W., R. Koenig and R. Troutbeck (1999), 'Useful estimation procedures for critical gaps', *Transportation Research A*, 33, 61–186.

Buckley, D. (1968), 'A semi-Poisson model for traffic flow', *Transportation Science*, 2, 107–133.

Cassidy, M.J. and R.L. Bertini (1999), 'Some traffic features at freeway bottlenecks', *Transportation Research Part B: Methodological*, 33 (1), 25–42.

Chandler, R.E., R. Herman and E.W. Montroll (1958), 'Traffic dynamics: studies in car following', *Operations Research*, 6, 165–184.

Chung, K., J. Rudjanakanoknada and M.J. Cassidy (2007), 'Relation between traffic density and capacity drop at three freeway bottlenecks', *Transportation Research Part B: Methodological*, 41 (1), 82–95.

Cowan, R.J. (1975), 'Useful headway models', *Transportation Research*, 9 (6), 371–375.

Daganzo, C.F. (2002a), 'A behavioral theory of multi-lane traffic flow, part I: long homogeneous freeway sections', *Transportation Research Part B: Methodological*, 36 (2), 131–158.

Daganzo, C.F. (2002b), 'A behavioral theory of multi-lane traffic flow, part II: merges and the onset of congestion', *Transportation Research Part B: Methodological*, 36 (2), 159–169.

Daganzo, C.F. and N. Geroliminis (2008), 'An analytical approximation for the macroscopic fundamental diagram of urban traffic', *Transportation Research Part B: Methodological*, 42 (9), 771–781.

Edie, L. (1965) 'Discussion of traffic stream measurements and definitions', in *Proceedings of the 2nd International Symposium on the Theory of Traffic Flow*, Paris: OECD, pp. 139–154.

Forbes, T., H. Zagorski, E. Holshousera and W. Deterline (1958), 'Measurement of driver reactions to tunnel conditions', *Highway Research Board Proceedings*, 37, 60–66.

Geroliminis, N. and C.F. Daganzo (2008), 'Existence of urban-scale macroscopic fundamental diagrams: some experimental findings', *Transportation Research Part B: Methodological*, 42 (9), 759–770.

Hall, F.L. and K. Agyemang-Duah (1991), 'Freeway capacity drop and the definition of capacity', *Transportation Research Record: Journal of the Transportation Research Board*, 1320, 91–98.

Herman, R. (1959), 'Traffic dynamics: analysis of stability in car-following', *Operation Research*, 1, 86–106.

Hoogendoorn, S.P. (2005), 'Unified approach to estimating free speed distributions', *Transportation Research Part B: Methodological*, 39 (8), 709–727.

Hoogendoorn, S.P. and H. Botma (1997), 'Modeling and estimation of headway distributions', *Transportation Research Record: Journal of the Transportation Research Board*, 1591, 14–22.

Hoogendoorn, S.P. and P.H.L. Bovy (2001), 'State-of-the-art of vehicular traffic modeling', Proceedings of the Institution of Mechanical Engineers, Part I, *Journal of Systems and Control Engineering*, 4 (215), 283–304.

Jepsen, M. (1998), 'On the speed–flow relationships in road traffic: a model of driver behaviour', *Proceedings of the Third International Symposium on Highway Capacity*, Copenhagen, pp. 297–319.

Kerner, B.S. (2004), *The Physics of Traffic: Empirical Freeway Pattern Features, Engineering Applications, and Theory*, Berlin: Springer.

Leutzbach, W. and R. Wiedemann (1986), 'Development and applications of traffic simulation models at the Karlsruhe Institut für Verkehrswesen', *Traffic Engineering and Control*, 27, 270–278.

May, A.D. (1990), *Traffic Flow Fundamentals*, Englewood Cliffs, NJ: Prentice Hall.

Michaels, R. (1963), 'Perceptual factors in car following', in *Proceedings of the Second International Symposium on the Theory of Road Traffic Flow*, Paris: OECD, pp. 44–59.

Minderhoud, M. (1999), 'Supported driving: impacts on motorway traffic flow', Ph.D. thesis, Delft University of Technology.

Minderhoud, M.M., H. Botma and P.H.L. Bovy (1996), 'An assessment of roadway capacity estimation methods', Technical Report vk2201.302, Delft University of Technology.

Nelson, W. (1982), *Applied Life Time Analysis*, New York: Wiley.

Newell, G.F. (1961), 'A theory of traffic flow in tunnels', in R. Herman (ed.), *Theory of Traffic Flow: Proceedings of the Symposium on the Theory of Traffic Flow*, Amsterdam: Elsevier, pp. 193–206.

Pignataro, L. (1973), *Traffic Engineering: Theory and Practice*, Englewood Cliffs, NJ: Prentice Hall.

Pipes, L. (1953), 'Car following models and the fundamental diagram of road traffic', *Transportation Research*, 1, 21–29.

Thomson, J.M. (1967), 'Speeds and flows of traffic in central London: speed–flow relations', *Traffic Engineering and Control*, 8 (12), 721–725.

Treiterer, J. and J.A. Myers (1974), 'The hysteresis phenomenon in traffic flow', in D.J. Buckley (ed.), *Proceedings of the 6th International Symposium on Transportation and Traffic Theory*, New York: Elsevier, pp. 13–38.

Wardrop, J.G. (1968), 'Journey speed and flow in central urban areas', *Traffic Engineering and Control*, 9 (11), 528–532.

Wasielewski, P. (1974), 'An integral equation for the semi-Poisson headway distribution', *Transportation Science*, 8, 237–247.

Zahavi, Y. (1972), 'Traffic performance evaluation of road networks by the α-relationship', Parts I and II, *Traffic Engineering and Control*, 14 (5 and 6), 228–231 and 292–293.

Part II

Impacts of the transport system

8

Transport technology to reduce transport's negative impacts

Jan Anne Annema, Robert van den Brink and Leonie Walta

8.1 Introduction

Mobility has changed tremendously over history. About 200 years ago, people walked, rode horses, sat in carriages and used barges. At the same time, goods and mail were transported on people's back, in carriages and by barge and boat. Two centuries later the transport system has revolutionized, as can be illustrated with three examples.

Today, over 600 million passenger cars travel the streets and the roads of the world (Worldometers, 2011). In 1900 this number was nearly zero. In 1950 containers did not exist. Sixty years later the modern container has transformed worldwide trade and economy (Levinson, 2008). According to Levinson, by making shipping so cheap the container paved the way for Asia to become the world's workshop, and brought consumers a previously unimaginable variety of low-cost products from around the globe. Finally, one of the first jet airliners (the Boeing 707) was introduced in 1959. In 2009, the world's airlines carried around an amazing 2.3 billion passengers on scheduled services (ICAO, 2010).

Technological progress in vehicles and infrastructure (see section 8.2) has resulted in more speed (thus reducing travel times), cheaper transport and more comfort. In relation to Chapter 6 of this book, this means that transport technology progress, broadly speaking, has often lowered transport resistance and, thereby, increased transport volumes. This chapter is not about transport technologies that have made transport faster, cheaper and more comfortable, because this is such a huge topic that the chapter would become far too long. We refer to Levinson (2008), Oliver Wyman (2007) and Williams and Weiss (2005) for some brief insights into innovations in freight transport, the automotive industry and the aircraft industry respectively.

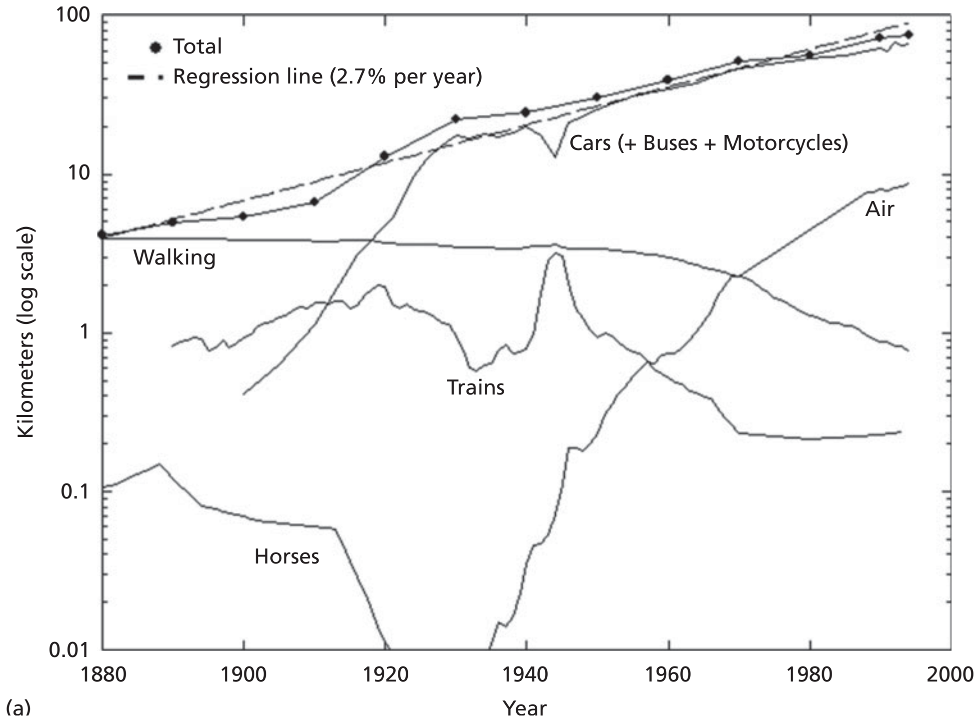
This chapter is about transport technologies which have already been implemented or which might be implemented in the future to reduce transport's drawbacks. The goal of this chapter is to explain the role technology has played and might play in the future to decrease transport's external effects (see Chapter 12 for an explanation of this concept). Over the last 50 years it has become more apparent that there are also severe problems related to the emerging transport system, such as air pollution, noise nuisance and increasing road traffic casualties. Thus transport technology in this chapter is defined as a set of tools, both hardware (physical) and software (algorithms or procedures), that have helped us or may help us in the future to make the transport system less congested, cleaner, quieter, safer and more fuel-efficient. As government policies always play an important role in reducing external effects (see Chapter 12), this chapter also explains the role industry and governments play and can play to minimize transport's external effects by the use of new vehicle technology.

The strong focus on vehicle technology and external effects is a limited approach. Therefore section 8.2 gives a brief description of the evolution of transport technology in general. In section 8.3 we start focusing on transport's external effects by explaining why a technologically imperfect transport system has emerged and what is required from a political economy perspective to implement new technologies aimed at decreasing the undesirable negative impacts of transport. In the following sections technologies are described that are aimed at decreasing air pollution (section 8.4), oil dependency and climate change impacts (section 8.5), noise nuisance (section 8.6), traffic fatalities and injuries (section 8.7) and congestion (section 8.8). Some reflections on technology dynamics and the nature of transportation technologies are then presented (section 8.9), together with some conclusions (section 8.10).

8.2 The evolution of transport technology

Figures 8.1a and 8.1b show the development in US passenger travel between 1880 and 2000 (Figure 8.1a) and the share of the total length of US transport infrastructure (Figure 8.1b).

From a transport technological perspective we can conclude from Figures 8.1a and 8.1b that the past trend is relatively straightforward: humans have

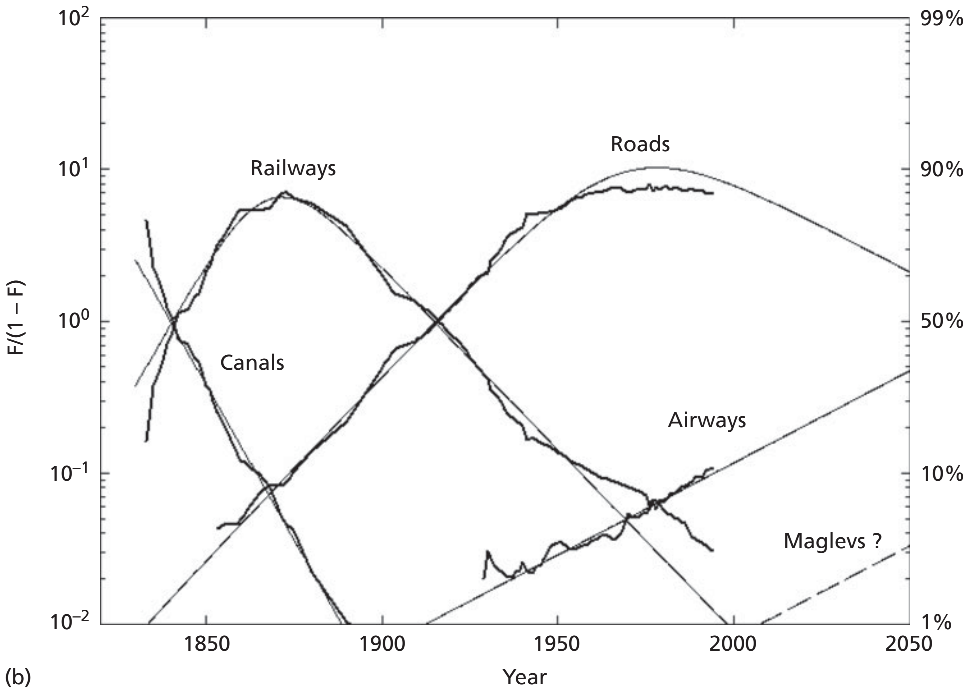


Source: Ausubel et al. (1998).

Figure 8.1a US passenger travel per capita per day by all modes

always aimed for higher-speed technologies. As already explained in Chapter 6, travel time is an important resistance factor for transport. Thus humankind apparently has chosen (consciously or unconsciously) to beat this resistance factor by developing higher-speed transport vehicles and infrastructure. By doing so, people have increased their mobility without violating the 'law' of the constant travel time budget (see Chapter 6). Americans have increased their mobility by 4.6 per cent each year (kilometres travelled per person per day) since 1870 (Ausubel et al., 1998). Gruebler (1990) estimated that the French have increased their mobility by about 4 per cent per year since 1800.

Figure 8.1b shows the authors speculating about the fast Maglevs (with a question mark) as the possible future new dominating transport means. Perhaps they are right or perhaps it will be high-speed rail in general not the specific Maglev[1] technology. Of course, new transport technology for the future is highly uncertain. Nevertheless, increased demand for high-speed transport technologies is expected to continue in the next few decades by most scholars (e.g. Schafer, 1998, 2006; Nakicenovic, 2008). Nakicenovic (2008) thinks that the next-generation transport modes of dominance can already be seen



Note: Squiggly lines: US transport infrastructure; smooth lines: model by the authors. F is the fraction of total length.

Source: Ausubel et al. (1998).

Figure 8.1b Shares of the actual length of the US transport infrastructure analysed with a model by the authors

around us, but most likely in a very embryonic form – suborbital space travel? – so that it is not predictable which one might be a winner.

8.3 Implementing new transportation technology to solve negative impacts: a policy perspective

In the early 1950s the car started to dominate the Western world transport system (Figure 8.1a). Sperling and Gordon (2009) talk about the US baby boom generation which came of age in comfortable car-dependent families in the late 1950s. However, as these authors point out, the 1960s and 1970s also brought about a rather sudden new attitude and new consciousness. In the 1960s, Ralph Nader campaigned against the reluctance of the car manufacturers to spend money on safety measures (Nader, 1965). Meadows et al. (1972) published their famous book with the telling title *The Limits to Growth*. This book, commissioned by the Club of Rome, modelled future population growth and the use of natural resources, showing that oil is a finite resource. Indeed, two worldwide oil crises, in 1973 and 1979, and an oil-price peak in 2008 showed the Western world that the supply of cheap oil is less self-evident than perhaps previously thought.

Despite new thinking and a greater awareness of the challenges facing transport and the need for technological change, the system is still, by far, not perfect (Table 8.1). Economists explain transport imperfections by pointing towards the existence of external effects (see also Chapter 12). External effects are impacts of a transport activity on third parties. The people who decide to travel or transport freight unintentionally do not (or not fully) take these impacts into account. For example, a third party can be people living close to a (rail)road or airport, and be confronted with traffic noise. Other road users can also be third parties, such as car-drivers who end up in a traffic jam owing to another car-driver's choice to travel on the same road at the same time, or cyclists who are involved in an accident because of a car-driver's choice to use the same road. Economists state that the existence of external effects is a kind of market failure, which can be a reason for governments to intervene.

Why do we use an imperfect transport system? Perhaps evolutionary economics can help here. On the one hand, this theory can help to explain how we have ended up with an imperfect transport system (Table 8.1). On the other hand, the theory can also help to explain why it is so difficult for governments to change it.

Innovation and selection towards an imperfect system[2]

In evolutionary economics all actors are assumed to have bounded rationality. Bounded rationality implies that actors, amongst others, have routines and habits, that they are satisfiers rather than optimizers and that they have a limited time horizon. Bounded rationality results in heterogeneity in behavioural strategies. Innovation is the result of this diversity. Sovacool (2009) describes how, by the end of the nineteenth century, a person seeking transport in the United States (and many other corners of the world) could choose between a bewildering array of different options: the horse, bicycle, train, subway, new steam-powered horseless 'carriage', petrol automobile and electric-powered vehicles – this is true diversity.

In evolutionary economics, serendipity plays an important role in explaining the innovation process. This means that a combination of chance, luck and knowledge results in an invention. Knowledge is important because empirical evidence shows that creative innovations are in most cases the result of a new combination of existing knowledge, techniques or concepts.

Table 8.1 Examples of external effects of transport

	Current statistics, some examples from all over the world	Historical and expected future trends
Traffic jams	In and around large US and European cities a car traveller may lose more than 60 hours in congestion per year [a),b)].	An increasing trend in the past. Without additional policies an increase in traffic jams in urbanized areas is to be expected.
Oil dependency	Worldwide transportation is responsible for approximately 60 per cent of oil consumption. The transport sector is the most exposed part of the economy to oil prices. Large amounts of oil produced come from countries with politically unstable regimes or from cartels such as the Organization of the Petroleum Exporting Countries (OPEC).	OPEC market power, particularly the Middle East, is expected to increase.
Global warming	Transportation's share in greenhouse gas emissions is estimated at 20 to 30 per cent.	Growing share in the past. Emissions will increase approximately 2 per cent per year with business as usual (so without additional policies).
Acidification and local air pollution	For example, European road transport contributed 13 per cent to the total emissions (i.e. from all sectors) of acidifying substances in 2006. Road transport is the dominant source of ozone precursors and accounted for 28 per cent of total ozone precursor emissions in 2007 in Europe. Road transport is a main source of emissions of fine particulates, contributing 21per cent to total emission of fine particulates in Europe.	In Europe, emissions of road transport air pollutants decreased by roughly 30 to 40 per cent between 1990 and 2007. A further emission reduction is expected.
Traffic safety	Worldwide an estimated 1.2 million people are killed in road crashes each year and as many as 50 million are injured.	In OECD countries traffic casualties decreased in the period 1970–2010. However, in economically developing countries an increase has taken place. If present trends continue, road traffic injuries are predicted to be the third-leading contributor to the global burden of disease and injury by 2020.

Table 8.1 (continued)

	Current statistics, some examples from all over the world	Historical and expected future trends
Noise nuisance	In 2006, the global population exposed to 55 DNL[c] of aircraft noise was approximately 21 million people. In the European Union 40 per cent of the inhabitants are exposed to equivalent road noise levels exceeding 55 dB(A) in daytime and more than 30 per cent in night-time.	In OECD countries road noise burdens have remained relatively constant since the 1990s; aircraft noise burdens have increased. In the future a further increase of transport noise is expected with business as usual.

Notes:
[a] http://euscorecard.inrix.com/scorecard_eu/
[b] http://mobility.tamu.edu/resources/ttt_measure_2010.pdf
[c] DNL (day night level) is a descriptor of noise level based on energy equivalent noise level (Leq) over a whole day with a penalty of 10 dB(A) for night-time noise.

Source: Bluhm et al. (2004); WHO (2004); Worley (2006); OECD/ITF (2008); IEA/OECD (2009); EEA (2011).

Within the scope of behavioural diversity, 'knowledge' has many faces and it is unavoidable that a lot of knowledge, and therefore money, will be wasted. This means that knowledge 'waste' in the form of trial and error and culs-de-sac is needed, to get 'fit' technologies (one may even wonder if 'waste' is the right phrase here). In other words, according to evolutionary economics, petrol and diesel vehicles have emerged as fit technology thanks to early competition with the train, the steam-powered carriage and electric-powered vehicles.

Selection processes reduce the innovation diversity. The innovation and selection processes together determine the 'fitness' of a certain new technological alternative. Fitness is a measure for survival and reproduction. In the selection processes, the new innovations are put to the test of survival. Selection relates to many different factors: physical possibilities or impossibilities of a new technology, technical usage pros and cons, economic factors such as price and the possibilities of producing the technology on a large scale, psychological factors (do people actually like the new invention?), institutional barriers (will governments allow the new invention to enter the market?) and so forth. Sovacool (2009) in his history of early modes of transport in the US argues that all of these factors have played a role in explaining why the petrol automobile finally became the winner. For example, he shows that, even though the electric-powered vehicles initially (1895–1905) had many advantages, they did not break through because they were more expensive than petrol cars, had slower top speeds, were difficult to charge, were mostly confined to urban areas, and came to be seen as old-fashioned and feminine.

Lock-in and co-evolution

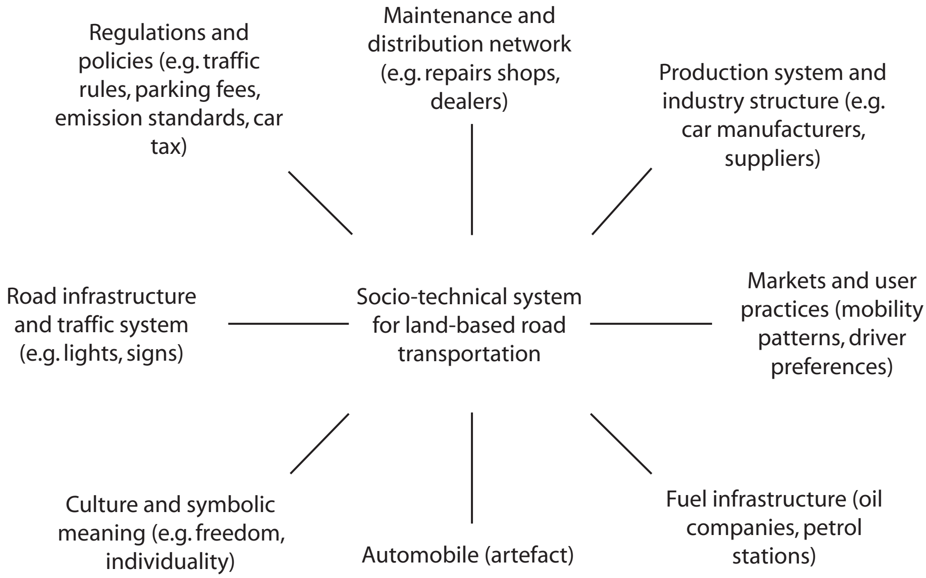
The dynamics of evolutionary systems as described here result in three important concepts for this chapter: path dependency, lock-in and co-evolution. Path dependency means that, for a certain technology, as a result of increasing economies of scale, a self-reinforcing feedback loop may emerge which ends up in the dominance of that technology. With economies of scale economists mean that the more one specific technology is used and produced, the more we learn to improve our use and production of this technology, and share the same network. For example, all people and shippers using internal combustion engines (ICEs) share the same fuel network and make use of the same maintenance and distribution networks (repair shops and dealers). Thus economies of scale result in substantially lower costs.

Unfortunately, the consequence of path dependency means that there may be a historically unavoidable path towards the complete dominance of one technology. Disadvantages that did not occur, or were not seen as disadvantages at the early stage of its development, can make it difficult to change technologies. A situation of so-called technological 'lock-in' has unintentionally been created. In many ways the current dominant transport technology (the ICE fuelled by oil products) can be regarded as such a lock-in situation (Table 8.1).

Co-evolution is related to the evolutionary notion that innovations are in most cases the consequence of combining already existing ideas or systems. Co-evolution focuses on the ways partial systems (such as cars, vans and lorries on the one hand, and the road network or the fuel network on the other) develop, work together and to an increasing extent influence each other's evolution. One may say that co-evolution of different partial technical systems working increasingly together will often result in improved synchronization and extra benefits for the users. However, if the resulting co-evolved system has societal disadvantages (Table 8.1), it seems even more difficult to escape the lock-in of the closely intertwined system.

System innovations (transitions)

Systems can change via system innovation or transition. In system innovations theory (Geels, 2005) socio-technical systems are defined. As an



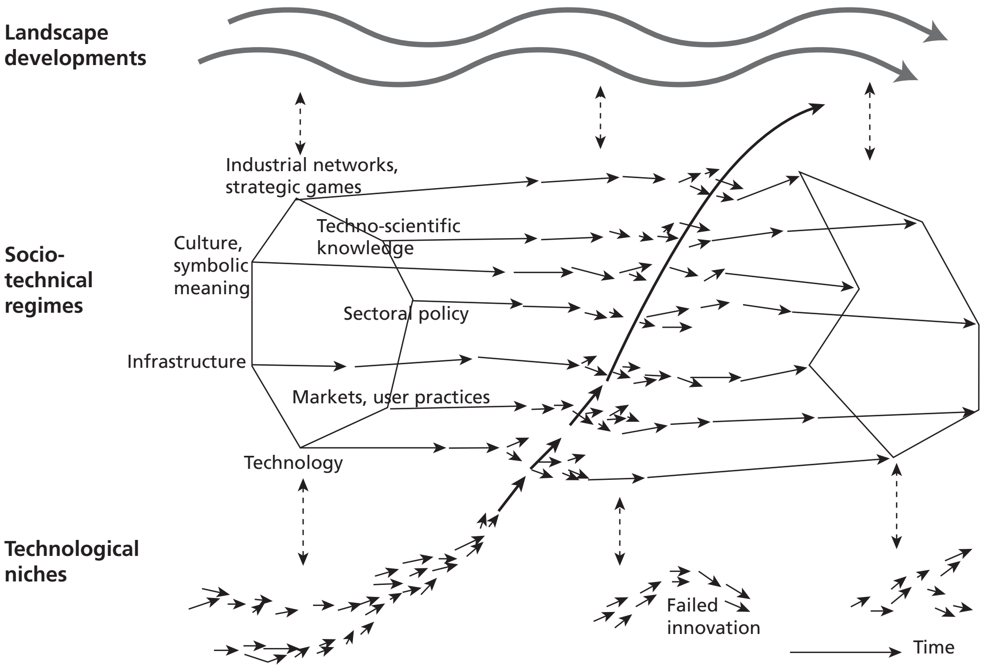
Source: Geels (2005).

Figure 8.2 Socio-technical system for road transport

example, Figure 8.2 illustrates the socio-technical system for road transport. The figure shows all interrelated entities within this socio-technical system, and explains the way transport technology is embedded in society in terms of physical infrastructure, institutions, market and culture. The notion of a co-evolved, closely intertwined system is clearly recognizable in Figure 8.2.

A system innovation is a transition from one socio-technical system to another, potentially characterized by a technological change (e.g. from sailing ships to steam ships). Transition is a process which can be explained by using the multi-level perspective. The multi-level perspective (see Figure 8.3) combines insights from evolutionary economics, innovation studies, and science and technology studies, in order to understand transitions and the dynamics in system innovation.

An existing socio-technical system is depicted as a heptagon somewhere to the middle left of Figure 8.3. The changed socio-technical system is symbolized at the right side (later in time) of Figure 8.3 as a differently shaped heptagon. In the multi-level model of system innovations three levels are distinguished:

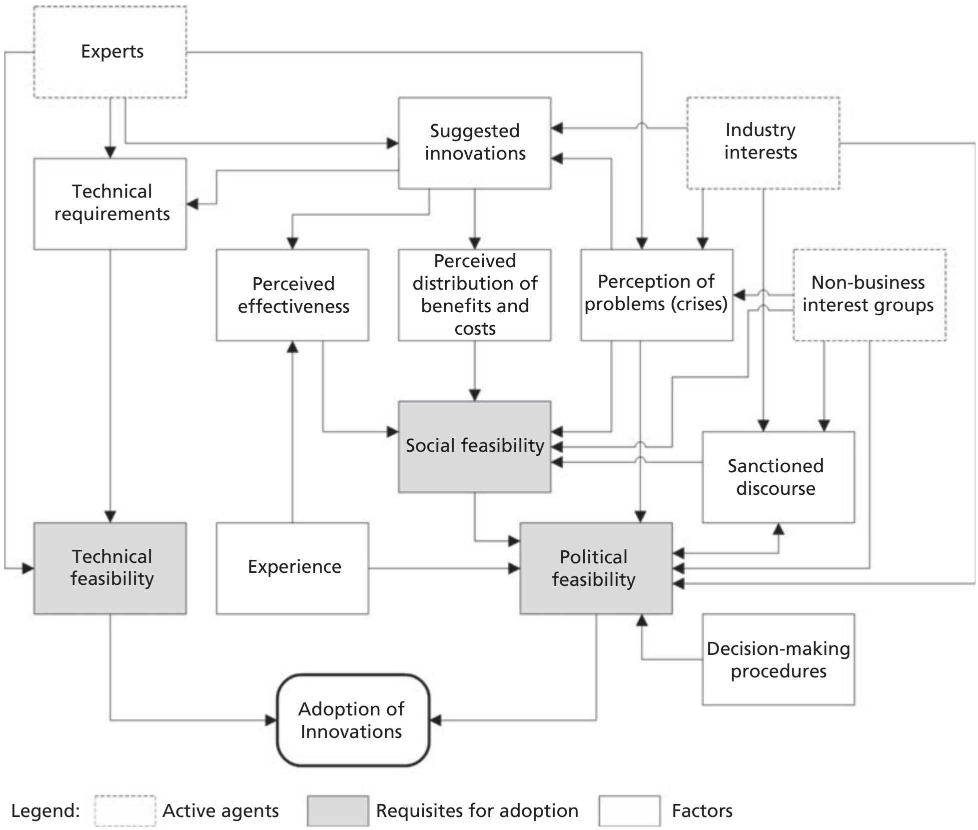


Source: Geels (2002).

Figure 8.3 Multi-level model of system innovations

1. **Technological niches.** These are the inventions (represented as small arrows) that try to penetrate the second level of the socio-technical regimes. Some inventions become a success and change the socio-technical system, and some fail.
2. **Socio-technical regimes.** See Figure 8.2 for an example.
3. **Landscape developments.** These are 'a set of deep structural trends external to the regime' (Geels, 2005: 78). The socio-technical landscape cannot be changed by actors within the socio-technical regime (level 2), in our case transportation. The landscape developments include both tangible (such as the built environment) and non-tangible aspects (such as economic growth, migration, perhaps the biological needs of humans to expand their territory, environmental problems, and wars). There are slow developments in the landscape (e.g. demographical changes) as well as rapid developments (e.g. oil crises). Some landscape developments may help a technical niche to penetrate the socio-technical regime successfully, but other landscape developments may result in failure.

Public authorities can play a role in the technological niche level (for example, 'protect' some early but very promising inventions by giving subsidies), but



Source: Feitelson and Salomon (2004).

Figure 8.4 A political-economy model for explaining the adoption of innovations

they are also regime players. This means public authorities can have an accelerating as well as a decelerating role in system innovations. The area of system innovation research is still in development, but may become state-of-the art in innovation theory and the transition to sustainability (Elzen et al., 2004).

A political economy model

A pragmatic view on transport innovations and the role of public authorities has been put forward by Feitelson and Salomon (2004). They have developed a political economy framework for analysing the adoption of transport innovations (Figure 8.4).

The box 'perception of problems' (see Figure 8.4, top right) is an important factor for the successful adoption of new technology, as will be shown in the subsequent sections on technological innovations aimed at reducing the external costs of transport. The oil crises (in 1973 and 1979), severe smog periods in Los Angeles and London, and a continuing, increasing amount of traffic casualties in the 1960s have all spurred technical changes in transport. Experts such as scientists, advisers and consultants also play a role in the framework (see Figure 8.4, top left). Experts suggest technical innovations and the means to implement these innovations. They research technical issues and societal problems such as congestion, climate change and air pollution. They also perform formal cost–benefit analysis showing, for example, that certain policies such as implementing nationwide road charging using GPS technology will have a positive benefit-to-cost ratio. Nevertheless, the Feitelson and Salomon framework shows clearly that a favourable benefit-to-cost ratio for a policy does not mean that this policy will be adopted. For politicians the perceived effectiveness and the perceived distribution of benefits and costs (who wins, who loses) of policy also play an important role in their decision making.

Most important in their framework is that they think that the adoption of innovations is predicated on the economic, social, political and technical feasibility. Thus it is insufficient that an innovation is technologically superior, that it meets a strict benefit–cost criterion or that there are a majority of voters supporting it. In their view only a particular combination of these feasibility issues will result in successful innovation.

The Feitelson and Salomon political economy framework, with its emphasis on the importance of social and political feasibility for explaining successful adoption of innovations, relates to the notions of the current transport system as a co-evolved lock-in situation or as a socio-technical regime. Both concepts imply that the current situation not only means a strong dependence on one dominant technology, but also implies a strong position for the defenders of the existing socio-technical regime, such as vehicle manufacturers, the oil industry, unions and billions of consumers worldwide who, by using their voice and vote, influence the position of non-business interest groups and political parties (Figure 8.4). To illustrate the political power of certain actors: according to the European Automobile Manufacturers' Association (ACEA, 2011), the European automobile industry supports more than 12 million jobs (directly and indirectly), representing a work force amounting to more than the size of the population of a country such as Belgium. This is not meant to blame the defenders of an imperfect system. After all, it is clear that the current dominant transport technology also has many advantages related to economies of scale. So it seems obvious that industries and people profiting from these advantages are reluctant advocates for a fast and radical change.

8.4 Technological innovations to reduce transport air pollution

Sections 8.4 to 8.8 proceed to discuss the role of technology in the areas of the environment, accessibility and safety – see Chapters 9 to 11 for an introduction to these areas. We start with air pollution.

The problem

The air pollution problem results from oil products (e.g. petrol, diesel and kerosene) and air combining in the combustion processes to produce energy for vehicle propulsion, but at the same time producing undesired chemical by-products, such as nitrogen oxides (NO_x), carbon monoxide (CO), volatile organic compounds (VOCs), sulphur dioxide (SO_2), soot and fine particulate matter. Particulate matter (PM_{10}, $PM_{2.5}$ or $PM_{0.1}$) is of varying size, but is all microscopic. The subscripts define the different fractions of PM diameter sizes: respectively, smaller than 10 μm, 2.5 μm and 0.1 μm. Tyre wear, brake linings and road surfaces all contribute to emissions of PM.

The story of transport related air pollution problems starts in California, USA. In 1940, California's population rose to 7 million, with the number of registered vehicles approaching 2.8 million and an amount of vehicle miles travelled of 24 billion (California Environmental Protection Agency, Air Resources Board, 2011). Three years later the first of many recognized episodes of smog occurred in Los Angeles. In the 1950s it became clear that NO_x and VOCs in the presence of ultraviolet radiation from the sun formed the smog (a key component of which is O_3, ozone). In that period it also became clear that the continuous increase in road vehicles produced more of these pollutants (see also Chapter 10). Later, in the 1980s, the phenomenon of 'acid rain' provided an additional reason for air pollution policies (see Chapter 10). Acid rain occurs when sulphur dioxide (SO_2) and NO_x are emitted into the atmosphere. NO_x and SO_2 undergo chemical transformations in the air and are absorbed by water droplets in clouds. The droplets then fall to earth as rain, snow or sleet. This can increase the acidity of the soil, and affect the chemical balance of lakes and streams, resulting in ecological damage.

Past and current technological changes

Owing to the perception of problems (severe smog), it is not surprising that the first legislation requiring controls of vehicle emissions was passed in California in 1959, followed within a year by the creation of the statewide

Motor Vehicle Pollution Control Board, the first of its kind, to test and certify devices to clean up California's cars (Sperling and Gordon, 2009). This policy led to the use of positive crankcase ventilation[3] in 1961, the first automotive emission control technology ever required. Tailpipe standards for CO and VOCs were adopted by California in 1966 and then for NO_x in 1971. The rest of the USA, Europe and Japan lagged somewhat behind the Californian initiative, but today all large economies of the world have implemented air pollution emission standards for all kinds of road vehicles (cars, vans and lorries). A comprehensive overview of vehicle standards from all over the world can be found on Dieselnet (2011).

The emission standards policy started in most countries with the control of air polluting substances such as CO and VOCs. Today, there is almost universal agreement over the need to control vehicle regulated substances such as CO, VOCs, NO_x and particulate matter (PM_{10}). An important feature of this policy is that the standards are tightened over time (see Figure 8.5 for the European Union passenger car standards). The reason for this tightening over time is that the emission standards are the result of political negotiations between governments and the vehicle industry. The result is a compromise between tightness and the time when the vehicle industry has to apply the stricter rules. Of course, one could criticize this approach as being slow. However, this approach can also be regarded as perhaps the only way in which economic, technical, social and political feasibility can be met. Still, even this 'slow' policy is not perfect (see below).

For petrol cars an important technological innovation of this policy has been the introduction of the three-way catalyst or catalytic converter. The role of the catalyst is simultaneously to remove the primary pollutants CO, VOCs and NO_x by catalysing their conversion to carbon dioxide (CO_2), steam (H_2O) and nitrogen (N_2). In optimal circumstances the three-way catalyst has a conversion efficiency of 90 to 95 per cent. The catalyst could technically not cope with lead in petrol. Lead was a petrol additive to increase its octane rating, and it retarded valve wear. Therefore, as a side-effect of the tailpipe standards, the highly toxic lead was banned. Nowadays, in OECD countries leaded petrol can no longer be bought. Sulphur also reduces the effectiveness of the catalyst irreversibly. This has led to policies lowering the sulphur content of petrol and diesel. Since 2009 petrol and diesel have been practically sulphur-free.

In Europe, diesel cars had a market share of roughly 30 per cent in the period 2005 to 2011. Three-way catalysts cannot be used for diesel engines because diesel engines operate with excess air which hinders the NO_x reduction

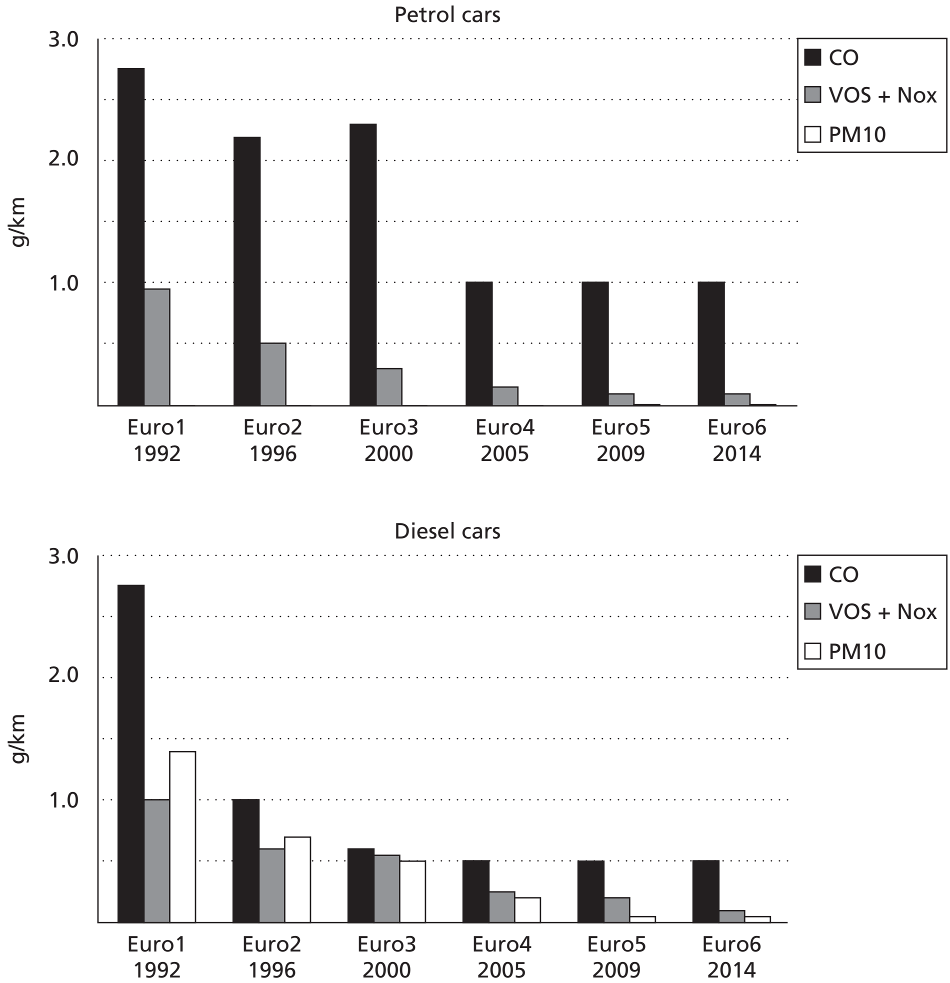


Figure 8.5 Emission standards for passenger cars in the European Union

process in the tailpipe gas. Therefore, in diesel passenger cars, so-called oxidation catalyst technology has been used since the mid-1990s. In the 1990s it also became clearer that the emission of very fine particles by diesel cars (PM_{10}, $PM_{2.5}$ and $PM_{0.1}$) was harmful to human health. Although the oxidation catalysts reduce PM_{10}, it became obvious that extra technologies were needed to meet the relatively strict PM_{10} Euro 4 and 5 emission standards (Figure 8.5). This technology became the particulate filter. First introduced by Peugeot in a passenger car in 1999, it has become universal, and it is now impossible to buy a new diesel passenger car without such a filter. Modern particulate filters are self-regenerative and have a particle efficiency of approximately 90 per cent.

For diesel engines it is particularly challenging to decrease NO_x and PM_{10} together. Broadly speaking, it is valid to state that, for diesel engines, measures to reduce PM_{10} emissions will result in an increase of NO_x emissions. Therefore, the use of so-called exhaust gas recirculation (EGR) is increasingly applied in car diesel engines to reduce NO_x emission. EGR recirculates a portion of the engine's tailpipe gas back to the cylinders. There, the exhaust gas will replace some of the excess air in the pre-combustion mixture. The EGR technology will not be sufficient to meet the Euro 6 car standards in 2014 (Figure 8.5), and the so-called selective catalytic reduction catalysts (SCR catalysts) are required for diesel passenger cars.

This SCR technology together with particulate filters is already used in heavy duty vehicles. An SCR catalyst can reduce NO_x to N_2 and, for example, water (H_2O) despite the relatively high presence of oxygen in the exhaust gas. However, a disadvantage of this technology is that, to carry out the chemical conversion, the SCR catalysts require a reductant, in most cases urea ($(NH_2)_2CO$). Urea functions as the 'helper' reductant, reducing NO_x to N_2, water and CO_2. The helper role implies that urea has to be available at garages or filling stations. Smart additional technology (on-board diagnostic systems) is also required to warn the diesel vehicle user automatically to refill the urea tank in time, because without this reductant the NO_x emissions will increase substantially.

Although the vehicle emission standard policy and the subsequent technical innovations have been successful, they are not perfect. Air quality has not improved as much as predicted with the tightening of emissions standards, especially in respect of NO_x (ECMT, 2006). An important reason for this is the gap between the performance of emission control measures during type approval tests and their effectiveness under real operating conditions. To be clear, the vehicle emission standard policy implies that vehicle or engine types are only allowed to be put on the market from a certain date if they pass the type approval tests. In the approval tests driving conditions are simulated in controlled circumstances (test cycle). New car types are tested on roller benches where the emitted mass of a regulated substance (in grams or milligrams) is measured per kilometre driven in a test cycle. For heavy duty vehicles, new engine types are tested to establish how much of a regulated substance (in grams or milligrams) per delivered amount of engine power in a test cycle is emitted (for example, in grams per kilowatt hours). ECMT (2006) points out three factors that can explain the gap between the exhaust emissions during type approval and those during in-use operation:

1. Manufacturers may use cycle by-pass measures in order to pass the type approval tests, but achieve better fuel efficiency or other performance enhancement at the cost of higher emissions during operation on the road.
2. On-board diagnostic systems (OBD) related gaps. Properly functioning OBD warn drivers that their exhaust emission control devices do not function optimally so that they can have their control devices repaired as soon as possible. In approval tests better functioning OBD are assumed, compared to how they are in reality.
3. Driver behaviour related gaps.

An example of a gap between real-world and test results has recently been found in Europe. Real-world emission measurements on Euro III, IV and V heavy duty vehicles[4] (Hausberger et al., 2009; Ligterink et al., 2009) showed hardly any improvement in NO_x emissions at low speed for Euro IV and V heavy duty vehicles equipped with SCR. PBL (2010) give two explanations for the relatively high NO_x emission factors (on urban roads) for Euro IV and V heavy duty vehicles. Firstly, the SCR catalysts do not seem to function optimally on urban roads because of the relatively low motor load on this type of road resulting in relatively low engine exhaust temperatures. Secondly, it is probable that truck manufacturers have optimized their motors and emission reduction technologies in order to comply with the European emission test cycle, while in the real world the effectiveness of the emission reduction techniques applied is lower (cycle by-pass). As mentioned by ECMT (2006), the main avenue for improving regulations is to modify test cycles to mirror more closely driving as it is in the real world.

Future

For the future, the use of advanced technologies such as electric vehicles or fuel cell vehicles holds the promise of zero air pollutants emission, at least at the stage of using the vehicles. To produce electricity and hydrogen (the probable feedstock for fuel cells), air pollution may still arise. For further explanation on these advanced technologies, see the next section.

8.5 Climate change and oil dependency

The problem

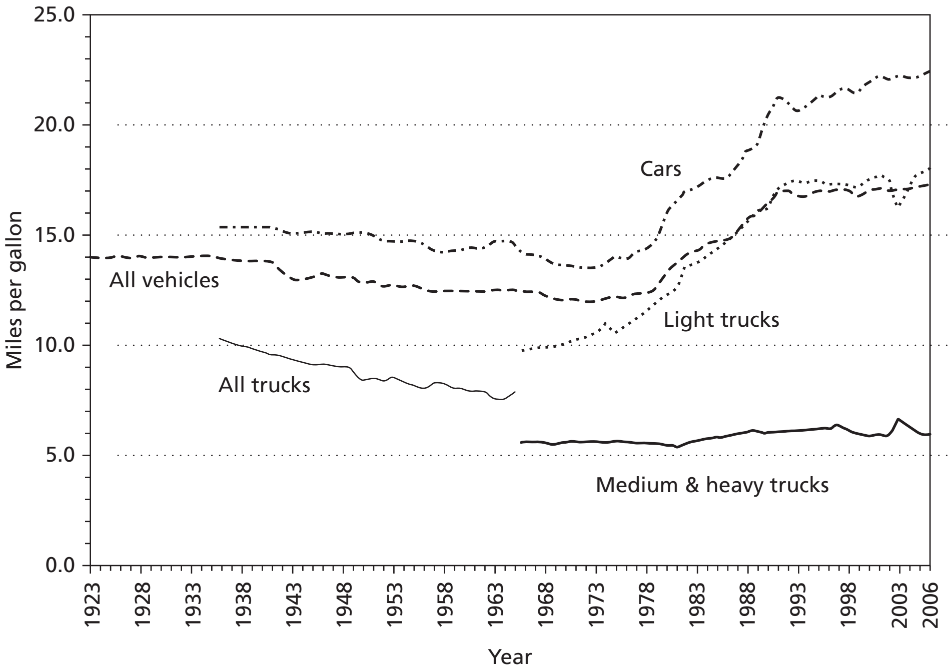
Climate change and oil dependency are interrelated problems because climate change is a result of the use of oil (a fossil fuel). Crude oil and its products such as petrol, diesel and kerosene are high-energetic long-chain hydrocarbons, which when burnt in engines deliver propulsion energy. However, as a by-product of the burning process, CO_2 is formed, which is held responsible for contributing to climate change. To limit climate change to an acceptable level, strong reductions in CO_2 emissions are required (Pachauri and Reisinger, 2007).

Oil dependency has been seen as a political problem in the Western world for a long time. The American comedian Jon Stewart shows in a very funny but also insightful history lesson that President Obama, who in 2010 called for America to stop depending on foreign oil, is just the latest in a long line of American presidents who have tried to make the US less dependent on oil and failed (Stewart, 2011). Oil dependence is seen as a problem not only because of environmental concerns but also because oil consumption can make a country's economy highly dependent on the behaviour of unstable or politically undesirable oil-producing regimes or cartels, such as OPEC.

Past and current technological changes

The past trend is that the transportation system increasingly used oil products and emitted CO_2, despite efforts to curb the trend. Perhaps this is because the perception of these problems has not been strong enough to really push innovations forward towards a far less oil-dependent transport system. Additionally, it is also true that to beat vehicle CO_2 emissions no relatively cheap end-of-pipe technologies are available such as the ones implemented to reduce air pollution (section 8.4). This lack of 'easy' technical end-of-pipe measures has made it difficult for governments to pursue strong CO_2 emission policies in transport; these policies were not feasible from an economic, political and social point of view.

Nevertheless, what has happened in the past is not totally bleak. Fuel efficiency – energy use per kilometre driven – of vehicles has improved in the past, but these improvements were not sufficiently strong to result in absolute CO_2 emission or oil usage reduction. Figure 8.6 shows the changes in fuel efficiency between 1923 and 2006 on US roads (Sivak and Tsimhoni, 2009). For cars, the data show a gradual decline to an all-time low of 13.4 miles per gallon (mpg)[5] around 1973. After the oil embargo of 1973 the fuel economy rose relatively quickly to 21.2 mpg in 1991. This was due to technical measures such as engine improvements, together with measures to lower the car air and rolling resistance. The improvements from 1991 were substantially smaller, achieving an average 22.4 mpg in 2006. The Arab oil embargo in 1973 played a role in the improvement and also led to a policy in the USA called Corporate Average Fuel Economy (CAFE) standards. CAFE



Source: Sivak and Tsimhoni (2009).

Figure 8.6 Fuel efficiency of US vehicles from 1923 to 2006

is the sales weighted average fuel economy, expressed in mpg, of a manufacturer's fleet of passenger cars or light trucks manufactured for sale in the United States, for any given model year (http://www.nhtsa.gov/cars/rules/cafe/overview.htm). By comparison, the CAFE standard for 2006 model passenger cars was 27.5 mpg (Sivak and Tsimhoni, 2009).

In Europe the average car is more fuel-efficient compared to the US average car. This is at least partly a result of the much higher fuel taxes in Europe. For comparison, in 2008 average car fuel economy for European countries such as the UK, Germany, France and Spain was roughly 35 per cent higher compared to the 2008 US average car fuel economy (Schipper, 2011) (the US fuel economy relates to cars and household light trucks and SUVs). Nevertheless, in Europe the same trend in fuel-efficiency improvement can be observed as in the USA, with major improvements right after 1973, followed by a sharp decline of the slope of improvement. For example, measured for the period 1985–95, the average Dutch car on the market has shown no improvement in fuel economy. The explanation is that in this period cars per class (a certain weight and cylinder capacity) became more fuel-efficient but, at the same time, the average consumer bought a bigger and heavier car. Between 1980 and 1999 the average new Dutch car bought became approximately 20 per cent heavier. And the average consumer also bought a car with a bigger engine, expressed in cylinder capacity, and this amounted to an increase of 15 per cent for the average car between 1980 and 1999 (van den Brink and van Wee, 2001). To summarize, between the mid-1980s and 1995 technical improvements in the Netherlands were traded off against consumer preferences for bigger and heavier cars.

Since 1995, new cars in Europe have again become more fuel-efficient. The European Commission has made voluntary agreements with European, Japanese and Korean car manufacturers to achieve, for new cars, a reduction in CO_2 emissions per kilometre driven of 25 per cent compared to 1995. This means that the average new car in 2008 would emit around 140 g CO_2 per kilometre (equivalent to a fuel efficiency of 40 mpg). The 2008 goal was not met, but the improvements made were still significant. The EU average CO_2 emission factor for new cars decreased from around 185 g/km in 1995 to 154 g/km in 2008 (equivalent to 36 mpg) (T&E, 2009). Note that these values are test values, and real-world values are significantly higher (see further below). The technical improvements included more efficient petrol and diesel engines, hybrid power trains, more sophisticated transmissions, low rolling-resistance tyres, improved attention to detail, aerodynamics, stop–start technology and regenerative charging systems. An increased share of diesel car sales on the European market has also played a role, for modern diesel engines are more fuel-efficient than petrol engines.

In order to meet climate change goals, the use of low-carbon fuels has been stimulated by governments all over the world. The most important example is the so-called biofuels. Biodiesel and bio-ethanol blended with petrol were used in vehicles with up to a share of 1.8 per cent in 2008 worldwide (UNEP, 2009). The idea is that biofuels are made from plants, which take up CO_2 from the air to grow. When the biofuels are burnt to give heat for propulsion this CO_2 is again released. The net CO_2 emission of this cycle is zero, at least in theory.

The advantage of biofuels is that they can relatively easily be implemented into the current transport system. Most vehicles on the road today can operate on up to a 10 per cent ethanol blend (US Department of Transportation, 2010). However, there are also disadvantages related to this technology. Currently, first-generation biofuels are commercially produced. The basic feedstock used is seeds, grains or whole plants from crops such as corn, sugar cane, rapeseed, wheat, sunflower seeds or oil palm. Greenhouse gases are emitted to produce the liquid biofuels from the feedstock, making the net emission of biofuels not equal to zero. In fact, it is highly uncertain what the actual CO_2 emission reduction is from using first-generation biofuels compared to fossil fuels. Some sources (for example, IFEU, 2007; Liska et al., 2009; Menichetti and Otto, 2009; Bureau et al., 2010) even report an increase in CO_2 emissions for certain biofuel feedstock-production method combinations as compared to oil. Additionally, first-generation biofuels have disadvantages related to their land use, which results in biofuel production being in competition with food and fodder production and nature. In future, therefore, experts expect more from so-called second-generation and even third-generation biofuels in their capability to reduce transport related CO_2 emissions (see 'Advanced (second-generation or third-generation) biofuels' in 'Future technology', below).

Future technology

During 2008, the crude oil price peaked at nearly $150 per barrel. At the same time a worldwide economic crisis emerged which hit the car market hard. Several governments (e.g. the US) decided to support their home car industry financially. Perhaps related to this atmosphere of crisis, it seems that governments took on a somewhat bolder position in order to force the car industry to produce more fuel-efficient cars than in the past. In December 2008, the European Commission agreed upon a regulation stating that the EU car manufacturers' fleet average had to be aligned with 130 g CO_2/km – partially as of 2012 and completely by 2015. Two years later the European Parliament and the Council agreed to a target of 175 g/km in 2017 and of 147 g/km in 2020 for vans. New American fuel standards were also put in place: for new passenger cars (39 mpg by 2016), light duty trucks (30 mpg) and medium duty passenger vehicles (35.5 mpg) (IEA/OECD, 2009). Japan and Korea have set targets of 16.8 km/l (JC08 test cycle) and 17 km/l (US city cycle) respectively by 2015.

Which technologies will be used in the future? It can be expected that car manufacturers will try to improve the conventional ICE, by combining it with technologies to reduce the vehicle weight and the air and rolling resistance. Hybrid technology will come to the fore. Hybrid vehicles have two power sources, in most cases a combustion engine and an electric motor. The electric motor is powered by a battery pack which is recharged by the combustion engine and by recovering brake energy. Hybrid vehicles are more fuel-efficient compared to conventional vehicles because they have a smaller combustion engine and because the brake energy can be recovered. Furthermore, in so-called serial and combined hybrids, the ICE is used to charge the battery pack of the electric motor. In this configuration, the ICE can be used at a nearly constant load where its efficiency is best (IEA/OECD, 2009). In test cycles, hybrid cars used over 10 per cent less fuel compared to conventional cars in the same class (size, luxury, engine characteristics). However, the real-world emission factors of hybrid cars can be much worse compared to the test values (see also section 8.4). Based on fuel refilling data (actual amount of fuel tanked compared to number of kilometres driven) of lease drivers in the Netherlands, the difference between real-world and test fuel efficiency for a Toyota Prius (a combined hybrid) amounted to nearly 50 per cent (Ligterink and Bos, 2010).

In the longer term advanced technologies may be used in transport (IEA/OECD, 2009), such as:

1. **Plug-in hybrids.** These vehicles are essentially similar to conventional hybrids except that they also have the capacity to draw electricity from the grid to charge the batteries. The advantage is that these vehicles can run a few kilometres solely on electricity, which is favourable for urban air quality. If the electricity is produced using (partly) biomass, sun or wind power, then additional CO_2 emission reduction takes place. These vehicles are already on the market at the time of writing this chapter, but it is not certain that they will become as popular as the old ICE technology.
2. **Electric vehicles (EVs).** These old acquaintances are entirely powered by batteries and an electric motor without the need for an internal conventional engine. Like plug-in hybrids, they offer the prospect of zero vehicle emissions (CO_2 as well as air pollutants; sections 8.4 and 8.5), at least in the vehicle usage phase. Of course, electricity has to be produced, which may involve CO_2 emissions, air pollutant emissions or nuclear waste depending on the source and the conversion technology chosen. Mellios et al. (2011) estimate a European real-world CO_2 emission factor of 50 g/km for new electric vehicles in 2020, taking the average European carbon intensity of electricity production into account. This compares to the average real-world CO_2 emission factors for new small petrol and diesel cars in 2020 of around 130 and 115 g/km, respectively (Mellios et al., 2011). Electric cars are also already on the market at the time of writing this chapter, but – again – it is not certain that they will become as popular as the old ICE technology.
3. **Fuel cell vehicles (FCVs).** A fuel cell is an electrochemical cell which is capable of converting chemical energy (in a source fuel such as hydrogen, H_2) into electricity. The electricity is used to power an electric motor that drives the wheels and supports other vehicle func-

tions. Currently, the most suitable fuel cell for vehicle application is the proton exchange membrane (PEM) fuel cell. These fuel cells need to be sourced by very pure H_2. Simply stated, in the PEM fuel cell, H_2 is combined with oxygen (O_2) to produce water (H_2O) and electricity for propulsion. However, the H_2 refuelling of FCVs raises some difficult technical issues. These vehicles need an on-board reformation process, which is an on-board small chemical factory to convert a hydrocarbon based fuel (methane, ethanol, petrol and so forth) into very pure H_2. This route is expensive. Alternatively, the vehicles need on-board H_2 storage, which requires a lot of space (H_2 has a much lower energy density than petrol, for instance, even in a highly compressed state: up to a factor 25 lower for 150 Bar H_2) and an extensive H_2 production and distribution infrastructure. Even when petrol is used as the H_2 source, fuel FCVs produce fewer CO_2 emissions compared to ICEs because they are more fuel-efficient. However, the real CO_2 gain of this technology is when H_2 is produced using biomass, sun or wind power, as in the case of electricity.

4. **Advanced (second-generation or third-generation) biofuels.** The essence of advanced biofuel-producing technology compared to the first generation is that non-food crops (such as grasses and trees), waste and perhaps even algae are used as feedstock to produce liquid or gaseous fuels such as biogas, bio-ethanol or biodiesel. Although the technologies to convert non-food solid feedstock to liquid fuels are well known, commercial application is not yet possible.

A lot of research is being done on these four advanced technologies. In all four cases major barriers can be identified which have to be tackled by research before large scale use can be foreseen. These barriers relate to the social, political and economic feasibility factors as proposed by Feitelson and Salomon (2004) (see section 8.3), mostly related to the currently relatively high costs of the advanced technologies compared to the still dominating ICE technology costs. Additionally, technological feasibility is also a barrier, such as the limited range of the current vehicle battery technology before recharging is required.

The State of California has tried to advance zero-emission vehicles (ZEV) on their market by implementing ZEV regulation since 1996. This idea started with relatively simple 2.5 per cent and 10 per cent requirements: x per cent of car sales on the Californian market had to be ZEV from a certain future date. However, as Sperling and Gordon (2009) point out, the ZEV rule led a tortured life, undergoing industry lawsuits and continuing modifications. The simple 2.5 per cent and 10 per cent requirements have given way to a

Table 8.2 Future technologies to improve CO_2 emission performance of trucks and freight movement, aviation and shipping

	Future
Trucks and freight movement	– Through better technologies (such as advanced engines, light-weighting, improved aerodynamics, better tyres) new trucks can probably be made 30 to 40 per cent more fuel-efficient by 2030. – Second-generation biofuels may be the only way to decarbonize trucking fuel substantially.
Aviation	– A wide range of fuel-efficiency technologies, including aerodynamic improvements, weight reduction and engine efficiency, might have the potential to make the average aircraft nearly twice as efficient in 2050 as it is today. Note that this efficiency is measured in energy use per passenger – so larger planes are a major factor here. Some of the planes flying today will still be in the air in 2050, implying that aviation improvements face a longer time scale than road transport in terms of replacement. – Improved air traffic control may result in fuel savings to the order of 5 to 10 per cent. – Second-generation biofuels could have an important potential to decarbonize aviation fuel.
Shipping	– The average size of ships may rise so that shipping becomes steadily more efficient per tonne kilometre moved. Measures here include cleaner fuels and slower speeds – also use of wind to assist. As in aviation, the life of a boat is 40–50 years, resulting in slow replacement rates. – Technical options related to vessel design, engine design and propulsion design (e.g. the propeller) have been identified which could achieve a 50 per cent or greater reduction in energy use per tonne kilometre. – Second-generation biofuels could hold an important potential to decarbonize shipping fuel.

Source: IEA/OECD (2009); Christensen et al. (2010).

complex of arcane rules. The latest revision in 2008 requires automakers to produce a total of 7500 fuel cell vehicles or 12 500 battery electric vehicles (or some combination thereof) between 2012 and 2014, along with 58 000 plug-in hybrids (Sperling and Gordon, 2009).

Most emphasis has been placed on future car technology. However, for other transport categories also, CO_2 emission reduction measures can be taken, and these are summarized in Table 8.2.

8.6 Noise

The problem

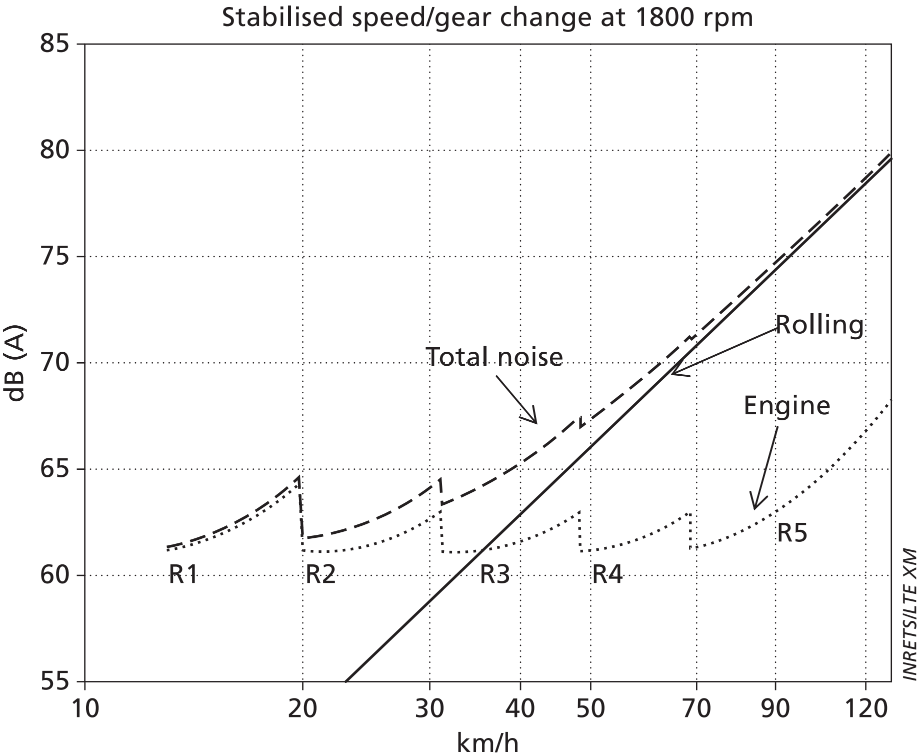
Road transport and aviation cause air pressure pulsations which can travel as waves at around 330 km/h in the air to a recipient. If the air pressure changes occur between 20 and 20 000 times per second (the frequency), the human ear, as the recipient, can hear a sound. Louder noise is caused by greater changes in the air pressure. The perception of the pitch of the sound is determined by the frequency. The problem is that exposure to noise can result in direct negative well-being impacts (nuisance) and in negative health impacts such as hearing damage and/or heart and coronary diseases (Berglund et al., 1999). Stress, distraction and sleeping disorders are seen by these authors as the biological factors which might explain why noise – a physical phenomenon – can result in health problems.

Noise exposure can be prevented in roughly two ways. Firstly, the noise production at the source can be decreased by taking volume or technical measures. This section is mainly about these technical measures. Secondly, it is possible to shield the recipient from the noise source. In transport the most well-known shield example is the noise barriers along roads. Although these barriers can be regarded as a kind of technology, we will not discuss them further. Specific information can be found, for example, in Kotzen and English (1999).

Road transport

In road transport noise is caused by the engine and the contact between tyres and road. At higher speeds (roughly more than 40 km/h), tyre–road surface contact noise is dominant (Figure 8.7 for cars, for example).

The European Union has set noise emission standards for light and heavy duty vehicles since 1970 (Directive 70/150). The standards in the type approval tests were tightened between 1970 and 1995 in steps, with a total reduction of 8 to 11 dB(A), similar to the air pollution policy (section 8.4). The logarithm unit decibel (dB) is a common unit used to express noise loads. Using so-called A-weighing the dB unit can be corrected according to the sensitivity of the human ear (dB(A)). A 3 dB(A) reduction by technical measures is equivalent to a 50 per cent reduction of traffic volume. A noise reduction of 8 to 11 dB(A) corresponds to an 85 to 90 per cent reduction of traffic volume.



Note: R1 to R5 represent gear shifts.

Source: Mitchell (2009).

Figure 8.7 Relation between car speed and noise causes

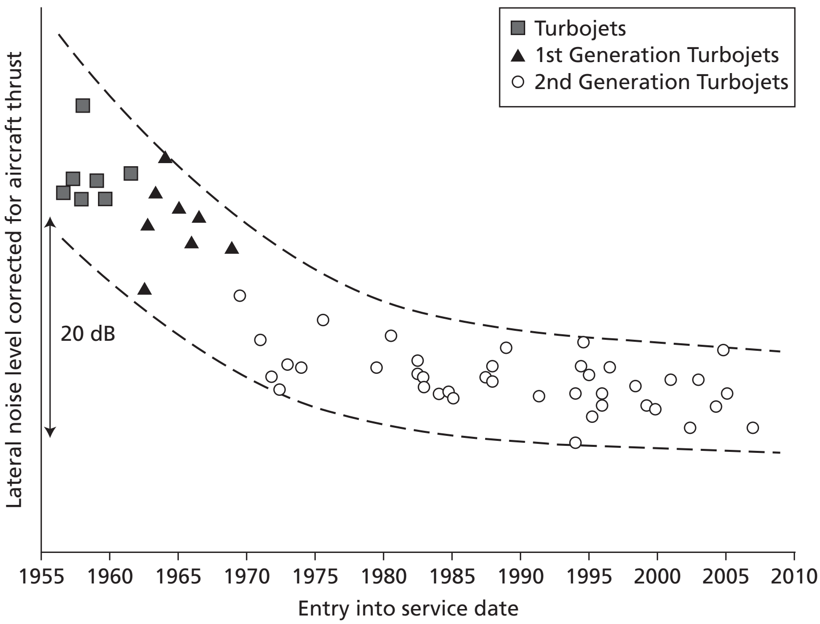
In the type approval tests, improvements were indeed realized in the 1980s and 1990s: 2 to 3 dB(A) for cars and 4 to 6 dB(A) for heavy duty vehicles. These improvements in the tests were reached by improved insulation of the engine and drive-line. However, in the same period (1974–1996) measurements alongside roads showed that passenger cars became hardly any quieter and heavy duty vehicles only 3 dB(A) (van der Toorn et al., 1997; van der Toorn and van Vliet, 2000; Heinz, 2005). The reasons for this difference between test and real world is, in the first place, that in the real world tyre–road surface contact noise is much more dominant (Figure 8.7) compared to the type in the approval test, where very silent tyres can be used. Secondly, the circumstances to test the maximum noise loads in the approval tests hardly ever occur in reality. And, thirdly, the steps to tighten the noise standards are partly offset by relaxing the required test circumstances at the same time. So the actual progress made is rather disappointing.

Since 2001, European standards have been put in place for tyres (Directive 2001/41/EC). Tyres can be made to be more silent, for example by making them smaller (less broad) and less stiff. However, these changes are more easily said than done, because the adaptations could lead to an unwanted higher rolling resistance (and, thus, lower fuel efficiency; section 8.5) and less grip on wet surfaces (hampering safety; section 8.7). From 1 October 2011 only tyres which apply to the type used in approval tests may be sold on the European market. Nevertheless, the effectiveness of this policy is doubted. The reasons are twofold. Firstly, the directive uses standards which most tyre manufacturers were already meeting in 2000. Secondly, the directive is based on a reduction of 1 dB(A) from the measured value in the test because of possible measurement errors and, in the directive, decimal truncation is allowed. These rules imply that a measured value in the test of 77.9 means that a 76 dB(A) standard is met.

Another way to take measures against the dominant tyre–road surface contact noise is to use low road noise surfaces. Several techniques are possible, which are related to a) optimizing the texture of roads, b) making porous surfaces and c) making road surfaces elastic (Descornet and Goubert, 2006). For example, in the Netherlands one-layer porous asphalt has become the standard highway road surface type. It has become the standard not only for its noise reduction but also because these types of asphalt are safer owing to their improved water permeability. The low noise aspect of porous asphalt is related to its good absorption of both rolling and power train noise, which leads to a noise reduction of on average 3 dB(A) at higher speeds. Rubber road surfaces can even reduce the noise emissions alongside the road by more than 8 dB(A). However, these surfaces can only be applied on roads with low traffic volumes.

Aviation (based on ICAO, 2008)

Aircraft noise emission control is regulated via the International Civil Aviation Organization (ICAO) of the United Nations. The ICAO started to develop a set of noise limit standards for aircraft in 1971. The ICAO standards are aimed at ensuring that any new aircraft entering service would have to use the best noise reduction technology available at that time. The first standards appeared in Chapter 2 of an annex of a report,[6] which explains why aircraft noise standards are called 'chapters'. The Boeing 727 and the Douglas DC-9 are examples of aircraft covered by the early Chapter 2 standards. More stringent standards became applicable in October 1977 (Chapter 3). The Boeing 737-300/400, Boeing 767 and Airbus 319 are examples of aircraft covered by this chapter. From 2002, Chapter 4 demands are applicable.



Source: ICAO (2008).

Figure 8.8 Aircraft lateral noise development 1955–2007

Noise from a single aircraft is primarily produced by the engine as air is sucked into the engine and exits from the exhaust at high velocity. Noise is also created by the airframe as it moves through the air. Figure 8.8 shows that aircraft have been producing less noise, especially in the period 1955–70. An important technical explanation for the improvements realized is the introduction of much higher bypass ratio jet engines into service in the 1970s (second-generation turbofans) to achieve better fuel economy. The bypass ratio (BPR) is the relationship between the amount of air passing through the core of a turbofan engine, compared with the amount of air drawn in by the fan but bypassing the core. The inclusion of noise absorbing material in engines and engine nacelles, as well as overall nacelle design, and mechanical refinements on engines, together with airframe adjustments, have also all contributed incrementally to further reducing the noise of jet-powered aircraft. According to ICAO it appears likely that the future will be similar to the past, with steady incremental progress in a number of areas, but with no dramatic improvement in any one area.

ICAO (2008) points out that it is also possible to achieve noise reductions through changes in aircraft operational procedures. There are several operational measures that can be adopted, such as changing runways and routes, and special noise abatement manoeuvres during take-off and approach, thereby reducing the number of people exposed to noise in specific areas around airports.

8.7 Safety

The problem

Traffic injuries and accidents form a major problem (Table 8.1), especially related to road transport. The causes of accidents can be manifold, but accidents are all related to the simple fact that road users driving at high speed, or where there is a big difference in their speeds, share the same space. Lack of attention, carelessness, bad luck, unclear situations, drowsiness, distraction (e.g. using mobile phones while driving) and alcohol use can all be factors resulting in small or very severe accidents. Two geographically different road safety trends can be observed. In most economic developed countries, road trauma levels peaked during the 1960s and early 1970s. After that, large reductions in casualties were achieved. The second trend is far less positive, namely that in economic developing countries transport related casualties have steadily increased. WHO (2004) estimated that in 2002 around 85 per cent of all global road deaths and 96 per cent of all children killed worldwide as a result of road traffic injuries occur in low-income and middle-income countries.

Past and current technological changes

The many traffic deaths in the economic developed world in the early 1970s spurred many governments to take action (related to a feeling of crisis; see Figure 8.4, the Feitelson and Salomon framework). The best-known actions include more stringent drink-driving laws, the mandatory wearing of seat belts and the use of crash helmets for mopeds and motorcycles. These three technologies may not seem to be the most advanced appliances in transport. However, based on a very rough calculation, these technologies might have saved hundreds of thousands of lives worldwide since the 1970s.

In the 1990s, protection of car occupants was substantially improved. For example, in the European Union several directives were issued on car frontal and side impact safety. Furthermore, EURO NCAP was established by European governments and consumer organizations in 1997. This organization provides standardized crash tests and independent consumer information about car safety. These initiatives led to the wide use of safety innovations such as airbags, crumple zones, safety cages and antilock braking systems

(ABS). Interestingly, EURO NCAP was established because proposals for the adoption in European legislation of the European Enhanced Vehicle-safety Committee (EEVC) test proposals were being strongly resisted by the car industry (EURO NCAP, 2011).

All the technologies mentioned up to now, except ABS, are called passive safety measures because they aim to limit the damage of an accident. In addition, active safety measures can be taken. These measures aim to prevent accidents. Table 8.3 gives an overview of passive and active technological safety measures, illustrated with some impact assessment results.

Future

For the future, technology related to so-called intelligent transport systems (ITS), in particular, seems promising (for an overview see www.itsoverview.its.dot.gov). These technologies can be used not only in cars and heavy duty vehicles but also in motorcycles (Ambak et al., 2009). ITS are active safety measures and are based on information, communication and satellite technologies. The aim of ITS is not only to enhance safety but also to mitigate traffic congestion (see section 8.8) and to improve the quality of the environment (Shah and Dal, 2007). Advanced driver assistance systems (ADAS) are one example of ITS. Applications in this ITS family, like lane departure warning, automatic cruise control, monitoring and warning systems for driver vigilance (intervenes when driver drowsiness, fatigue and inattention occur) or night vision, have already been introduced to some vehicles (Techmer, 2007). Collision warning and avoidance systems are also already in use. Collision warning systems use radar and internal-vehicular information to detect any collision risk. Intelligent speed adaptation (ISA) (see also Table 8.3) is another ITS family. With ISA, the vehicle has information on the permitted or recommended maximum speed for the road along which it is travelling. The standard ISA system uses an in-vehicle digital road map on to which speed limits have been coded, combined with a satellite positioning system. If the driver exceeds the permitted or recommended maximum speed limit a system intervenes to control the speed of the vehicle. The intervention system can be one of the following (Carsten and Tate, 2005):

1. **An advisory system.** The driver is informed by a voice, for example, of the speed limit and when it is being exceeded.
2. **A voluntary system.** Now the ISA system can directly intervene when speed limits are exceeded because it is linked to the vehicle controls (such as to the acceleration pedal or throttle). However, the driver can choose whether and when to override the intervention.

Table 8.3 Safety technology measures

Technology	Example of use or impact assessment
Passive safety measures related to infrastructure:	
Crash-protective roadsides	Collapsible lighting columns and other devices that break away on impact were first introduced in the United States in the 1970s and are now used widely throughout the world.
Crash cushions	In Birmingham, England, installing crash cushions resulted locally in a 40 per cent reduction in injury crashes and a reduction (from 67 to 14 per cent) in the number of fatal and serious crashes at the treated sites.
Passive safety measures related to vehicles:	
Daytime running lights for cars and motorized vehicles (energy-saving but visible lighting technology); high-mounted stop lamps in cars	A study of data over four years from nine American states concluded that, on average, cars fitted with automatic daytime running lights were involved in 3.2 per cent fewer multiple crashes than vehicles without.
Crash-protective vehicle design (crumple zones, safety cages)	A review of the effectiveness of casualty reduction measures in the United Kingdom between 1980 and 1996 found that crash protection improvements to vehicles contributed significantly to reducing casualties. These improvements counted for around 15 per cent of the reduction observed.
Safer car fronts to protect pedestrians and cyclists	It has been estimated that take-up of tests aimed at improving the front of passenger cars with respect to pedestrian and cyclist safety could avoid 20 per cent of deaths and serious injuries to pedestrians and cyclists in European Union countries annually.
Car occupant protection (e.g. structural design, the design and fitting of seat belts, child restraints, airbags, anti-burst door latches, laminated glass windscreens, seat and head restraints)	When used, seatbelts have been found to reduce the risk of serious and fatal injury by between 40 and 65 per cent. Estimates of the general effectiveness of airbags in reducing deaths in all types of crashes range from 8 to 14 per cent.
'Smart', audible seatbelt reminders	In Sweden investigations have revealed that reminders in all cars could lead to national levels of seatbelt use of around 97 per cent, contributing to a reduction of some 20 per cent in car occupant deaths.
Helmet wearing (bicycles, mopeds, motorcycles)	In Malaysia, where legislation on the use of helmets was introduced in 1973, it was estimated that the law led to a reduction of about 30 per cent in motorcycle deaths.

Table 8.3 (continued)

Technology	Example of use or impact assessment
On-board electronic stability programmes (the electronics help to stabilize vehicles when they suddenly have to change direction, for example to avoid a collision or when the vehicle skids)	A recent Swedish evaluation of the effects of this technology produced promising results, especially for bad weather conditions, with reductions in injury crashes of 32 per cent and 38 per cent on ice and snow, respectively.
Active safety measures related to infrastructure:	
Lanes for overtaking, median barriers, rumble strips, speed humps, traffic lights, other road signs	The use of speed bumps, in the form of rumble strips and speed humps, has been found to be effective on Ghanaian roads. For instance, rumble strips on the main Accra–Kumasi highway at the crash hot spot of Suhum Junction reduced the number of traffic crashes by around 35 per cent. Fatalities fell by some 55 per cent and serious injuries by 76 per cent, between January 2000 and April 2001.
Speed cameras	The cost–benefit ratios of speed cameras have been reported to range between 1:3 and 1:27. See Allsop (2010) also for their effectiveness.
Active safety measures related to vehicles:	
Speed adaptation such as intelligent speed adaptation (ISA)	In Sweden several trials with ISA were carried out. One trial was conducted with ISA primarily in built-up areas with speed limits of 50 km/h or 30 km/h, and the test drivers were both private car and commercial drivers. If the driver exceeded the speed limit, light and sound signals were activated. The Swedish National Road Administration reported a high level of driver acceptance of devices in urban areas and suggested that they could reduce crash injuries by 20–30 per cent in urban areas.
Alcohol ignition interlocks (vehicles can only start after the driver has proven that he or she has not drunk alcoholic beverages)	Around half of Canada's provinces and territories have embarked on alcohol interlock programmes, and in the United States most states have passed legislation for enabling such devices.
Speed limiters in heavy duty vehicles	It has been estimated that speed governors on heavy goods vehicles could contribute to a reduction of 2 per cent in the total number of injury crashes.

Source: WHO (2004).

3. **A mandatory system.** In this case, no override of the intervention system is possible. If the maximum speed limit is exceeded, the system will irrevocably slow down the vehicle.

ISA, as well as other in-car safety systems, is only slowly penetrating the market, although the technology has been known for years (de Kievit et al., 2008).

8.8 Congestion

The problem

Congestion arises if the number of vehicles exceeds the handling capacity of a road, station, airport or seaport. In this section only road congestion will be presented. Data on road congestion levels are widespread. However, the current situation is relatively clear, as stated, for example, by the European Conference of Ministers of Transport (ECMT, 2007): 'Congestion is increasing in many urban areas across the OECD (and elsewhere) and in locations where populations and city economies are growing it is likely to continue to increase.'

Past and current technology

The standard measures to keep urbanized areas uncongested are to build more roads and lanes and to invest in public transit such as subways, buses and railways.

Since the late 1970s, other technological means to achieve more road capacity have been applied, and these can be summarized under the term '(dynamic) traffic management'. With traffic management the road authority applies devices, such as traffic signals, dynamic signs and sometimes gates, to inform or regulate traffic. An example of regulation is ramp management in which the number of vehicles entering a highway is restricted. The idea is that overall congestion levels can be decreased by preventing too many vehicles entering the highway in too short a time span. Another example of dynamic traffic management is using large dynamic information signs. The information on these signs can be changed according to the actual traffic situation to inform road users to adapt their speed or to change routes because there is a traffic jam ahead. Since the 1990s, additional real-time information about the location of traffic jams has also been given by commercial vehicle navigation systems. The more advanced navigation systems use so-called floating car data. These data are gathered using the mobile phones of the occupants of road vehicles. Mobile phones continuously search for contact with the nearest transmit–receive antennas, which are positioned along the roads. By measuring the speed of mobile phones travelling to a next antenna along a road stretch, the mobile phone provider can determine a free flow or congested traffic situation. Accordingly, it can, together with the navigation system provider, inform users of the advanced navigation system about the actual traffic situation ahead. Consequently, these users may choose to change their route. These technologies can have major drawbacks for urban liveability and traffic safety where the alternative routes go through urban areas.

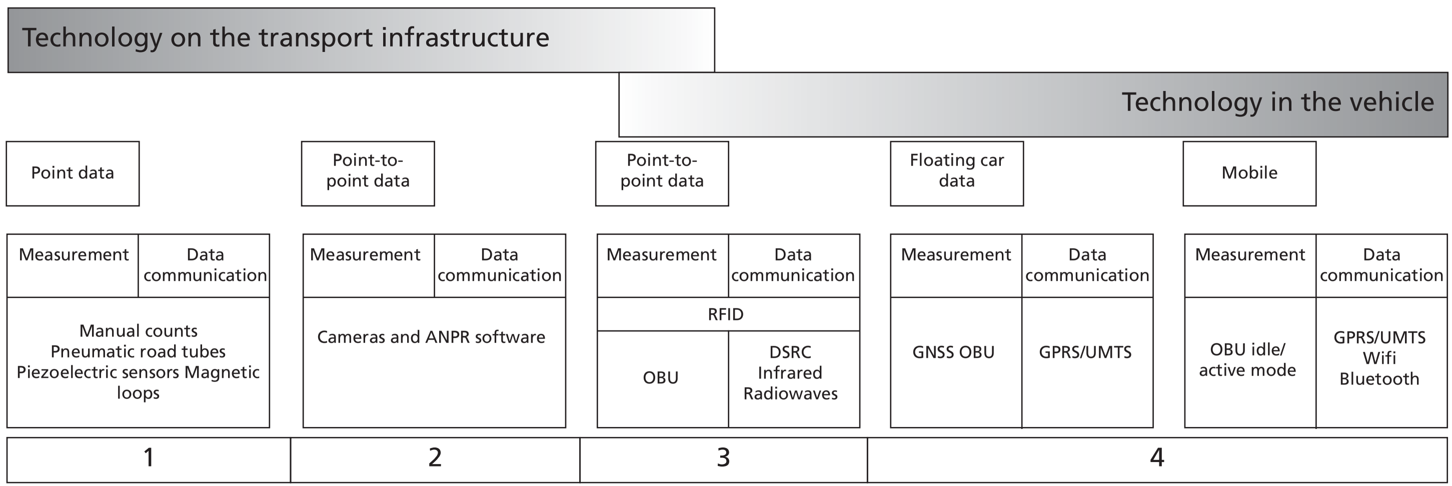
Pricing policies

Several countries, cities or regions worldwide have implemented road charging (McFadden, 2007). Well-known examples are the London congestion charge, Singapore, and several projects in Sweden and Norway, including, most recently, Stockholm. The concept of congestion charging or road pricing is relatively old. For example, in the UK in 1964 the Smeed Report (Smeed, 1964) appeared on the technical feasibility of various methods for improving the pricing system for the use of roads.

The core policy idea is to charge road users according to time and place. In order to be able to charge people 'correctly', or what economists call 'first-best charging' (e.g. Verhoef, 1996), rather complicated technology is required. With first-best, economists mean that road users should be charged in such a way that they pay their full marginal societal costs[7] up to the level that the marginal social costs equal the marginal social benefits (see Chapter 12). In the process, the charging authority needs to know exactly where, when and which people drive and with which vehicle they drive and, subsequently, they have to confront these people with a charge related to, amongst other things, the congestion weight and they have to collect the charge (billing). Figure 8.9 illustrates the different technologies and combinations of technologies that are available to perform these tasks (as of spring 2011).

Four main categories of technology are distinguished in Figure 8.9, which range from technology on the transport infrastructure to technology in the vehicle:

1. These technologies (manual counts, pneumatic road tubes and so forth) are not suitable for road pricing itself because they are too labour intensive (manual counts) or they cannot identify who is driving. However, these technologies can be used to support a road charging scheme, for



Source: Vonk Noordegraaf et al. (2009).

Figure 8.9 Technologies that can be used for road pricing example by measuring the actual congestion levels, which can be used to determine the charge levels.

2. Cameras and automatic number plate recognition (ANPR) software. This technology is used in London and Stockholm. Video cameras (positioned above or alongside roads) take pictures of the number plates of vehicles using roads or entering a certain city zone. When this is combined with a large number plate database, people can be billed.
3. Radio-frequency identification (RFID) or beacon transponder systems combine technology by the roadside and in the vehicle. A transponder, also called on-board unit (OBU) or 'tag', is placed in the vehicle and transmits data to a roadside beacon. Examples of types of data communication are radio waves, infrared and so-called dedicated short range communication (DSRC). The OBUs can perform several functions, including charging and billing.
4. In-vehicle-only systems use OBUs connected to a wide-area data communication system such as general packet radio system (GPRS) or universal mobile telecommunications system (UMTS). The data communication systems are combined with the global navigations satellite system (GNSS) or cellular positioning (through the cellular network) (see above under floating car data, p. 195) to measure location and time. In these systems, the OBUs can also perform several functions, including charging and billing.

It is not clear what the future will bring related to congestion charging. Although some road charging systems have been successfully implemented worldwide, there are also failed and cumbersome implementation processes (see Ison et al., 2008; Schaller 2010; Vonk Noordegraaf et al., 2012). The societal and political feasibility (see section 8.3, the framework of Feitelson and Salomon) of road charging is often difficult when related to issues such as the question about its fairness, as some people lose and others win, people see the charge as just another tax, people have a distrust related to its effectiveness, and they have concerns about privacy issues ('Big Brother is watching you'). In fact, the Feitelson and Salomon scheme (Figure 8.4) is based, amongst other things, on an analysis of road pricing implementation.

Future

Intelligent transport systems can also help to overcome congestion, just as they can in the case of safety. An example is adaptive cruise control (ACC) technology. With ACC the vehicle speed and the distance to the vehicle ahead are automatically adjusted. It uses a radar sensor to measure the distance between the vehicles. The current commercially available ACC systems are working only within a limited speed range (i.e. not at slower speeds) and they mainly increase driving comfort. New ACC systems that can work on the full speed range are now becoming available and include a 'stop-and-go' function that can automatically stop the vehicle and start driving again in stop-and-go traffic. These full-speed-range ACC systems, also referred to as congestion assistants, can have a substantial positive impact on traffic flow and congestion (van Driel and van Arem, 2010). The next step in ACC development is to make it cooperative, in which the first step could be the infrastructure-to-vehicle communication of advice on the speed and following distance, and the second step the vehicle-to-vehicle communication of speed and following distance, which would then be directly used by the ACC system. This cooperative adaptive cruise control can further improve traffic flow (van Arem et al., 2006).

8.9 The nature of technologies and the dynamics

Sections 8.4 to 8.8 show that, thanks to technological innovations, air quality and traffic safety (in the Western world) have actually improved since the 1970s, despite substantial transport volume growth. Related to the other external transport issues mentioned in these sections, the new technologies have merely slowed down the growth problem. We can make three observations about the nature of the new technologies and their dynamics.

Firstly, we can conclude that all important technological changes to date are optimizing technologies (also called incremental technologies). These kinds of technologies do not change the existing system but improve the performance of the existing dominant transport technology. Examples are the engine fuel-efficiency improvements, the three-way catalysts, seat belts and porous asphalt. Redesign is another innovation type where a part of an existing product or process is radically changed. The system as a whole is kept the same but certain parts are radically changed. Examples are fuel cell technology, advanced ACC or electric cars. Redesign innovations have not been popular so far but may be seen more often in the future, as shown in the previous sections. The conditions for successful adoption of these redesign innovations seem to be whether the transport related problems are perceived as 'a large problem', and the redesign innovations must be technologically, economically, socially and politically feasible (Figure 8.4).

Secondly, we can observe that it takes years before effective new technologies are implemented. In the case of a success story (air pollution), the time between the first recognition of the problem and the implementation of highly effective technologies (for example, the ones which meet the strict

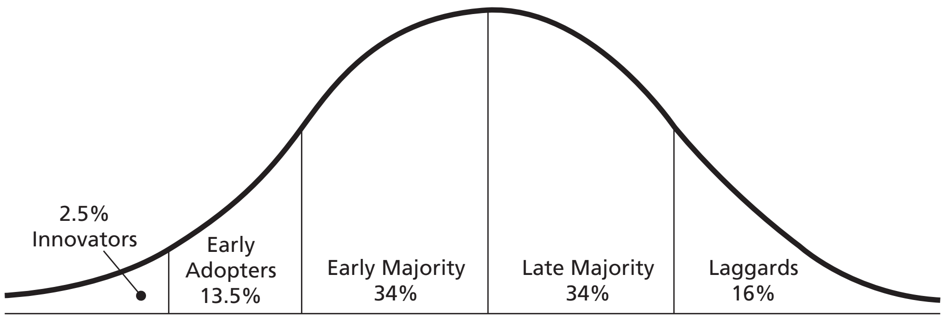


Figure 8.10 The diffusion of innovations according to Rogers (2003)

Euro 5 and 6 standards) amounts to approximately 50 years. The explanation for slow progress can be based on the Feitelson and Salomon framework (Figure 8.4). They stress that technological, economic, social and political feasibility is required for adopting innovations. So, perhaps 50 years ago the really effective air quality improving technologies were not available or economically feasible. On the other hand, governments could have forced vehicle manufacturers to invest more in research on cleaner technologies to speed up the process by imposing strict emission standards much earlier compared to when they actually took place. However, this was politically not feasible. Thus the dynamics of large scale adoption of new transport technologies aimed at reducing external transport issues, as described in this chapter, are mainly politically driven.[8] This adoption dynamic is different from the famous theory of Rogers (2003), who looked at the way individuals adopt innovations. Rogers distinguishes five classes in the adoption process of innovations: innovators, early adopters, early majority, late majority and laggards (Figure 8.10).

Some features of this model can be found in sections 8.4 to 8.8, especially relating to the 'early adopters'. Hybrid cars, advisory speed warning systems, seat belts and low-air-resistant cars are examples of innovations which were first used by some specific car makers, even before policies were in place to stimulate the use of these innovations. However, in the case of the transport technology innovations covered in this chapter, after the 'early adopters' period, often a political process takes over the further adoption process by implementing (stricter) standards or making certain technologies mandatory. And all of these political processes take time.

Thirdly, we can note that, for congestion, noise, fossil fuel use and CO_2 emissions, it seems difficult to reduce absolute levels of the negative outcomes of the current dominant transport technology. This is difficult because transportation volumes keep on growing. Thus optimizing technological changes can increase road capacity or can make vehicles more silent, more fuel-efficient (more mpg) and less carbon intensive (fewer grams of CO_2 emissions per miles driven). However, if at the same time people drive more miles, the congestion and noise levels, the amount of gallons used and the amount of carbon emitted could still increase.

8.10 Conclusions

The conclusions of this chapter are:

1. In the Western world, because of technological innovations, air pollution (NO_x and PM_{10}) and casualties caused by transport have strongly decreased over time, despite much higher transport volumes.
2. For climate change (CO_2 emissions), oil dependency, noise and congestion, technological innovations have merely slowed the growth of these problems in the recent past.
3. In the period 1960–2011, optimizing technologies have been applied. These technologies do not change the existing system, but improve the performance of the existing dominant transport technology. In the future, perhaps, redesign technical innovations will also be seen, and this would include electric vehicles.
4. For all transport related problems in this chapter, government policies have played a major role in achieving technological changes. This has been particularly important when implementing new or stricter vehicle or fuel standards or norms.
5. Broadly speaking, transport vehicles are relatively difficult to regulate. In every section that addressed the policy of setting emissions, fuel, noise or safety standards, the problem of the differences – often large differences – between the actual, real-world and the test results was noted. In the tests, the technically adapted vehicles performed well according to the requirements, but in reality they did worse.
6. It is difficult to change the dominant transportation technology. This can be explained with the concept of 'lock-in' from the theory of evolutionary economics or with the notion of transport as a socio-technical system from the theory of system dynamics. Socio-technical systems theory points at the way transport technology is embedded in society in terms of physical infrastructure, institutions, market and culture.
7. The way transport technology is embedded in society means that, for technological change, more than a problem and a technology which might solve the problem are needed. Adoption of a technological

innovation aimed at decreasing transportation's negative external effects is dependent on technical, economic, social and political feasibility.

NOTES

1 Maglev stands for magnetic levitation and propulsion. Essentially, the idea is to use frictionless trains that are levitated and propelled by the force of magnetic fields. Germany, South Korea and Japan have experimented with Maglev technology. In Shanghai, China, a 30 km Maglev track is in operation. For a review of high-speed rail, including Maglev technology, see Givoni (2006).

2 The most cited modern book in this area of economic research is by Nelson and Winter (1982). We will now only summarize some important notions (based on van den Bergh et al., 2005).

3 Before the 1960s car engines were vented to the atmosphere. That is, toxic vapours were simply allowed to flow out of the engine. Positive crankcase ventilation (or PCV) is a technology which prevents this from happening.

4 It is customary to denote the emission standards as Euro 1, 2, 3 and so forth for cars and light duty vehicles, and as Euro I, II, III for heavy duty vehicles.

5 Miles per gallon (mpg) is the inverse of litre per kilometre (l/km). High mpg values mean fuel-efficient vehicles. However, high l/km values mean relatively lower fuel-efficient vehicles. A value of 20 mpg equals 11.8 l/km.

6 *Environmental Technical Manual on the Use of Procedures in the Noise Certification of Aircraft* (Doc. 9501 of ICAO). The annex is Annex 16, 'Environmental protection', vol. I.

7 In first-best regulatory road charges the tax can be regulated according to the various dimensions affecting the actual marginal external costs of each trip, such as the length of the trip, the time of driving, the route followed and the vehicle used (Verhoef et al., 1995).

8 This is somewhat different compared to an often used distinction between innovation processes (van der Duin, 2006): 'technology push' processes versus 'market pull' processes. 'Technology push' processes were the first generation of innovation processes in the 1950s and mid-1960s when scientific researchers within the R&D department of an organization were pretty much free to investigate any subject they wanted. The ensuing ideas or inventions often had a technological character and were more or less pushed on to the market. In the later 'market pull innovation processes' the R&D processes were initiated by market requirements.

REFERENCES

ACEA (2011), European Automobile Manufacturers' Association, accessed 30 August 2011 at http://www.acea.be/images/uploads/files/20100902_2010_KEY_FIGURES_1_Employment.pdf.

Allsop, R. (2010), *The Effectiveness of Speed Cameras: A Review of Evidence*, London: RAC Foundation.

Ambak, K., R. Atiq and R. Ismail (2009), 'Intelligent transport system for motorcycle safety and issues', *European Journal of Scientific Research*, 28 (4), 600–611.

Ausubel, J.H., C. Marchetti and P. Meyer (1998), 'Toward green mobility: the evolution of transport', *European Review*, 6 (2), 137–156.

Berglund, B., T. Lindvall and D.H. Schwela (eds) (1999), *Guidelines for Community Noise*, Geneva: World Health Organization (WHO).

Bluhm, G., E. Nordling and N. Berglind (2004), 'Road traffic noise and annoyance: an increasing environmental health problem', *Noise and Health*, 6 (24), 43–49.

Bureau, J.-C., A.-C. Disdier, C. Gauroy and D. Treguer (2010), 'A quantitative assessment of the determinants of the net energy value of biofuels', *Energy Policy*, 38, 2282–2290.

California Environmental Protection Agency, Air Resources Board (2011), *Key Events in the History of Air Quality in California*, accessed February 2011 at http://www.arb.ca.gov/html/brochure/history.htm.

Carsten, O.M.J. and F.N. Tate (2005), 'Intelligent speed adaptation: accident savings and cost–benefit analysis', *Accident Analysis and Prevention*, 37, 407–416.

Christensen, J., K.P. Glaeser, T. Shelton, B. Moore and L. Aarts (2010), 'Innovation in truck technologies', Discussion Paper 2010-10, Organisation for Economic Co-operation and Development (OECD)/International Transport Forum (ITF), Paris.

Descornet, G. and L. Goubert (2006), *Noise Classification of Road Pavements, Task 1: Technical Background Information*, Kongens Lyngby: COWI.

Dieselnet (2011), accessed February 2011 at http://www.dieselnet.com.

ECMT (2006), *Reducing NO_x Emissions on the Road: Ensuring Future Exhausts Emission Limits Deliver Air Quality Standards*, Paris: European Conference of Ministers of Transport, OECD.

ECMT (2007), *Managing Urban Traffic Congestion*, Paris: European Conference of Ministers of Transport, OECD.

EEA (2009), *Transport at a Crossroads. TERM 2008: Indicators Tracking Transport and Environment in the European Union*, EEA Report No 3/2009, Copenhagen: European Environment Agency.

EEA (2011), European Environmental Agency, Copenhagen, accessed February 2011 at http://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-air-pollutants/transport-emissions-of-air-pollutants-2#toc-0.

EEA (2011), European Environmental Agency, Copenhagen, accessed November 2011 at http://www.eea.europa.eu/data-and-maps/figures/overview-of-total-annual-subisidies.

Elzen, B., F.W. Geels and K. Green (2004), *System Innovation and the Transition to Sustainability: Theory, Evidence and Policy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

EURO NCAP (2011), *History*, accessed February 2011 at http://www.euroncap.com/history.aspx.

Feitelson, E. and I. Salomon (2004), 'The political economy of transport innovations', in M. Beuthe, V. Himanen, A. Reggiani and L. Zamparini (eds), *Transport Developments and Innovations in an Evolving World*, Berlin: Springer.

Geels, F.W. (2002), 'Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study', *Research Policy*, 31, 1257–1274.

Geels, F. (2005), *Technological Transitions and System Innovations: A Co-Evolutionary and Socio-Technical Analysis*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Givoni, M. (2006), 'Development and impact of the modern high-speed train: a review', *Transport Reviews*, 26 (5), 593–611.

Gruebler, A. (1990), *The Rise and Fall of Infrastructure: Dynamics of Evolution and Technological Change in Transport*, Heidelberg: Physica.

Hausberger, S., M. Rexeis, M. Zallinger and R. Luz (2009), *Emission Factors from the Model PHEM for the HBEFA Version 3*, Graz: Graz University of Technology.

Heinz, S. (2005), *Ermittlung der Geräuschemission von Kfz im Straßenverkehr*, Würselen: TÜV Nord Mobilität.

High Speed Two (HS2) Limited (2010), *High Speed Rail London to the West Midlands and Beyond, HS2 Demand Model Analysis*, London: HS2 Limited.

ICAO (2008), *ICAO Environmental Report 2007*, International Civil Aviation Organization, accessed at http://www.icao.int/env/pubs/env_report_07.pdf.

ICAO (2010), *ICAO Environmental Report*, International Civil Aviation Organization, accessed February 2011 at http://www.icao.int/icao/en/env2010/Pubs/EnvReport2010/ICAO_EnvReport10-Outlook_en.pdf.

IEA/OECD (2009), *Transport, Energy and CO_2: Moving towards Sustainability*, Paris: International Energy Agency/Organisation for Economic Co-operation and Development.

IFEU (2007), *Nachwachsende Rohstoffe für die chemische Industrie: Optionen und Potenziale für* *die Zukunft* [Renewable resources for the chemical industry: options and potentials for the future], Frankfurt am Main: Fachvereinigung Organische Chemie im VCI [German Chemical Industry Association].

Ison, S., G. Hughes and R. Tuckwell (2008), 'Cambridge's experience of road user charging: lessons learned', *Transport Proceedings of the Institution of Civil Engineers*, 161, 135–141.

de Kievit, M., K. Malone, H. Zwijnenberg and B. van Arem (2008), 'Benchmarking promotion and deployment activities regarding intelligent vehicle systems in the EU', paper presented at the 2008 Annual Meeting of the Transportation Research Board, Washington, DC.

Kotzen, B. and C. English (1999), *Environmental Noise Barriers: A Guide to Their Acoustic and Visual Design*, Oxford: Taylor & Francis.

Levinson, M. (2008), *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger*, Princeton, NJ: Princeton University Press.

Ligterink, N.E. and B. Bos (2010), *CO_2 uitstoot van personenwagens in norm en praktijk: analyse van gegevens van zakelijke rijders* [Passenger cars' CO_2 emission in test and practice: analysis based on business drivers], TNO-report MON-RPT-2010-00114, Delft: TNO.

Ligterink, N., R. de Lange, R. Vermeulen and H. Dekker (2009), *On-Road Emissions of Euro-V Trucks*, MON-RPT-033-DTS-2009-03840, Delft: TNO.

Liska, A.J., H.S. Yang, V.R. Bremer, T.J. Klopfenstein, D.T. Walters, G.E. Erickson and K.G. Cassman (2009), 'Improvements in life cycle energy efficiency and greenhouse gas emissions of corn-ethanol', *Journal of Industrial Ecology*, 13 (1), 58–74.

McFadden, D. (2007), 'The behavioral science of transportation', *Transport Policy*, 14, 269–274.

Meadows, D.H., D.L. Meadows, J. Randers and W.W. Behrens III (1972), *The Limits to Growth*, New York: Universe Books.

Mellios, G., C. Samaras and L. Ntziachristos (2011), 'Projection of CO_2 emissions from road transport', Report no. 10.RE.0034.V4, University of Thessaloniki.

Menichetti, E. and M. Otto (2009), 'Energy balance and greenhouse gas emissions of biofuels from a life-cycle perspective', in R.W. Howarth and S. Bringezu (eds), *Biofuels: Environmental Consequences and Interactions with Changing Land Use*, Scientific Committee on Problems of the Environment (SCOPE) International Biofuels Project Rapid Assessment, accessed at http://cip.cornell.edu/biofuels/S, pp. 81–109.

Mitchell, P. (2009), 'Speed and road traffic noise: the role that lower speeds could play in cutting noise from traffic', report commissioned by the UK Noise Association.

Nader, R. (1965), *Unsafe at Any Speed: The Designed-In Dangers of the American Automobile*, New York: Grossman Publishers.

Nakicenovic, N. (2008), 'The mobility drive', *Elektrotechnik und Informationstechnik*, 125 (11), 362–366.

Nelson, R.R. and S.G. Winter (1982), *An Evolutionary Theory of Economic Change*, Cambridge, MA: Belknap Press of Harvard University Press.

OECD/ITF (2008), 'Oil dependence: is transport running out of affordable fuel?', Discussion Paper no. 2008-5, Organisation for Economic Co-operation and Development/International Transport Forum, accessed at http://www.internationaltransportforum.org/jtrc/DiscussionPapers/DP200805.pdf.

Oliver Wyman (2007), *Car Innovation 2015: A Comprehensive Study on Innovation in the Automotive Industry*, New York: Oliver Wyman Automotive.

Pachauri, R.K. and A. Reisinger (eds) (2007), *Climate Change 2007*, Synthesis Report, Geneva: Intergovernmental Panel on Climate Change.

PBL (2010), *Balans van de leefomgeving 2010* [Environmental balance 2010], Bilthoven: Planbureau voor de Leefomgeving.

Rogers, E.M. (2003), *The Diffusion of Innovations*, 5th edition, New York: Free Press.

Schafer, A. (1998), 'The global demand for motorized mobility', *Transportation Research A*, 32 (6), 455–477.

Schafer, A. (2006), 'Long-term trends in global passenger mobility', *The Bridge*, 36 (4), 24–32.

Schaller, B. (2010), 'New York City's congestion pricing experience and implications for road pricing acceptance in the United States', *Transport Policy*, 17, 266-273.

Schipper, L. (2011), 'Automobile use, fuel economy and CO_2 emissions in industrialized countries: encouraging trends through 2008', *Transport Policy*, 18, 358–372.

Shah, A.A. and L.J. Dal (2007), 'Intelligent transportation systems in transitional and developing countries', *IEEE A&E Systems Magazine*, 27–33.

Sivak, M. and O. Tsimhoni (2009), 'Fuel efficiency of vehicles on US roads: 1923–2006', *Energy Policy*, 37, 3168–3170.

Smeed, R.J. (1964), *Road Pricing: The Economic and Technical Possibilities*, London: HMSO.

Sovacool, B.K. (2009), 'Early modes of transport in the United States: lessons for modern energy policymakers', *Policy and Society*, 27 (4), 411–427.

Sperling, D. and D. Gordon (2009), *Two Billion Cars: Driving towards Sustainability*, Oxford: Oxford University Press.

Stewart, J. (2011), accessed February 2011 at http://videocafe.crooksandliars.com/david/stewart-teaches-history-lesson-oil-dependenc.

T&E (2009), *Reducing CO_2 Emissions from New Cars: A Study of Major Car Manufacturers' Progress in 2008*, Brussels: T&E – European Federation for Transport and Environment AiSBL.

Techmer, A. (2007), 'Application development of camera-based driver assistance systems on a programmable multi-processor', *Architect Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, FrE1.9, Istanbul, pp. 1211–1216.

UNEP (2009), *Assessing Biofuels: Towards Sustainable Production and Use of Resources*, Geneva: United Nations Environment Programme.

US Department of Transportation (2010), *Transportation's Role in Reducing U.S. Greenhouse Gas Emissions*, Report to Congress, 29 April, Washington, DC: Secretary of Transportation.

Van Arem, B., C.J.G van Driel and R. Visser (2006), 'The impact of cooperative adaptive cruise control on traffic flow characteristics', *IEEE Transactions on Intelligent Transportation Systems*, 7 (4), 429–436.

Van den Bergh, J.C.J.M., A. Faber, A.M. Idenburg and F.H. Oosterhuis (2005), *Survival of the Greenest: Evolutionaire economie als inspiratie voor energie- en transitiebeleid* [Survival of the greenest: evolutionary economy as inspiration for energy and transition policy], RIVM-rapport 550006002/2005, Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Van den Brink, R.M.M., and B. van Wee (2001), 'Why has car-fleet specific consumption not shown any decrease since 1990? Quantitative analysis of Dutch passenger car-fleet specific fuel consumption', *Transportation Research D*, 6 (2), 75–93.

Van der Duin, P. (2006), *Qualitative Futures Research for Innovation*, Delft: Eburon Academic Publishers.

Van der Toorn, J.D., T.C. Dool and F. de Roo (1997), *Geluidemissie door motorvoertuigen: klassieke metingen met de Syntakan* [Noise emissions from motor vehicles: classic measurements using Syntaken], report TPD-HAG-RPT-950033, Delft: TNO-TPD.

Van der Toorn, J.D. and J.W.A. van Vliet (2000), *Emissiekentallen motorvoertuigen 2000, gebaseerd op metingen uit 1996 en 1999* [Noise emission indicators for motor vehicles 1996 and 1999], Report HAG-RPT-000048. Delft: TNO-TPD.

Van Driel, C.J.G. and B. van Arem (2010), 'The impact of a congestion assistant on traffic flow efficiency and safety in congested traffic caused by a lane drop', *Journal of Intelligent Transportation Systems*, 14 (4), 197–208.

Verhoef, E.T. (1996), *The Economics of Regulating Road Transport*, Cheltenham, UK and Brookfield, VT, USA: Edward Elgar.

Verhoef, E., P. Nijkamp and P. Rietveld (1995), 'Second-best regulation of road transport externalities', *Journal of Transport Economics and Policy*, 29 (2), 147–167.

Vonk Noordegraaf, D.M., B.M.R. Heijligers, O.A.W.T. van de Riet and G.P. van Wee (2009), 'Technology options for distance-based user charging schemes', paper presented at the Transportation Research Board (TRB) 88th Annual Meeting, Washington, DC, January 11–15, 2009.

Vonk Noordegraaf, D.M., J.A. Annema and O.A.M.T. van de Riet (2012), 'The policy implementation process for road pricing in the Netherlands', in H. Geerlings, Y. Shiftan and D. Stead (eds), *Transition Towards Sustainable Mobility: the Role of Instruments, Individuals and Institutions*, Farnham: Ashgate.

WHO (2004), *World Report on Road Traffic Injury Prevention*, Geneva: World Health Organization.

Williams, J.R. and M. Weiss (2005), 'Competitive innovation in the U.S. airline industry', Tepper School of Business, Paper 551, accessed August 2011 at http://repository.cmu.edu/tepper/551.

Worldometers (2011), accessed February 2011 at http://www.worldometers.info/cars/.

Worley, H. (2006), 'Road Traffic Accidents Increase Dramatically Worldwide', publication retrieved from http://www.prb.org/Articles/2006/RoadTrafficAccidentsIncreaseDramatically Worldwide.aspx.

9

Accessibility: perspectives, measures and applications

Karst Geurs and Bert van Wee

9.1 Introduction

A principal goal of transport policy is to improve accessibility: the transport system should allow people to travel and participate in activities, and firms to transport goods between locations (from mining, via stages of production, to distribution centres and finally to clients, such as shops or other firms). Several authors have written review articles on accessibility measures, focusing on certain perspectives, such as location accessibility (e.g. Song, 1996; Handy and Niemeier, 1997), individual accessibility (e.g. Pirie, 1979; Kwan, 1998), economic benefits of accessibility (e.g. Koenig, 1980; Niemeier, 1997) or other, different perspectives (Geurs and van Wee, 2004).

However, despite the crucial role of accessibility in transport policy making throughout the world, accessibility is often a misunderstood, poorly defined and poorly measured construct. Accessibility is defined and operationalized in several ways, and thus has taken on a variety of meanings. Gould (1969) noted that one of the problems with accessibility is that 'accessibility is a slippery notion . . . one of those common terms that everyone uses until faced with the problem of defining and measuring it'. Indeed, defining and operationalizing accessibility can be rather complex. This is problematic because the choice and operationalization of an accessibility measure may strongly affect the conclusions on accessibility. For example, Linneker and Spence (1992) illustrated that inner London has the highest access costs (in terms of time and vehicle operation costs) in the UK, but also the highest level of potential accessibility to jobs, despite the high travel cost.

Handy and Niemeier (1997) have stated that 'a distinct gap currently exists between the academic literature and the practical application of accessibility measures. It is important that accessibility measures used in practice are theoretically and behaviourally sound and that innovative approaches to measuring accessibility are made practical.' This statement is still valid today. Land-use and infrastructure policy plans are often evaluated with accessibility measures that are easy to interpret for researchers and policy makers, such as congestion levels or travel speed on the road network, but which have strong methodological disadvantages. Theoretically sound accessibility measures typically involve huge amounts of data or complex transport models, which may restrict analysis to a relatively small region or regions or countries where advanced transport models are available.

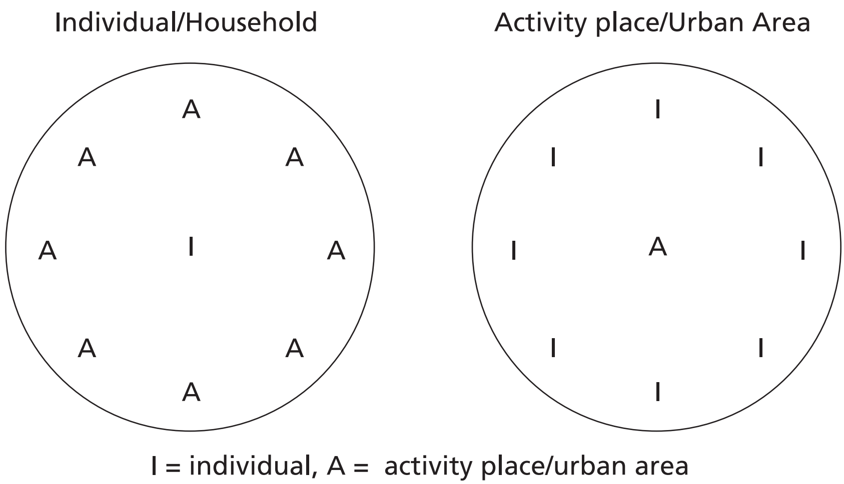
In this chapter we describe the different perspectives on accessibility (section 9.2), the different components of accessibility (section 9.3), the different means by which accessibility can be operationalized (section 9.4) and the different criteria for choosing accessibility measures (section 9.5). Two examples of accessibility measures – potential accessibility and logsum accessibility – are described in section 9.6, and section 9.7 presents the conclusions and future trends in accessibility studies. Chapter 8 has already discussed the relevance of some key technologies for accessibility.

9.2 Perspectives on accessibility

As already noted, accessibility is defined and operationalized in several ways and thereby has taken on a variety of meanings. These include such well-known definitions as 'the potential of opportunities for interaction' (Hansen, 1959), 'the ease with which any land-use activity can be reached from a location using a particular transport system' (Dalvi and Martin, 1976), 'the freedom of individuals to decide whether or not to participate in different activities' (Burns, 1979) and 'the benefits provided by a transportation/land-use system' (Ben-Akiva and Lerman, 1979).

Here, accessibility measures are interpreted as indicators for the impact of land-use and transport developments and policy plans on the functioning of the society in general. This means that accessibility should relate to the role of the land-use and transport systems in society which will give individuals or groups of individuals the opportunity to participate in activities in different locations. Subsequently, we define accessibility as:

> The extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s) at various times of the day (*perspective of persons*), and the extent to which land-use and transport systems enable companies, facilities and other activity



Source: Dijst et al. (2002).

Figure 9.1 Individual and location perspective on accessibility

places to receive people, goods and information at various times of the day (*perspective of locations of activities*).

The terms 'access' and 'accessibility' in the literature are often used indiscriminately. Here, 'access' is used when talking about a person's perspective: the area that a person can reach from his or her origin location to participate in one of more activities at destination locations at certain times. The size of the area depends, for example, on the time, costs and effort that an individual is willing to accept (the transportation component of accessibility; see section 9.3). From the perspective of location, the 'accessibility' is the catchment area where people, goods and information are that can access the destination location from a certain origin location. The size of the catchment area also depends on the time, cost and effort acceptable to bridge the distance from origin to destination (Figure 9.1).

9.3 Components of accessibility

Four components of accessibility can be distinguished: a land-use, transportation, temporal and individual component (Geurs and van Wee, 2004):

1. **The land-use component** reflects the land-use system, consisting of (a) the amount, quality and spatial distribution opportunities supplied at each destination (jobs, shops, health, social and recreational facilities,

etc.), (b) the demand for these opportunities at origin locations (e.g. where inhabitants live), and (c) the confrontation of supply of and demand for opportunities, which may result in competition for activities with restricted capacity such as job and school vacancies and hospital beds (see van Wee et al., 2001). See also Chapter 5.

2. **The transportation component** describes the transport system, expressed as the disutility for an individual to cover the distance between an origin and a destination using a specific transport mode; included are the amount of time (travel, waiting and parking), costs (fixed and variable) and effort (including reliability, level of comfort, accident risk, etc.). This disutility results from the confrontation between supply and demand. The supply of infrastructure includes its location and characteristics (e.g. maximum travel speed, number of lanes, public transport timetables, travel costs). The demand relates to both passenger and freight travel. See also Chapter 6.
3. **The temporal component** reflects the temporal constraints, that is, the availability of opportunities at different times of the day and the time available for individuals to participate in certain activities (e.g. work, recreation). See also Chapter 3.
4. **The individual component** reflects the needs (depending on age, income, educational level, household situation, etc.), abilities (depending on people's physical condition, availability of travel modes, etc.) and opportunities (depending on people's income, travel budget, educational level, etc.) of individuals. These characteristics influence a person's level of access to transport modes (e.g. being able to drive and borrow or use a car) and spatially distributed opportunities (e.g. having the skills or education to qualify for jobs near the home residential area) and may strongly influence the total aggregate accessibility result. Several studies (e.g. Cervero et al., 1997; Shen, 1998; Geurs and Ritsema van Eck, 2003) have shown that, in the case of job accessibility, inclusion of occupational matching strongly affects the resulting accessibility indicators. See also Chapter 3.

The different components have a direct influence on accessibility but also indirectly through interactions between the components. For example, the land-use component (distribution of activities) is an important factor determining travel demand (transport component) and may also introduce time restrictions (temporal component) and influence people's opportunities (individual component). The individual component interacts with all other components: a person's needs and abilities that influence the (valuation of) time, cost and effort of movement, types of relevant activities and the times at which one engages in specific activities.

9.4 Operationalization of accessibility measures

Following our definition of accessibility, an accessibility measure should ideally take all components and elements within these components into account. In practice, applied accessibility measures focus on one or more components of accessibility, depending on the perspective taken. There are four basic types of accessibility measures generally used:

1. **Infrastructure-based accessibility measures**, analysing the (observed or simulated) performance or service level of transport infrastructure, such as the length of infrastructure networks, the density of those networks (e.g. kilometre road length per square kilometre), level of congestion, and average travel speed on the road network. This type of accessibility measure is typically used in transport planning. Some of these measures focus only on the supply of infrastructure, while others also use demand factors.
2. **Location-based accessibility measures**, analysing accessibility at locations, typically on a macro-level. The measures describe the level of accessibility to spatially distributed activities, such as 'the number of jobs within 30 minutes' travel time from origin locations'. More complex location-based measures explicitly incorporate capacity restrictions of supplied activity characteristics to include competition effects.
3. **Person-based accessibility measures**, analysing accessibility at the individual level, such as 'the activities in which an individual can participate at a given time'. This type of measure is founded in the space–time geography (Hägerstrand, 1970) that measures limitations on an individual's freedom of action in the environment, that is, the location and duration of mandatory activities, the time budgets for flexible activities and travel speed allowed by the transport system. See also Chapter 3.
4. **Utility-based accessibility measures**, analysing the (economic) benefits that people derive from access to the spatially distributed activities. This type of measure has its origin in economic studies and is increasingly receiving attention in accessibility studies (e.g. de Jong et al., 2007; Geurs et al., 2010).

Table 9.1 presents an overview of different types of accessibility measures, applications and examples, with brief comments on the advantages and disadvantages of the measures used.

The different accessibility perspectives focus on different components of accessibility, often ignoring other relevant elements of accessibility.

Table 9.1 Accessibility indicators, applications and examples

Accessibility type	Applications	Examples	Disadvantages and comments
Infrastructure-based accessibility measures:			
Supply-oriented measures – network level	Description and comparison of characteristics of infrastructure supply in a region or country	Length of motorways, density of rail network	Partial measure of accessibility; does not include land-use and individual components of accessibility, e.g. it does not say anything about the number of opportunities that can be accessed.
Supply-oriented measures – connectivity of locations to transport networks	Analysis of how well locations are connected to transport networks	Distance to nearest railway station, exit point of a motorway	Partial measure of accessibility; measures are not suited for a comparison of transport modes, taking available opportunities into account.
Supply-oriented measures – network connectivity	Describing network connectivity, expressing how well each node in a network is connected to each adjacent node	Connectivity or centrality of a node relative to the rest of the network	Partial measure of accessibility. It also does not provide plausible results in complex networks with many indirect linkages between nodes.
Demand- and supply-oriented measures	Describing actual quality of performance of infrastructure networks	Actual travel times on the road network	Partial measure of accessibility; does not include land-use and individual components of accessibility, e.g. it does not say anything about the number of opportunities that can be accessed.
Location-based accessibility measures:			
Cumulative opportunities	Counts the number of opportunities that can be reached from an original location within a given travel time, distance or cost (fixed costs); or a measure of the (average or total) time or cost required to access a fixed number of opportunities (fixed opportunities)	Number of jobs within 30 minutes' travel time by car; average travel time or cost to reach 1 million jobs	These measures are relatively undemanding of data and are easy to interpret for researchers and policy makers, as no assumptions are made on a person's perception of transport, land use and their interaction. The measure is extremely sensitive to travel time changes and is not suited to describing accessibility developments in time.
Potential or gravity-based accessibility	Estimates the number of opportunities in destination locations that	Index of jobs, population or services which	The measure evaluates the combined effect of land-use and transport elements, and

Table 9.1 (continued)

Accessibility type	Applications	Examples	Disadvantages and comments
Location-based accessibility measures:			
	can be accessed from an original location, weighted by a distance decay function, which describes how more distant opportunities provide diminishing influences	can be accessed from an original location	incorporates assumptions on a person's perceptions of transport by using a distance decay function. The measure has no meaning in absolute terms (index). For plausible results, the form of the function should be carefully chosen and the parameters should be estimated using empirical data on travel behaviour in the study area.
Actual accessibility	Estimates total travel distances, times or costs from an original location to all destinations, weighted by the actual number of trips on an original destination location	Analysis of competition between different transport modes	Detailed information of spatial patterns of travel behaviour is needed.
Person-based accessibility measures:			
Space–time approach	The measures analyse accessibility from the viewpoint of individuals, incorporating spatial and temporal constraints	The number of household activity programmes that can be carried out by individuals, given personal and time constraints	Founded in space–time geography. Measure is theoretically advanced but is very data demanding.
Utility-based accessibility measures:			
Utility of accessibility	The measures estimate the utility or monetary value (when utility is converted into monetary terms)	Logsum accessibility describing the direct economic benefits of having access to spatially distributed activities	Founded in microeconomic theory. More difficult to communicate to non-experts.

Source: Based on van Wee et al. (2001); Geurs and van Wee (2004).

Table 9.2 Types of accessibility measures and components

Measure	Component			
	Transport component	Land-use component	Temporal component	Individual component
Infrastructure-based measures	Travelling speed; vehicle-hours lost in congestion		Peak-hour period; 24-hour period	Trip-based stratification, e.g. home to work, business
Location-based measures	Travel time and/or costs between locations of activities	Amount and spatial distribution of the demand for and/or supply of opportunities	Travel time and costs may differ, e.g. between hours of the day, days of the week, or seasons	Stratification of the population (e.g. by income, educational level)
Person-based measures	Travel time between locations of activities	Amount and spatial distribution of supplied opportunities	Temporal constraints for activities and time available for activities	Accessibility is analysed at individual level
Utility-based measures	Travel costs between locations of activities	Amount and spatial distribution of supplied opportunities	Travel time and costs may differ, e.g. between hours of the day, days of the week, or seasons	Utility is derived at the individual or homogeneous population group level

Note: Dark shading: primary focus of measures; light shading: non-primary focus.

Source: Geurs and van Wee (2004).

Table 9.2 presents a matrix of the different accessibility measures and components. Infrastructure-based measures do not include a land-use component; that is, they are not sensitive to changes in the spatial distribution of activities if service levels (e.g. travel speed, times or costs) remain constant. The temporal component is explicitly treated in person-based measures and is generally not considered in the other perspectives, or is treated only implicitly, for example by computing peak- and off-peak-hour accessibility levels. Person-based and utility-based measures typically focus on the individual component, analysing accessibility on an individual level. Location-based measures typically analyse accessibility on a macro-level but focus more on incorporating spatial constraints in the supply of opportunities, usually excluded in the other approaches (see the dark-shaded cells in Table 9.2).

To operationalize accessibility measures, the most suitable type of accessibility measure needs to be chosen (the columns in Table 9.2), and then the different elements within the different components need to be determined (the rows in Table 9.2). A few examples can illustrate this process:

1. In determining travel times between origin and destination locations, one can choose whether or not to weigh the different time components of a trip, such as access and egress times to and from boarding points, in-vehicle travel times, waiting times and so on. Generally speaking, access and egress and waiting time will incur greater disutility to travellers than in-vehicle time (e.g. see Balcombe et al., 2004).
2. In determining the costs of car trips, one can include only fuel costs, but also total variable costs, including for example parking costs and fixed costs (e.g. depreciation of the car).
3. One can use either objective costs or perceived costs, which may differ greatly (e.g. see van Exel and Rietveld, 2009).
4. In determining the land-use component, one needs to consider whether available opportunities have capacity limitations (such as in the case of school locations and health-care facilities), and where accessibility measures need to account for differences in the spatial distribution of the demand and supply of these opportunities (competition effects).

9.5 Choosing and using accessibility measures

In defining and operationalizing accessibility, there is no one best approach because different situations and purposes demand different approaches (Handy and Niemeier, 1997). However, several criteria can be derived to evaluate the usefulness and limitations of accessibility measures for different study purposes. Such criteria can, for example, be found in Black and Conroy (1977), Jones (1981), Handy and Niemeier (1997) and Geurs and van Wee (2004).

Purpose of the study

This is the starting point of the operationalization process. What is the purpose of the study and, following from that, what is the main reason for analysing accessibility? All other choices essentially follow on from this. The definition and operationalization would, for example, strongly differ when the study purpose is to evaluate accessibility impacts of a transport project, or to analyse social equity effects, or the economic benefits that people derive from having access to opportunities. This means that the analysis of transport policy can be carried out through more aggregate, location-based accessibility measures, whereas the analysis of social equity effects requires a highly spatially differentiated and disaggregated analysis. The analysis of economic benefits would require choosing a utility-based accessibility measure that is directly linked to microeconomic theory.

Scientific quality

An accessibility measure should ideally take all of the components and elements within these components into account (section 9.2). Thus an accessibility measure should firstly be sensitive to the changes in the transport and land-use systems and the temporal constraints of opportunities, and it should take individual needs, abilities and opportunities into account. Geurs and van Wee (2004) derived the following five criteria which an accessibility measure should behave in accordance with, keeping all other conditions constant:

1. If the service level (travel time, costs, effort) of any transport mode in an area increases (decreases), accessibility to any activity in that area or from any point within that area should increase (decrease).
2. If the number of opportunities for an activity increases (decreases) anywhere, accessibility to that activity should increase (decrease) from any place.
3. If the demand for opportunities for an activity with certain capacity restrictions increases (decreases), accessibility to that activity should decrease (increase).
4. An increase in the number of opportunities for an activity at any location should not alter the accessibility to that activity for an individual (or groups of individuals) not able to participate in that activity given the time budget.
5. Improvements in one transport mode or an increase of the number of opportunities for an activity should not alter the accessibility to any individual (or groups of individuals) with insufficient abilities or capacities (e.g. driving licence, educational level) to use that mode or participate in that activity.

These criteria should not be regarded as absolute, but more in the line of what accessibility studies should strive for. Applying the full set of criteria would imply a level of complexity and detail that can probably never be achieved in practice. However, it is important that the implications of violating one or more theoretical criteria should be recognized and described.

Operationalization (cost, ease)

This is the ease with which the measure can be used in practice, for example in ascertaining availability of data, models and techniques, and time and budget. This criterion will usually be in conflict with one or more of the theoretical criteria described above.

Interpretability and communicability

The literature shows a trend towards more complex and disaggregated accessibility measures, partly in response to the recognition that the aggregate measures lack many important details. However, increased complexity increases the effort for calculations and the difficulty of interpretation. Clearly, researchers, planners and policy makers should be able to understand and interpret the measure, and communicate results to clients, as otherwise it is not likely to be used in evaluation studies of land-use and/or transport developments or policies and will thus have no impact on the policy making process.

The interpretations of more complex accessibility measures can firstly be improved by comparing accessibility across place or time, or both place and time, rather than focusing on absolute levels of accessibility. Secondly, the interpretation can be much improved by estimating the separate influence of the different components of accessibility. For example, Geurs and Ritsema van Eck (2003) computed the separate influence of land-use changes, infrastructure investments and congestion on the development of (job) accessibility for the Netherlands. Computation of the different components of accessibility facilitates both the explanation of overall accessibility changes and the relative position of regions. Thirdly, the more complex utility-based accessibility measures can be expressed in monetary values, and this strongly improves the interpretation and communication to planners and policy makers. Fourthly, for measures that are difficult to interpret, such as utility-based measures, the output could be indexed. For example, the base year value or a reference scenario can be indexed at the level of 100. The value of the accessibility indicators could then be indexed and compared to this base level value.

However, there are no guarantees that accessibility measures will be used in public policy even when they are easy to interpret and communicate. Pirie (1981) clearly points out that there is no guaranteed or easy transition from accessibility research to the formulation of public policy and its implementation; public policy on accessibility will only be forthcoming if accessibility is a well-politicized issue.

9.6 Two examples of accessibility measures

Potential accessibility measures

Several types of location-based measures are used in accessibility studies. The distinguishable groups of measures are distance based and contour based, along with potential measures and the balancing factors of spatial interaction models. Potential accessibility measures (also called gravity-based measures) have been widely used in urban and geographical studies since the late 1940s; well-known studies are from Hansen (1959), Ingram (1971) and Vickerman (1974). The potential accessibility measure estimates the accessibility of opportunities in zone i to all other zones (n) in which smaller and/or more distant opportunities provide diminishing influences. The measure has the following form, assuming a negative exponential cost function:

$$A_i = \sum_{j=1}^{n} D_j e^{-\beta c_{ij}} \tag{1}$$

where A_i is a measure of accessibility in zone i to all opportunities D in zone j, c_{ij} the costs of travel between i and j, and β the cost sensitivity parameter. The cost sensitivity function used has a significant influence on the results of the accessibility measure. For plausible results, the form of the function should be carefully chosen and the parameters of the function should be estimated using recent empirical data of spatial travel behaviour in the study area. The function (form, parameters) is generally referred to as the impedance function. Impedance functions show a decrease in value if costs increase.

Several studies have used different impedance functions, such as the power, Gaussian or logistic functions. However, the negative exponential function is the most widely used and the most closely tied to travel behaviour theory (Handy and Niemeier, 1997). The potential measure overcomes some of the theoretical shortcomings of the contour measure, as it evaluates the combined effect of land-use and transport elements and incorporates assumptions on a person's perceptions of transport by using a distance decay function. The measures are appropriate as social indicators for analysing the level of access to social and economic opportunities for different socio-economic groups. Potential measures have the practical advantage that they can be easily computed using existing land-use and transport data (and/or models), and they have been traditionally employed as an input for estimating infrastructure-based measures. The disadvantages of potential measures are related to more difficult interpretation and communication. The measure is not easily interpreted and communicated, as it combines land-use and transport elements and weighs opportunities (according to the cost sensitivity function).

Standard potential accessibility measures ignore so-called competition effects. For example, the labour force compete for jobs; firms compete for the labour force. Ignoring such effects could lead to misleading conclusions. For example, locating all jobs at the 'best' location in a country shows the highest value for potential accessibility of the labour force, but if all employers were to locate there they would compete strongly for employees. To incorporate competition effects, several authors have adapted potential accessibility measures. Here, we summarize the different approaches. (See Geurs and Ritsema van Eck, 2003, for a more elaborate description.) Firstly, a number of authors have tried to incorporate the effects of competition on opportunities in accessibility measures by dividing the opportunities within reach from origin zone i (the 'supply' potential) by a demand potential from zone i (see Weibull, 1976; Knox, 1978; van Wee et al., 2001). This approach is useful if the travel distance between origins and destinations is relatively small, such as for elementary schools. A second approach is to use the quotient of opportunities within reach from origin i (supply potential) and potential demand of those opportunities from each destination j (see Breheny, 1978; Joseph and Bantock, 1982). This approach is useful for the analysis of accessibility to destinations where competition effects occur on destination locations (e.g. nature areas) or where available opportunities have capacity limitations (e.g. in the analysis of recreational or health-care facilities). A third approach is based on the balancing factors of Wilson's double constrained spatial interaction model (Wilson, 1971). The balancing factors a_i and b_j ensure that the magnitude of flows (e.g. trips) originating at zone i and destined for zone j equals the number of activities in zones i (e.g. workers) and j (e.g. jobs). The balancing factors of this model can be interpreted as accessibility measures, modified to account for competition effects. The balancing factors are mutually dependent, so they have to be estimated iteratively. As the balancing factors are dependent and estimated in an iterative procedure, they incorporate the competition on supplied opportunities and the competition on demand. The balancing factors are not often applied but are useful in analysing accessibility for opportunities where competition effects occur on both the origin and the destination location, such as job accessibility, where workers compete with each other for jobs and employers compete with each other for employees (Geurs and Ritsema van Eck, 2003).

Logsum accessibility measure

Publications on the logsum as a measure of consumer surplus (the difference between the market value of a product or service and the value for the user – see Chapter 12) date back to the early 1970s. One of the earliest references to the logsum as an accessibility measure is from Ben-Akiva and Lerman (1979). An introduction can be found in the textbooks on discrete choice models (e.g. Train, 2003). Here, we base our description of the logsum on De Jong et al. (2005, 2007), who present a contemporary review on the theoretical and applied literature on the logsum as an evaluation measure.

The utility that decision maker n obtains from alternative j is decomposed into an observed and an unobserved (random) component:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \tag{2}$$

where U_{nj} is the utility that decision maker n obtains from alternative j (n = 1,..N ; j = 1, . . . ,J), V_{nj} = 'representative utility' and ε_{nj} captures the factors that affect utility but are not measured by the researcher. In a standard multinomial logit (MNL) model, the choice probabilities are given by:

$$P_{nj} = \frac{e^{V_{nj}}}{\sum_j e^{V_{nj}}} \tag{3}$$

The 'logsum' now is the log of the denominator of this logit choice probability. It gives the expected utility from a choice (from a set of alternatives). It is defined as the integral with respect to the utility of an alternative, and provides an exact measure of transport user benefits, assuming the marginal value of money is constant. In the field of policy analysis, the researcher is mostly interested in measuring a change in consumer surplus that results from a particular policy. By definition, a person's consumer surplus is the utility in money terms that a person receives in the choice situation (also taking account of the disutility of travel time and costs). The decision maker n chooses the alternative that provides the greatest utility so that, provided that utility is linear in income, the consumer surplus (CS_n) can be calculated in money terms as:

$$CS_n = (1/\alpha_n)\, U_n = (1/\alpha_n)\, max_j\, (U_{nj} \,\forall\, j) \tag{4}$$

where α_n is the marginal utility of income and equal to $dU_{nj}/\, dY_n$ if j is chosen, Y_n is the income of person n, and U_n is the overall utility for the person n. The division by α_n in the consumer surplus formula translates utility into money units (e.g. dollars, euros), since $1/\alpha_n = dY_n/dU_{nj}$. If the model is an MNL model and the income utility is linear (that is, α_n is constant with respect to income), the change in expected consumer surplus for decision maker n can be calculated as the difference between $E(CS_n)$ under the conditions before the change and after the change (e.g. introduction of policy):

$$\Delta\, E(CS_n) = (1/\alpha_n)\,[ln\,(\sum_{j=1}^{J^1} e^{V^1_{nj}}) - ln\,(\sum_{j=1}^{J0} e^{V0_{nj}})] \tag{5}$$

where superscript 0 and 1 refer to before and after the change. To calculate this change in consumer surplus, the researcher must know (or have estimated) the marginal utility of income α_n. Usually, a price or cost variable enters the representative utility and, in case that happens in a linear additive fashion, the negative of its coefficient is α_n by definition (McFadden, 1981). The above equations for calculating the expected consumer surplus depend critically on the assumption that the marginal utility of income is constant with respect to income. If this is not the case, a far more complex formula is needed. However, for policy analysis, absolute levels are not required; rather only changes in consumer surplus are relevant, and the formula for calculating the expected consumer surplus can be used if the marginal utility of income is constant over the range of implicit changes that are considered by the policy. So, for policy changes that change the consumer surplus by small amounts per person, relative to their income, the formula can be used – even though in reality the marginal utility of income varies with income.

The logsum measure seems to be getting more attention in academic studies but is less popular among practitioners. An explanation for this is that the measure cannot be easily explained without reference to relatively complex theories of which most planners and political decision makers may not have a complete understanding (Koenig, 1980). See de Jong et al. (2007) for an overview of applications, and Dong et al. (2006) and Geurs et al. (2010) for recent examples. Dong et al. presented a novel approach, developing a logsum accessibility measure within a space–time framework, expressing the individual's expected maximum utility over the choices of all available activity patterns. Furthermore, attention has been paid to the behavioural assumptions underlying the logsum approach to measuring accessibility: Chorus and Timmermans (2009) show how logsum-based measures of user benefits associated with changes in the transport system (such as increases in accessibility) can be extended to allow for limited awareness among travellers.

Applying utility-based accessibility measures such as the logsum in policy appraisal may generate different results (and therefore conclusions) from infrastructure-based or location-based accessibility approaches. Suppose the researcher evaluates the accessibility effects of certain candidate policy options related to transport and land use; then there can be at least two important reasons for different results. Firstly, in the case of land-use policies, utility-based accessibility benefits from these policies can be quite large compared to investment programmes for road and public transport infrastructure. This is shown by Geurs et al. (2010), who applied the logsum accessibility benefit measure to examine the accessibility benefits from land-use and transport policy strategies. The accessibility impacts from the land-use scenarios are largely due to changes in trip production and destination utility. Ignoring accessibility benefits from land-use changes resulting from transport investments may lead to serious biases. Secondly, in the case of land-use policies, infrastructure policies or combined land-use and infrastructure policies, utility-based measures show diminishing returns, as the measures incorporate non-linear relationships between accessibility improvements and user-benefit changes. As a result, the measure may indicate that it is better to improve accessibility for individuals at locations with low accessibility levels (e.g. peripheral regions) than at locations that are already highly accessible (e.g. central urban areas) (see, for example, Koenig, 1980; Geurs and Ritsema van Eck, 2001). This is clearly relevant for social and economic evaluations of land-use and transport projects.

9.7 Conclusions and future trends in accessibility studies

The most important conclusions of this chapter are summarized as follows:

1. The conclusions on accessibility strongly depend on the definition of accessibility used and operationalization of the accessibility measure. It is therefore very important to make careful decisions on the definition and operationalization of accessibility. The four criteria on which decisions can be based are (1) purpose of the study, (2) scientific quality, (3) operationalization (cost, effort) and (4) interpretation and communication.
2. In practice, the accessibility measures used are often those that are easy to operationalize and interpret, rather than those that satisfy more stringent theoretical criteria. Applying a full set of scientific quality criteria

would imply a level of complexity and detail that is difficult to achieve in practice. This means that different situations and study purposes demand different approaches. However, it is important to recognize the implications of ignoring one or more of these criteria.

3. Location- and utility-based accessibility measures can be considered effective measures of accessibility, which can also be used as input for social and economic evaluations. These measures overcome the most important shortcomings of infrastructure-based measures and can be computed with state-of-the-practice land-use and transport data and models. It is sensible to use accessibility measures which incorporate competition effects when interpreting the analysis of accessibility to destinations where available opportunities have capacity limitations (e.g. in the analysis of jobs or recreational or health-care facilities) or where competition effects might occur on destination locations (e.g. nature areas). Utility-based measures capture the valuation of accessibility by individuals, providing a useful basis for user-benefit evaluations of both land-use and transport investments.

We now discuss a potentially important trend in accessibility measures: in academic literature, there is a continuing trend towards more complex and disaggregated accessibility measures, partly in response to the recognition that the aggregate measures lack many important details. These trends partly result from improvements in the techniques to construct location-based accessibility indicators. These techniques have evolved from very simple calculations to more complex and detailed methods that use algorithms within a geographic information systems (GIS) platform to extract and assemble data from multiple spatial databases at very fine levels of spatial resolution (see, for example, Kwan, 2000; Chen et al., 2011). In addition, the temporal component of accessibility and person-based accessibility measures seems to be enjoying a rapid increase in popularity amongst academics in transportation and geography (see, for example, Ettema et al., 2007; Schwanen and Kwan, 2008; Neutens, 2010).

Indeed, this type of accessibility measure is potentially very useful in transport policy evaluations and social evaluations. The measures may also be tied to the utility-based approach, which opens up the possibility of using them in economic evaluations. There are two important trends which will better enable the development of disaggregated accessibility measures. Firstly, recent studies have shown the possibilities of synthetic data to reduce the enormous data collection effort, particularly for person-based accessibility measures (see, for example, Veldhuisen et al., 2000; Arentze et al., 2008). In California, an activity-based travel demand model is under development which synthetically generates the entire population and aims to simulate activities at a very fine level of spatial and temporal disaggregation (second-by-second and parcel-by-parcel level) (Chen et al., 2011). Secondly, ICT developments will hugely increase the possibilities for detailed data collection. The fast penetration of smartphones provides great opportunities for researchers to improve the quality of accessibility analysis. Berg Insight (2010), for example, forecasts that, by 2014, 60 per cent of new phones sold worldwide will be GPS enabled. This new type of ICT is likely to become more important than PC use at fixed locations. Data collection using smartphones can potentially reduce the burden for respondents in participating in longitudinal travel surveys, and this opens up new possibilities for monitoring travel behaviour and individual accessibility over relatively long time periods (months and years).

REFERENCES

Arentze, T.A., H.J.P. Timmermans and F. Hofman (2008), 'Creating synthetic household populations: problems and approach', *Transportation Research Record*, 2014, 85–91.

Balcombe, R., R. Macket, N. Paulley, J. Preston, J. Shires, H. Titheridge, M. Wardman and P. White (2004), *The Demand for Public Transport: A Practical Guide*, Wokingham: TRL.

Ben-Akiva, M. and S.R. Lerman (1979), 'Disaggregate travel and mobility choice models and measures of accessibility', in D.A. Hensher and P.R. Sopher (eds), *Behavioural Travel Modelling*, Andover: Croom Helm, pp. 654–679.

Berg Insight (2010), *GPS and Mobile Handsets*, Summary, Gothenburg: Berg Insight.

Black, J. and M. Conroy (1977), 'Accessibility measures and the social evaluation of urban structure', *Environment and Planning A*, 9, 1013–1031.

Breheny, M.J. (1978), 'The measurement of spatial opportunity in strategic planning', *Regional Studies*, 12, 463–479.

Burns, L.D. (1979), *Transportation: Temporal and Spatial Components of Accessibility*, Lexington, MA and Toronto: Lexington Books.

Cervero, R., T. Rood and B. Appleyard (1997), *Job Accessibility as a Performance Indicator: An Analysis of Trends and Their Social Policy Implications in the San Francisco Bay Area*, Berkeley: Institute of Urban and Regional Development, University of California.

Chen, Y., S. Ravulaparthy, P. Dalal, S.Y. Yoon, T. Lei, K.G. Goulias, R.M. Pendyala, C.R. Bhat and H.H. Hu (2011), 'Development of opportunity-based accessibility indicators', paper for the 90th Annual Meeting of the Transportation Research Board.

Chorus, C.G. and H.J.P. Timmermans (2009), 'Measuring user benefits of changes in the transport system when traveler awareness is limited', *Transportation Research Part A*, 43 (5), 536–547.

Dalvi, M.Q. and K.M. Martin (1976), 'The measurement of accessibility: some preliminary results', *Transportation*, 5, 17–42.

De Jong, G., A. Daly, M. Pieters and T. van der Hoorn (2007), 'The logsum as an evaluation measure: review of the literature and new results', *Transportation Research Part A: Policy and Practice*, 41, 874–889.

De Jong, G., M. Pieters, A. Daly, I. Graafland, E. Kroes and C. Koopmans (2005), *Using the Logsum as an Evaluation Measure: Literature and Case Study*, Leiden: RAND Europe/SEO.

Dijst, M., H. Jayet and I. Thomas (2002), 'Transportation and urban performance: accessibility, daily mobility and location of households and facilities', in M. Dijst, W. Schenkel and I. Thomas (eds), *Governing Cities on the Move: Functional and Management Perspectives on Transformations of Urban Infrastructures in European Agglomerations*, Aldershot: Ashgate.

Dong, X., M.E. Ben-Akiva, J.L. Bowman and J.L. Walker (2006), 'Moving from trip-based to activity-based measures of accessibility', *Transportation Research Part A: Policy and Practice*, 40, 163–180.

Ettema, D., T. Schwanen and H.J.P. Timmermans (2007), 'The effect of location, mobility and socio-demographic factors on task and time allocation of households', *Transportation*, 34, 89–105.

Geurs, K.T. and J. Ritsema van Eck (2001), *Accessibility Measures: Review and Applications. Evaluation of Accessibility Impacts of Land Use Transport Scenarios, and Related Social and Economic Impacts*, Bilthoven: National Institute of Public Health and the Environment, accessed at http://www.rivm.nl/bibliotheek/rapporten/408505006.pdf.

Geurs, K.T. and J.R. Ritsema van Eck (2003), 'Accessibility evaluation of land-use scenarios: the impact of job competition, land-use and infrastructure developments for the Netherlands', *Environment and Planning B: Planning and Design*, 30, 69–87.

Geurs, K.T. and B. van Wee (2004), 'Accessibility evaluation of land-use and transport strategies: review and research directions', *Journal of Transport Geography*, 12, 127–140.

Geurs, K., B. Zondag, G. de Jong and M. de Bok (2010), 'Accessibility appraisal of integrated land-use/transport policy strategies: more than just adding up travel time savings', *Transportation Research Part D*, 15, 382–393.

Gould, P. (1969), 'Spatial diffusion', Resource Paper no. 17, Commission on College Geography, Association of American Geographers, Washington, DC (cited by J.M. Morris, P.L. Dumble and M.R. Wigan, 1979, 'Accessibility indicators for transportation planning', *Transportation Research A*, 13, 91–109).

Hägerstrand, T. (1970), 'What about people in regional science?', *Papers of the Regional Science Association*, 24, 7–21.

Handy, S.L. and D.A. Niemeier (1997), 'Measuring accessibility: an exploration of issues and alternatives', *Environment and Planning A*, 29, 1175–1194.

Hansen, W.G. (1959), 'How accessibility shapes land use', *Journal of the American Institute of Planners*, 25, 73–76.

Ingram, D.R. (1971), 'The concept of accessibility: a search for an operational form', *Regional Studies*, 5, 101–107.

Jones, S.R. (1981), *Accessibility Measures: A Literature Review*, Crowthorne: Transport and Road Research Laboratory.

Joseph, A.E. and P.R. Bantock (1982), 'Measuring potential physical accessibility to general practitioners in rural areas: a method and case study', *Social Science and Medicine*, 16, 85–90.

Knox, P.L. (1978), 'The intraurban ecology of primary medical care: patterns of accessibility and their policy implications', *Environment and Planning A*, 10, 415–435.

Koenig, J.G. (1980), 'Indicators of urban accessibility: theory and applications', *Transportation*, 9, 145–172.

Kwan, M.-P. (1998), 'Space–time and integral measures of individual accessibility: a comparative analysis using a point-based framework', *Geographical Analysis*, 30, 191–216.

Kwan, M.-P. (2000), 'Interactive geovisualisation of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set', *Transportation Research Part C*, 8, 185–203.

Linneker, B.J. and N.A. Spence (1992), 'Accessibility measures compared in an analysis of the impact of the M25 London orbital motorway on Britain', *Environment and Planning A*, 24, 1137–1154.

McFadden, D. (1981), 'Econometric models of probabilistic choice', in C.F. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data with Economic Applications*, Cambridge, MA: MIT Press, pp. 198–272.

Neutens, T. (2010), *Space, Time and Accessibility: Analyzing Human Activities and Travel Possibilities from a Time-Geographic Perspective*, Gent: Department of Geography, University of Gent.

Niemeier, D.A. (1997), 'Accessibility: an evaluation using consumer welfare', *Transportation*, 24, 377–396.

Pirie, G.H. (1979), 'Measuring accessibility: a review and proposal', *Environment and Planning A*, 11, 299–312.

Pirie, G.H. (1981), 'The possibility and potential of public policy on accessibility', *Transportation Research A*, 15, 377–381.

Schwanen, T. and M.-P. Kwan (2008), 'The internet, mobile phone and space–time constraints', *GeoForum*, 39, 1362–1377.

Shen, Q. (1998), 'Location characteristics of inner-city neighborhoods and employment accessibility of low-wage workers', *Environment and Planning B: Planning and Design*, 25, 345–365.

Song, S. (1996), 'Some tests of alternative accessibility measures: a population density approach', *Land Economics*, 72 (4), 474–482.

Train, K. (2003), *Discrete Choice Methods with Simulation*, New York: Cambridge University Press.

Van Exel, N.J.A. and P. Rietveld (2009), 'Could you also have made this trip by another mode? An investigation of perceived travel possibilities of car and train travellers on the main travel corridors to the city of Amsterdam, The Netherlands', *Transportation Research Part A: Policy and Practice*, 43, 374–385.

Veldhuisen, J., H. Timmermans and L. Kapoen (2000), 'RAMBLAS: a regional planning model based on the microsimulation of daily activity travel patterns', *Environment and Planning A*, 32 (3), 427–443.

Vickerman, R.W. (1974), 'Accessibility, attraction, and potential: a review of some concepts and their use in determining mobility', *Environment and Planning A*, 6, 675–691.

van Wee, G.P., M. Hagoort and J.A. Annema (2001), 'Accessibility measures with competition', *Journal of Transport Geography*, 9, 199–208.

Weibull, J.W. (1976), 'An axiomatic approach to the measurement of accessibility', *Regional Science and Urban Economics*, 6, 357–379.

Wilson, A.G. (1971), 'A family of spatial interaction models, and associated developments', *Environment and Planning A*, 3, 1–32.

10

Transport and the environment

Bert van Wee, David Banister, Jan Anne Annema and Karst Geurs

10.1 Introduction

Over the last century transport has brought enormous benefits to society through facilitating the globalization of the world economy, promoting the internationalization of businesses, allowing for networking, and providing opportunities for many people to visit friends and relatives anywhere in the world. But this enormous increase in mobility and the potential to travel vast distances does not come without a cost, as there are negative impacts on society, safety, the environment, resource use and congestion. Much of this growth has been predicated on the availability of cheap and abundant oil.

This chapter gives an overview of the environmental impacts of transport and the factors that determine those impacts. Chapter 8 has already discussed the relevance of technology to these impacts. Table 10.1 shows that motorized transport affects the natural and built environment in many ways: by emitting greenhouse gases and air pollutants, by making noise and by taking land. These primary causes can have negative impacts on biodiversity, open spaces, health and so forth. It would take too much space to treat all these

Table 10.1 Environmental impacts of motorized transport

Natural environment		Built environment
Biodiversity	Greenhouse gas emissions	Liveability
Water run-off and biosystems impact	Air pollutant emissions	Health
Oil extraction	Land take for roads and urban sprawl	Community severance
Mining of iron and other materials		
Open space and landscape	Noise	

Source: Based on Ernst (2011).

primary causes and their final impacts in detail in this chapter. Therefore we will focus on three main environmental issues associated with transportation: greenhouse gas emission, air pollution and noise nuisance.

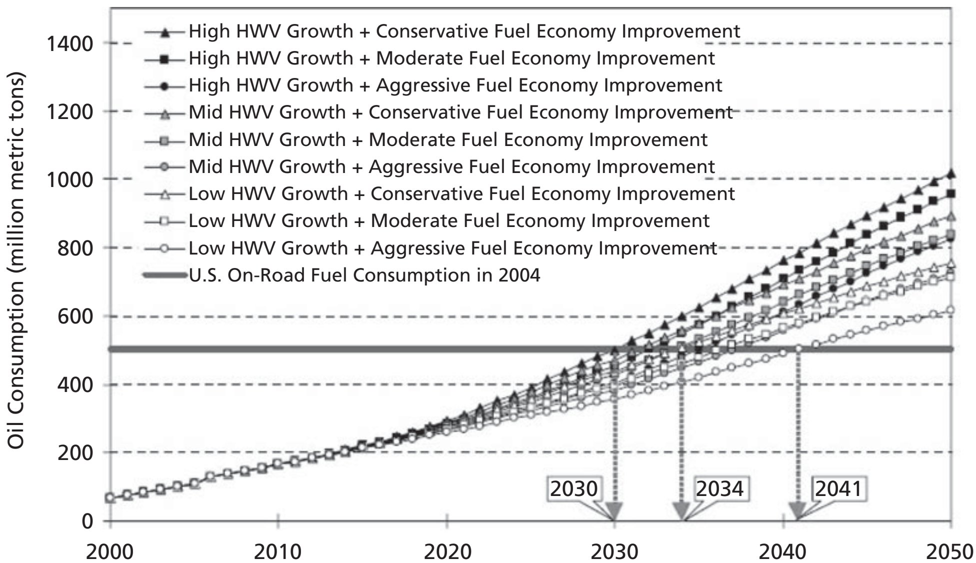
10.2 Overview of the environmental problems of transport

Concern over environmental problems increased rapidly in the late 1960s, mainly because of the rapid growth in population, industrialization, incomes, and consumption levels. In 1972 the Club of Rome produced their influential report *The Limits to Growth* (Meadows et al., 1972; see Meadows et al., 2005, for an update) arguing that exponential population and economic growth would be disastrous for the environment and the depletion of fossil fuels and raw materials. In 1987 another influential report was published: the report of the World Commission on Environment and Development (1987), often referred to as the Brundtland Report (Brundtland chaired the commission). It argued that the environment and environmental protection are conditions for future worldwide growth of the population and the economy, and that the unequal distribution of welfare could be a source of conflicts and ecological destruction, and that this could lead to stagnation in social, economic and technological development.

Transport, as one of the strongest-growing economic sectors in the twentieth century, has received a lot of political attention related to concerns over environmental problems. For example, as early as the 1950s it became clear in California, US, that the fast-growing car market was a major cause of periods of severe smog (Air Resources Board, 2011). Table 10.2 gives an overview

Table 10.2 Dominant environmental impacts by transport mode

	Road	Rail	Air	Water
Climate change (carbon dioxide emissions)	*	*	*	*
Energy use (oil depleting)	*	*	*	*
Use of raw materials and production of waste materials	*	*	*	*
Air pollution (nitrogen oxides, particulate matter and other emissions)	*	*	*	*
Soil and water pollution	*			
Open space damage, fragmentation and land take	*	*		
Odour pollution	*		*	
Noise pollution	*	*	*	
Non-emissions-related liveability	*			



Source: Wang et al. (2006).

Figure 10.1 Projected annual oil demand by Chinese motor vehicles under the nine combinations of scenarios

of dominant environmental impacts per transport mode. It shows that road, rail, air and waterborne transport has several impacts but not all modes contribute to all problems. However, road transport does contribute substantially to all impacts. 'Non-emissions-related liveability' impacts need further explanation. Let us assume that road vehicles would not use energy, not make any noise and not emit pollutants. Even then, they have negative impacts, as a result of land take (e.g. for parking) and community severance and by preventing streets being used for non-transport-related activities (e.g. for play). Such impacts are labelled 'non-emissions-related liveability' impacts.

In the USA, in 2009, the transport system accounted for 27 per cent of energy use (Davis et al., 2010). In the European Union (EU27) (EC, 2007) the transport sector's energy share is somewhat higher: around 32 per cent (EEA, 2010a). The USA is also the biggest consumer of oil for transport purposes worldwide, but this position will be challenged in the near future by upcoming economies, for example by China (Figure 10.1).

Table 10.3 presents some of transport's emission shares for the USA and Europe. It shows that in the USA and Europe the transport sector emits relatively large amounts of air pollutants such as nitrogen oxides (NO_x) and carbon monoxide (CO). The two continents largely differ in, among other things, par-

Table 10.3 Emissions share from transport in the USA and Europe (2008)

Pollutant		US (%)	Europe (EU27) (%)
Sulphur dioxide	SO_2	4.5	4
Nitrogen oxides	NO_x	57.9	49
Carbon monoxide	CO	73.2	36
Volatile organic compounds	VOC	37.7	18
Particulate matter, particles smaller than 2.5 μm	$PM_{2.5}$	0.2	18
Particulate matter, particles smaller than 10 μm	PM_{10}	3.2	16
Carbon dioxide	CO_2	33.2	24

Source: Davis et al. (2010) for US data; EEA (2010b) for European CO_2 estimate; EEA (2010c) for European air pollutant estimates.

ticulate matter ($PM_{2.5}$ and PM_{10}; the suffixes will be explained in section 10.4) emissions. This is mainly due to the use of diesel combustion technology in the European car fleet (slightly more than 50 per cent of all new cars sold on the European market are diesel), while in the US car fleet only petrol is used. Diesel cars sold before 2009 emitted far more $PM_{2.5}$ and PM_{10} per kilometre driven compared to petrol cars (more details in section 10.4). The share of transport SO_2 emissions is small on both continents, as most of the car fleet and heavy-duty vehicles use low-sulphur petrol and diesel. However, the emissions of ships at sea are excluded in Table 10.3, as these do not occur within a country's boundaries and are therefore not included in national statistics. Fuel oil for sea-going ships especially has a high sulphur content.

10.3 Data and trends in transport CO_2 emissions and oil use[1]

It is really only in the last 20 years that the issue of carbon has arrived on the international agenda, and the science of global warming is relatively young, even though the Swedish scientist Arrhenius had first speculated about the greenhouse effect over 100 years ago.[2] Since the 1990s climate change (Table 10.2, top row) has been prominent on the international agenda of researchers and policy makers, and the reports of the Intergovernmental Panel on Climate Change (IPCC) provide the best-known and most influential overviews about state-of-the-art scientific knowledge. The IPCC (2007: 5) concluded that 'Most of the observed increase in global average temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic Greenhouse Gases (GHG) concentrations. It is likely that there has been significant anthropogenic warming over the past 50 years

Table 10.4 CO_2 emissions 1990–2008, worldwide

	Change 1990–2008 (%)	2008 level (Mega tonnes)
Total CO_2	40	29381.43
Transport CO_2	44	6604.66
Road CO_2	48	4848.42
Aviation CO_2	40	752.19
Waterborne CO_2	56	706.59

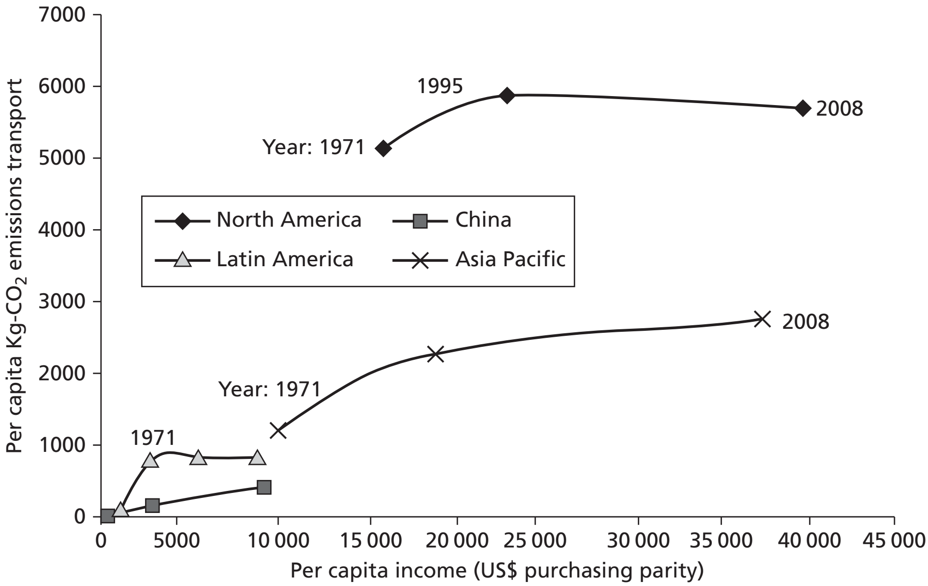
Source: ITF (2011).

averaged over each continent (except Antarctica).' Transport emits GHGs such as carbon dioxide (CO_2), nitrous oxide (N_2O) and methane (CH_4). As CO_2 is by far the most important GHG emitted by transport we will only focus on this GHG in this chapter.

Transport is the only major sector where CO_2 emissions continue to grow, with car use, road and maritime freight and aviation being the principal contributors to GHG emissions (Chapman, 2007). This results from its importance in terms of global trade and travel, the costs associated with any change, the global growth that is taking place in transport and its total dependence on carbon based energy sources. Transport CO_2 emission worldwide has grown steadily since 1990, and by a higher rate compared to total CO_2 emissions (Table 10.4).

Transport emits CO_2 because of the use of oil products such as fuel. Oil products consist of long chain hydrocarbons with a high energy content. In the internal combustion engines (ICEs) of cars, vans, lorries, ships and planes the hydrocarbons are burned, resulting in the desired energy for propulsion and in the undesired waste product CO_2. The popularity of oil as transport fuel can be explained technically because of its high energy content. This means that, with a relatively low fuel volume on board, vehicles can cover many kilometres. For example, a modern car can carry 40 litres of petrol or diesel and easily travel 500 to 700 kilometres before refuelling is required.

The growth in oil consumption between 1973 and 2008 has been more than 110 per cent (IEA, 2010), and global CO_2 emissions from transport have increased 44 per cent in the period 1990–2008 (Table 10.4). Motorized road transport is responsible for around 74 per cent of world transport CO_2 emissions (IEA, 2010), and per capita CO_2 emissions have grown with income growth (Figure 10.2). Figure 10.2 shows that the Asia Pacific region is more CO_2 efficient than the North American region. The question remains



Note: North America – USA, Canada; Asia Pacific – Japan, South Korea, Australia, New Zealand.

Source: Banister et al. (2011).

Figure 10.2 Carbon emissions as a function of income, 1971–2008

whether Latin America, China and other regions will move towards the North American group or towards the Asia Pacific group led by Japan.

CO_2 emissions and fuel use in road transport are a function of the number of vehicles on the road (affected by affordability, consumer incomes and preference (see Chapter 2) and fuel price), distances driven (affected by prices such as fuel price and tax policies such as car taxation, tolls or congestion charge (see Chapter 6), land use (see Chapter 5) and infrastructure characteristics (see Chapters 6 and 7)) and the level of fuel economy of the cars (amount of motor fuel consumed per kilometre driven per year). (See Chapter 2 of this book for a general overview of links between the determinants.) The personal vehicle has impeded the decarbonization of world transport activity because of, among other things, the increase in vehicle size (which requires more energy input; see Chapter 8 of this book for more details).

Oil dependency, peak oil

The debate on oil price and availability is closely related to CO_2 emission (about 3.15 kg of CO_2 is produced from each kilogram of petrol or diesel used). Although it is debatable whether or not oil availability is an environmental problem, we still discuss this topic here briefly because of the close relationship between oil and climate change.

Even though there have been concerns over the price and availability of oil since the 1970s, it is only recently that more fundamental questions have been raised over its future. Transport is almost totally dependent on oil, and the world price of oil has fluctuated between $40 and $140 a barrel between 2009 and 2011. It is uncertain as to how that might change in the short or longer term, as its price has an impact on travel demand and on the take-up of alternative technologies. In addition, there is the question of the availability of oil, which again influences the price of travel and alternative technologies. At present, transport accounts for over 61 per cent of global oil demand, and in the EU27 this level is even higher (over 71 per cent; IEA, 2009a). There are also no obvious alternative fuels that have a high energy density and are available at a reasonable price and in sufficient quantities.

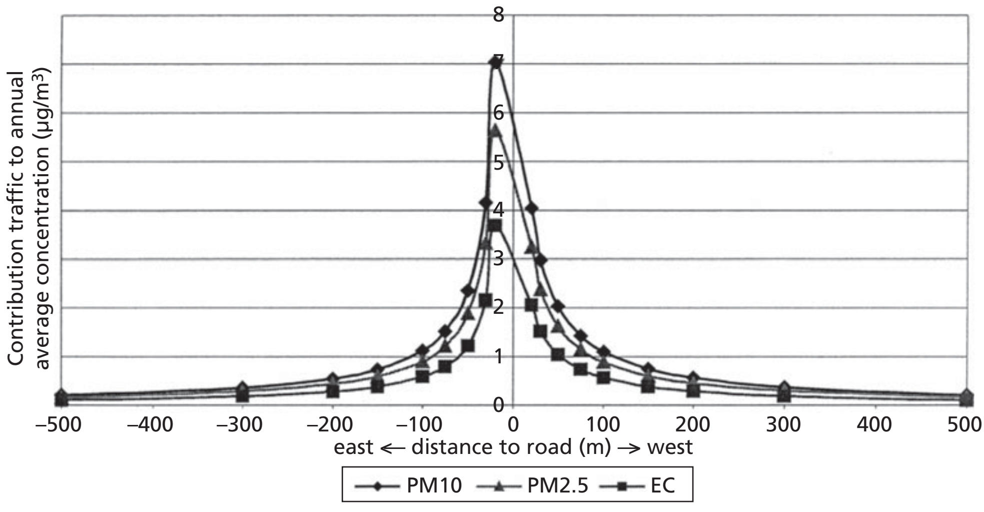
The depletion of fossil fuels, more specifically oil, raises major concerns, for example from the International Energy Agency (IEA). It is generally expected that the level of (proven and expected) reserves raises concerns, but, even if reserves were not a problem, production capacity may become a problem. Some experts expect worldwide oil production to peak between now and 2030 (Aleklett, 2008). Reducing oil production, combined with a worldwide increasing demand for oil, may result in high energy (or at least oil) prices. Developing commercially attractive alternatives for current oil products is already a major challenge. Options that are currently developed and discussed include electric vehicles, biomass fuels (including biofuels), hydrogen vehicles, and very efficient internal combustion engines. Electric vehicles can use energy produced from many sources, ranging from fossil fuels (oil, gas, coal) to renewables (solar, wind, biofuels and others). Non-oil based fossil fuels include gas-to-liquids and coal-to-liquids.

Looking at the future of oil, Gilbert and Perl (2008) claim that world oil production will peak in 2012. They are not alone in this view, even though it is extremely difficult to obtain accurate figures from many of the producers on their reserves. Simmons (2005) came to the same conclusion in 2003 through his detailed analysis of available information on Saudi Arabia's reserves, which account for about 25 per cent of the world's total. The main criticism of 'peak oil' is that, with higher prices and improvements in technology, new fields will be found, more oil will be recovered, and oil shale and tar sands will become economic. But the problem here is the speed at which change is needed and the levels of investment required to bridge the gap between cheap oil and expensive oil and the suitability of low-grade oil for transport.

10.4 Data and trends in air pollution

In addition to CO_2 emissions, the use of ICEs results also in emissions of air polluting substances such as NO_x, SO_2, VOCs and $PM_{2.5}$ and PM_{10} (Table 10.3). The emissions of these substances can lead to negative health impacts and to acid rain. Acid rain occurs when SO_2 and NO_x are emitted into the atmosphere. The NO_x and SO_2 molecules are transported in the air (short-range or relatively long-range, i.e. hundreds of kilometres), undergo chemical transformations and are absorbed by water droplets in clouds. The droplets then fall to earth as rain, snow or sleet. This can increase the acidity of the soil and affect the chemical balance of lakes and streams, resulting in ecological damage. Transport air pollutant emissions can also lead to local negative health impacts. Exposure to $PM_{2.5}$ and PM_{10}, especially, affects people's health (Pope and Dockery, 2006; Grahame and Schlesinger, 2010; WHO and JRC, European Commission, 2011). $PM_{2.5}$ and PM_{10} stand for different particulate matter fractions. 'Particulate matter' is a term for very small particles floating in the air, which humans can deeply inhale. The suffix 2.5 points at the fraction of particles which are smaller than 2.5 μm; the suffix 10 means smaller than 10 μm. Deep inhalation can result in severe health impacts (respiratory problems and even premature death). The World Health Organization considers air pollution a major environmental risk to health and estimates that it causes approximately 2 million premature deaths worldwide per year (WHO, 2009).

The 2 million premature deaths mentioned are not only because of transport emissions. However, transport's share in negative effects is relatively high, as (on average) the distance between road traffic and the people exposed is much shorter than from other emission sources of pollution, such as power plants. Traffic emissions therefore have a greater health impact per kilogram than average emissions, as people live close to the sources of pollution and are exposed to continuous levels, particularly along busy roads (Dorland and Jansen, 1997; Eyre et al., 1997; Newton, 1997). The importance of the distance between the source and the receptor is visualized by the concentrations of pollutants as a function of the distance between a location and the road. The greater the distance, the lower the concentration, and this is visualized (Figure 10.3) with the contribution of road-traffic-related emissions of PM_{10} and $PM_{2.5}$ in microgram/m^3 as a function of the distance from the road, using a motorway location in the Netherlands.

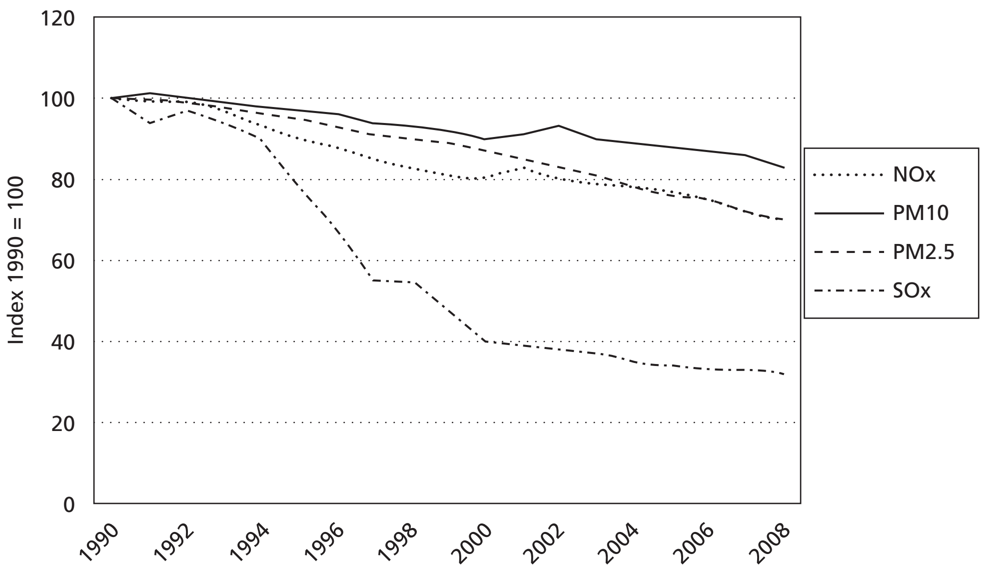


Source: Janssen et al. (2002).

Figure 10.3 Contribution of road-traffic-related emissions as a function of the distance to the road

Based on an overview of studies in this area and referring to the publications of Smith (1993a, 1993b) and Evans et al. (2002), we can conclude that the intake of emissions of particulates of vehicles is 10 times higher than those of a power plant. For further examples of this concept, see Marshall et al. (2003, 2005).

The air pollutant emissions of vehicles are caused by technical imperfections in the ICEs in combination with fuel characteristics such as the sulphur content. In contrast to CO_2 emissions vehicle manufacturers all over the world have succeeded in producing vehicles which emit fewer air pollutants per kilometre driven compared to 10, 20 or 30 years ago. Fuel producers worldwide also deliver far cleaner transport fuels (related, among other things, to sulphur contents) nowadays compared to 10 years ago. Both developments are policy driven (see section 10.6). The main reason for these technical successes seems to be the clear political recognition of air pollution being a major problem (which it had been since the 1950s), in combination with relatively cheap technical measures available to make vehicles cleaner (i.e. by implementing technical measures such as the three-way catalyst, particle soot filter and so forth; see Chapter 8). The total final air pollutant emissions by transport depend on transport and traffic volumes (such as the number of vehicle kilometres), the technology of the vehicles and fuels used, and driving behaviour. For example, the speed and acceleration characteristics matter: driving at high speeds and aggressively causes higher overall



Source: http://www.eea.europa.eu/data-and-maps/figures/transport-emissions-of-primary-and/transport-emissions-of-primary-and/Transport_particle_formation.JPG.75dpi.gif/at_download/image.

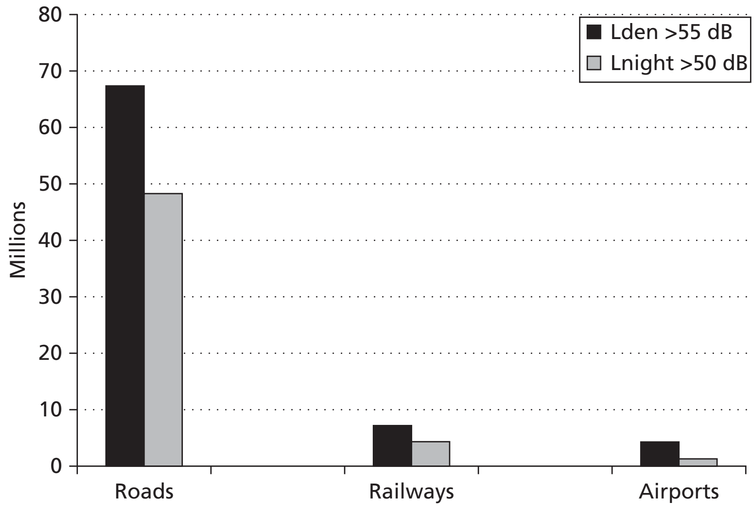
Figure 10.4 Road transport air pollutant emissions 1990–2008 for countries belonging to the European Economic Area (EEA; roughly all EU countries plus Iceland, Liechtenstein and Norway)

emission levels (El-Shawarby et al., 2005). As Figure 10.4 shows for Europe, the technical measures taken have been sufficiently strong to beat road transport's volume growth between 1990 and 2008; transport air pollutant emissions actually declined in this period.

10.5 Data and trends in noise

Another main environmental problem related to transport is noise. At first glance the main impact of transport noise seems to be annoyance. However, in WHO and JRC, European Commission (2011) it is estimated that health problems are also associated with exposure to transport noise. The report states (p. xvii) that:

> DALYs lost from environmental noise in the Western European countries are 61 000 years for ischaemic heart disease, 45 000 years for cognitive impairment of children, 903 000 years for sleep disturbance, 22 000 years for tinnitus and 587 000 years for annoyance. If all of these are considered together, the range of burden would be 1.0–1.6 million DALYs. This means that at least 1 million healthy life years are lost every year from traffic-related noise in the western European countries, including the EU Member States. Sleep disturbance and annoyance related to



Source: EEA (2009).

Figure 10.5 Amount of people (in millions) in the European Union exposed to daily road, railway and airport noise levels exceeding 55 dB L_{den} and 50 dB L_{night}

road-traffic noise constitute most of the burden of environmental noise in western Europe.

DALY stands for disability adjusted life years. They are the sum of the potential years of life lost due to premature death and the equivalent years of healthy life lost by virtue of being in a state of poor health or disability (WHO and JRC, European Commission, 2011).

Important transport noise sources are roads, railways and airports (Figure 10.5). Road and rail transport noise is caused by the engine and the contact between tyres and road or between wheels and rails. At higher speeds (roughly more than 40 km/h) tyre–road surface contact noise is dominant for road transport. For high-speed trains at high velocity, the frame also causes noise. Aircraft noise (most notable during landing and take-off) is primarily produced by the engine as air is sucked into the engine and exits from the exhaust at high velocity. Noise is also created by the airframe as it moves through the air.

L_{den} (see Figure 10.5) has been put forward as the noise metric for the prediction of annoyance during exposure in the daytime, evening and night (L_{night} defines exposure and annoyance at night; EC, 2002). At L_{den} 55 dB (decibel) around airports, it is estimated that 28 per cent of people exposed to this level of aircraft noise are annoyed; for roads this is 18 per cent and railways 10 per cent. These figures show that people experience the same level of noise from the three different modes of transport differently. Railway noise is the least annoying (ceteris paribus).

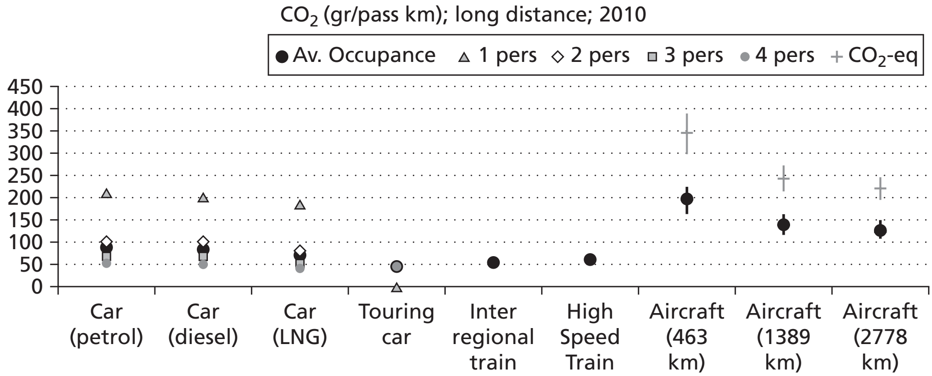
The trend in transport noise impact worldwide is not known. However, it is clear that transport noise impacts have been high on the political agenda for a long time. For example, in the US around 4.5 billion US dollars in total (in 2007 prices) was spent on noise barriers on federal highways in the period 1963–2007 (US Department of Transportation, Federal Highway Administration, 2011).

10.6 A comparison of environmental performance of modalities

From a policy perspective, it is interesting to know the environmental performance of different transport modalities. The focus in sections 10.2, 10.3 and 10.4 was on the use of ICEs and their environmental impacts, but electric trains also use energy and produce emissions. The policy question could be: is the use of an electric train preferable from an environmental perspective compared to the use of a car with an ICE? In this section we firstly focus on emissions.

Comparing the emissions performance of transport modes is not straightforward, as can be demonstrated through the following range of questions (based on van Wee et al., 2005):

1. Which emissions are selected for the comparison?
2. Which emissions are compared: only direct emissions in the use stage, or also indirect emissions, such as those arising from the production of vehicles or even infrastructure?
3. In the case of direct emissions, are emissions of electricity production and of refineries included?
4. For which year(s) are emissions compared?
5. Are average emissions compared, or emissions for sub-segments (such as for short- or long-distance travel only, or for containers only)?
6. Are average emissions compared or marginal emissions (extra emissions due to extra travel – Rietveld, 2002)?
7. Are average emissions compared, or for specific time periods (e.g. rush hour versus off-peak hours)? Note that, in the case of public transport, the choice for peak versus off-peak interacts with marginal versus average values: marginal peak hour emissions are relatively high, whereas average peak hour emissions are relatively low. For off-peak hours the opposite is true.



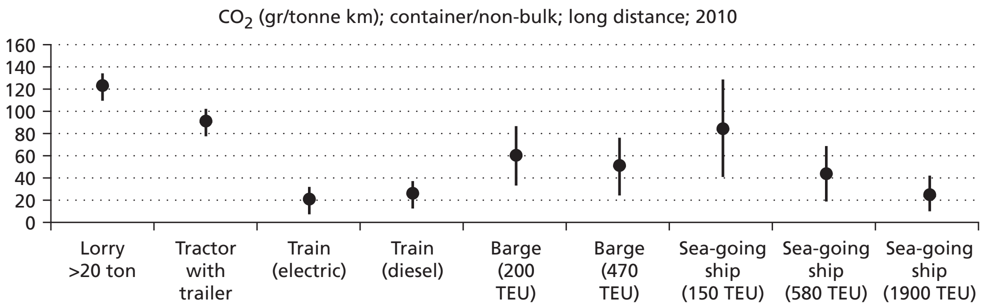
Source: den Boer et al. (2008).

Figure 10.6 CO_2 emissions for passenger transport per mode in the Netherlands, 2010

8. For the operational use of vehicles (e.g. driving circumstances of road vehicles), are real-world circumstances assumed, or test circumstances?
9. Which load factors or occupancy rates are assumed (real-world averages versus assumed factors)?

Generally speaking, van Wee et al. (2005) distinguish between direct and indirect factors. Direct factors include technical factors (such as vehicle characteristics), operational factors (such as speed) and logistical factors (such as occupancy rates and load factors), all having an impact on mode specific environmental performance. In current practice, two main categories of research methods for calculating direct energy use and emissions per transport unit can be distinguished: 'bottom-up' (BUMs) and 'top-down' (TDMs) methods. BUMs are engineering methods consisting of such variables as vehicle weight, drag resistance, speeds and so on. The TDMs calculate energy and emission factors by dividing total energy use and emissions by the transport indicator, for example CO_2 emissions/tonne km for barge transport or NO_x emissions/passenger km for rail transport. Both methods have their advantages and disadvantages (van Wee et al., 2005).

As an illustration, Figure 10.6 gives direct 'real-world average' emissions of CO_2-equivalents for long-distance (Dutch) passenger transport in 2010. It shows that the touring car, inter-regional train and high-speed rail have the lowest average emissions, but that a passenger car with four persons performs about equally. Aircraft and single-person cars have about 2.5–4 times higher emissions. For aircraft the flight distance is of major impact on the emissions per kilometre.



Source: den Boer et al. (2008).

Figure 10.7 CO_2 emissions for freight transport per mode in the Netherlands, 2010

Figure 10.7 presents the average CO_2 emissions for long-distance non-bulk container freight transport, showing that on average trains and large sea-going ships perform much better (up to a factor 6) than lorries, and also better (about a factor 4) than trailers and small sea-going ships.

One means of addressing the problems with definitions and inconsistency in determining the environmental effects of transport is to use life cycle assessments (LCA) methods, which were developed in the 1960s and have been used in transport since the late 1990s. They take account of the non-negligible impacts of transport activity through vehicle production and scrapping, energy (fuel) extraction and production and the full environmental impacts associated with infrastructure such as land take, materials use and the construction itself (Chester and Horvath, 2009). LCA has demonstrated the full range of transport activities associated with the entire supply chain of products (Browne et al., 2005). Even if these life cycle factors are included in the costs of transport (internalized), they have little effect on transport volumes (Maibach et al., 2007), as the demand elasticities for travel are low and rising income levels reduce the effectiveness of higher prices (Goodwin et al., 2004; Graham and Glaister, 2004; see also Chapter 6).

10.7 Transport environmental policy instruments[3]

This section focuses on the different types of policy measures – regulation, pricing, land-use planning, infrastructure policies, marketing, and information and communication (Table 10.5). These types of policy measures have been implemented all over the world, albeit every country or region differs in the specific policies chosen. Table 10.5 shows the relationship between types of policy measures and determinants for the environmental impact of transport. More information on transport policy can be found in Chapter 12.

Table 10.5 Dominant relationships between policy instruments and determinants for impact of transport on the environment

	Transport volume (e.g. 'less volume, less environmental pressure')	Modal split (e.g. 'car versus train')	Technology (e.g. 'petrol versus diesel cars')	Efficiency vehicles (e.g. 'light versus heavier petrol cars')	Driving behaviour (e.g. 'driving style')
Regulation	*	*	*		*
Prices	*	*	*	*	*
Land-use planning	*	*			*
Infrastructure	*	*			*
Public transport policies	*	*			
Marketing		*			
Information and communication	*	*		*	*

Source: Based on Blok and van Wee (1994).

Regulations

Regulations with respect to the access of motorized vehicles to central urban areas have been successful in improving the attractiveness of these areas by changing at least the modal shift, and probably also total transport volumes (owing to the increased travel times and costs of travelling by car). The distribution over time and space of transport volumes has also been influenced by regulations (e.g. limitations for access of lorries and vans, and loading and unloading times for these vehicles).

It is clear (at least in OECD countries) that emission regulations for road vehicles have been successful in reducing transport emissions of pollutants. Emissions of CO, VOC and NO_x per car and lorry kilometre have all decreased by a factor of between 5 and 10 since 1980 (see Chapter 8). Lead emissions from petrol cars have decreased to almost zero because of regulations in the 1970s and 1980s to ban lead in petrol. Regulations have also included speed regulations.

However, vehicle regulations for noise emissions have been much less effective (see Chapter 8). Although the emissions standards have been tightened and vehicles should have become much quieter, the reality is different, as car fleets have not become quieter (1980–96), and noise from fleets of lorries has only reduced by 3 dB(A) (van der Toorn and van den Dool, 1997).

There are two main reasons. Firstly, the potentially positive effects of new standards have been weakened because of changes in the measurement methods (as negotiated by the car industry). Secondly, there has been an increase in the share of noisier diesel cars and noisier tyres (increased width) in the car stock.

Pricing

Pricing policies include subsidies for public transport, levies on vehicles and fuels, prices for parking and road pricing. Subsidies on public transport (PT) have reduced the fares for PT, and so decreased transport resistance (Chapter 6), and increased PT patronage. However, the stated political reasons for subsidies have often included environmental improvement (by changing modal choice from car to PT), reduced congestion (also because of the change in modal choice) or improved access for people not having a car available (reduced social exclusion). The effectiveness discussion then becomes more complicated. For the impact of PT subsidies on the environment, the substitution effects (from car to PT) are important, as are the potential generation effects of additional travel ('induced demand' – see Chapter 6, and Goodwin, 1996). Because of the limited overlap in markets between PT and cars (Bovy et al., 1991), overall decreases in PT tariffs could have a negative impact on the environment, as the positive effect of mode change is more than compensated for by the negative effect of the generation of additional travel, at least for energy use and CO_2 emissions (van Wee et al., 2005).

Levies on fuels and taxes on cars can change the share between different fuel types. As there are differences between countries in these levies and taxes, the share of diesel cars is much lower in some countries, such as the UK or the Netherlands, as compared with others, such as Belgium and France. A higher share of diesel reduces CO_2 emissions (at least on a per-kilometre basis) and increases emissions of PM and NO_x as well as noise. Secondly, such levies and taxes increase transport resistance (see Chapter 6), and have had an impact on the overall level of car use and on modal split. Long-term fuel price elasticity for car use is in the magnitude of −0.25 (see Graham and Glaister, 2004, for a review of elasticities), which means that a 1 per cent increase in fuel prices reduces car use in the longer term by approximately 0.25 per cent. Note that most of the studies reviewed use data from the 1980s and 1990s, under conditions of lower incomes. Because people with higher incomes are less sensitive to price increases, the impact of fuel price increases now and in the future could be lower. Nevertheless these values indicate that prices do matter. Dargay (2007) estimates the car price elasticity of ownership to be

−0.12. For energy use and CO_2 emissions, elasticity values are higher than the fuel price elasticities because, in addition to the effect on car use, people buy more fuel efficient cars and tend to drive a little bit more efficiently if prices are higher (see Graham and Glaister, 2004), leading to an elasticity of approximately −0.77. Higher prices are a stimulus for higher load factors for transport.

Several cities and towns introduced parking charges in the 1960s and 1970s, often in combination with a reduction in the number of parking places in central urban areas. If parking is expensive, this increases transport resistance, leading to less car use and an increase in the use of alternatives.

Road pricing is probably the most controversial of all the categories of pricing policies. The controversy is not so much related to its effects but more to the difficulties in implementation. Nevertheless several examples of real-world implementation of road pricing schemes can be found, including Singapore, some Norwegian cities, London, Stockholm, and the German *Maut* system for lorries on motorways. Once road pricing is introduced, it may change transport resistance. On the one hand monetary costs increase; on the other hand, because of reduced traffic volumes, travel times might decrease (in Figure 2.1, the arrow from volume to resistance), with the first effect being generally stronger than the second effect, at least at the aggregate level (all road users).

Land-use planning

Land-use planning can have an impact on travel volumes and modal split (Chapter 5). Building at high densities can reduce travel distances, and building offices and residential neighbourhoods close to railway stations can increase the share of PT. Attractive neighbourhood design can increase the share of slow modes, and mixed land use (e.g. mixing housing, shops, schools and services) can reduce transport distances and increase the share of slow modes. In addition, because land-use planning can change the distribution of vehicle kilometres of road types and related driving behaviour, it can have an impact on emissions and safety levels, and on distances between the source of emissions (noise, pollutants) and the receivers. There is still debate on the impacts of land-use planning on travel behaviour (see Chapter 5), and this is partly explained by the fact that the actual impacts on travel behaviour seem to be less than expected. This conclusion may be due to the problems of measurement and the long-term effects that land use has on travel, but as with a lot of analysis it is the behavioural factors that are important, and our understanding of these responses is not good.

Infrastructure

In almost all countries infrastructure provision is a public matter. Even in the case of toll roads, such as in France and Portugal, the government decides which motorways will be built and where they will be located, and the government sets the institutional context including regulations, tendering and so on. It is beyond any doubt that the quality and quantity of road and rail networks have a major impact on transport volumes and modal split (see Chapters 3 and 6). In addition, these networks strongly influence the distribution of traffic across the networks, certainly in the case of roads, and therefore driving behaviour, route choice (see Chapter 7), safety levels and emission levels, and distances between the source of emissions and receivers. (See Figure 2.1 for the conceptualization of these impacts.)

Public transport policies

These refer to policies excluding policies involving public transport in one of the other categories of policies (such as pricing or subsidizing and infrastructure policies). Examples are routeing and frequency of buses and assistance of travellers at railway stations.

Marketing, education, information and communication

In marketing literature it is generally recognized that the independent role of these categories of policy instruments is limited, but it can have a supporting role. There is an impression that the government is much less successful in marketing than private parties. An exception might be policies related to information provision for public transport and dynamic traffic management: information in these categories might influence modal choice (in the case of information provision for PT) and route choice (dynamic traffic management).

10.8 Long-term environmental challenges

Over the next 50 years, the dominant environmental problems related to transport will be energy and resource use, climate change, the use of raw materials and noise nuisance. Local air pollution will still potentially be a problem, but technology will strongly reduce those impacts (see Chapter 8). Here, we only discuss climate change.

The European Union has expressed ambitious goals related to CO_2 emissions and transport: a reduction of at least 60 per cent of GHGs by 2050 compared to 1990 is required from the transport sector (EC, 2011). By 2030, the European goal for transport will be to reduce GHG emissions to around 20 per cent below their 2008 level (EC, 2011). Even if these European ambitions are attainable, it is expected that the worldwide CO_2 emission shares of aviation and maritime transport will still probably increase substantially (Banister, 2009). Aviation and maritime transport are, namely, difficult transport sectors to tackle because of their global character. Furthermore, it is expected that much of the future transport CO_2 emission growth will occur in the rapidly growing economies such as the BRIC countries (Brazil, Russia, India and China) (see also Figure 10.1).

If more countries follow the European Union's ambitions, it is clear that the task of achieving these ambitions will be hard. Owing to long lead times for low-CO_2 technologies, this would possibly mean a radical shift in lifestyle, which is not – generally speaking – a popular thought for most people or politicians. Generally, decoupling or decarbonizing means lowering the ratio of CO_2 per unit of GDP (Goldemberg and Tadeo Prado, 2010) or distance travelled (Banister and Stead, 2003). Decarbonization has been put on the agenda in the Kyoto Protocol Agreement (led by the United Nations Framework Convention on Climate Change) and more recently in the Copenhagen, Cancun and Durban summits. Yet these agreements have not resulted in internationally coordinated CO_2 mitigation action in transport.

Next to passenger transport, freight transport also faces major challenges that must be tackled for decarbonizing the global transport freight sector. The first challenge is to reduce off-shoring of manufacturing and CO_2 leakage to regions outside Europe and the USA so that CO_2 emissions in global transport can be reduced. The second is to find a mechanism to allocate emission reductions equitably, as a large quantity of CO_2 emissions from the BRIC countries result from the export of goods to the OECD countries. This leads to the third challenge, namely, to reduce consumption of key CO_2-intensive products, which stimulate the transport and manufacturing CO_2 footprints of China or other fast-growing regions. Finally, it seems important to support innovation in low-CO_2 transport in less developed nations.

Long-term options to reduce oil dependency and CO_2 emissions

Options to reduce oil dependency and CO_2 emissions can be clustered according to their potential emission reduction and the degree of uncertainty in emission reduction potential, the costs or the side-effects (Table 10.6).

Table 10.6 Examples of CO_2mitigation options clustered by emission reduction potential and uncertainty in potential, costs and/or side-effects

CO_2 emission reduction potential	Uncertainty in potential, costs or side-effects	
	Small to moderate	Moderate to large
Small to moderate	– Current carbon dioxide efficiency standards. – Pricing measures, e.g. ETS for aviation and shipping, EU-wide road pricing for trucks. – Energy efficiency measures for road freight, shipping, aviation. – Logistical efficiency measures, e.g. green logistics. – Land-use planning.	– First-generation biofuels (ethanol, bio-diesel). – Current commercial jet biofuels.
Moderate to large	– Plug-in hybrid cars. – Heavy oil biofuel substitutes for inland shipping and maritime transport.	– Full electric cars. – Fuel-cell hydrogen road vehicles. – Second-generation biofuels for road vehicles (ethanol, bio-diesel). – Second-generation and third-generation jet biofuels. – Biomass-to-liquid biofuels with carbon capture and storage.

Source: Netherlands Environmental Assessment Agency (2009).

The options include:

1. Technologies that may deliver large-scale emission reductions over time, but they are associated with a high degree of uncertainty, such as second-generation and third-generation biofuels, biomass-to-liquid jet fuels, hydrogen and fuel cells (see also Chapter 8). These technologies require further technical progress leading to performance improvement and cost reduction, and also require radical changes in areas such as vehicle production, fuel supply and agricultural systems.
2. Options that have a low emission reduction potential and a low degree of uncertainty in technology or cost. Uptake of these technologies may lead to a small increase in vehicle or transport prices, which may be offset by fuel savings. These technologies and measures are considered 'no-regret' measures on the way to low-carbon transport. For instance, a target of 90 g of CO_2 per km (new cars) could be adopted as a global target (IEA, 2009b). This would require global car manufacturers and car makers of

the BRIC economies to adopt the most stringent fuel economy standards. One possible disadvantage is that the CO_2 standards of new cars can lead to rebound effects (that is, an increase in energy use as its price falls because of improved fuel economy) on fuel consumption (Bonilla, 2009; van Dender, 2009). Another option in this class could also include adopting an active travel strategy aimed at increasing walking and cycling (Woodcock et al., 2007). Active travel is attractive because it brings co-benefits besides the fuel-saving effects of walking and cycling (not tabulated in Table 10.6). These policies involve interesting co-benefits: the mitigation of emissions (and climate change) also tackles other objectives, such as disease prevention and the promotion of public health (IEA, 2010).

3. Technologies and policy measures that have a high emission reduction potential and a low degree of uncertainty in technology or costs. In these cases, industry and consumers will have to accept the cost in order to benefit from reduction in carbon dioxide emissions. This is the case for biofuel substitutes in heavy-duty vehicles and shipping. Use of biofuels in maritime transport and heavy-duty road vehicles does not pose any fundamental or insurmountable technology challenges. The key barriers to biofuels are economic rather than technical, particularly for biofuels replacing marine diesel fuel (Christ, 2009).
4. Technologies that have low potential for emission reduction and major side-effects. This is the case for the first generation of biofuels and current commercial jet biofuels. These technologies have low and uncertain reduction potential, and it would not be wise to pursue them further.

Criteria to evaluate candidate policy options

To meet ambitious reductions in the transport sector, it is very likely that a transition to another energy system is needed. Candidate technological options for such a transition have been presented above. An important question is which criteria should play a role in the discussion on the advantages and disadvantages of the available technological options. These include expectations with respect to all of the following, both individually and in combination:

1. CO_2 emission reductions;
2. costs;
3. other environmental impacts (such as emissions of pollutants, noise);
4. risks, or at least risk perceptions (these could be relevant in the case of hydrogen and CO_2 capture and storage);
5. land-use and spatial impacts (especially in the case of biofuels – the

production of biofuels competes with food production and nature or biodiversity);

6. psychological factors (e.g. will the car driver accept hydrogen or electric vehicles?);
7. legal, institutional and political factors (important barriers might exist, e.g. legal barriers, the position of interest groups, political acceptability);
8. the position of important stakeholders, such as car producers, oil companies and the agricultural sector (in the case of biofuels);
9. the question of whether transitions can take place evolutionarily starting with the current system, or if radical changes are needed, resulting in a 'difficult' transition phase.

The wider context of CO_2 policies for transport

Achieving low-carbon transport implies a strong link between the transport, energy and agricultural sectors. Firstly, regardless of which non-carbon energy carrier is used for rail, urban transport and medium-distance transport (electricity or hydrogen), it has to be produced with low carbon emissions. Without a low-carbon energy sector, low-carbon transport is not feasible. Secondly, failure to decrease carbon dioxide emissions from transport will most likely need to be compensated in the energy sector. Thirdly, biofuel seems to be one of the few feasible low-carbon or zero-carbon options for aviation, shipping and road freight. But, given the limitation in bio-energy potential and likely negative side-effects of energy-crop production, bio-energy needs to be directed strategically to applications that maximize its contribution to decreasing carbon dioxide emissions and minimize the required inputs (Netherlands Environmental Assessment Agency, 2009).

The key important question here is that, if there is a perspective on an energy system that will be climate neutral (not emit CO_2 or other greenhouse gases), can other transport-related environmental policies be abandoned? We think the answer is no, for several reasons. Firstly, there are other environmental problems to be solved, such as noise nuisance, barrier effects of infrastructure, and possibly air pollution. Secondly, it will take decades before the transition is finalized, and conventional policies will be needed for the interim period. Thirdly, even if environmental problems are solved, accessibility-related problems (including congestion) will probably not be solved. Conventional policies, including infrastructure and land-use policies, might be needed to reduce such problems. Fourthly, energy production, in the case of a climate neutral energy system, may cause other problems, such as those related to the production of biofuels, which are reduced by conventional policies that aim to make vehicles more fuel efficient.

The use of non-renewable raw materials has recently received increasing attention because some modern technologies need 'rare' materials that are expensive and scarce and can often only be found in a few places worldwide. On the other hand, scarce materials may in the future be substituted by other materials. An example of substitution is from the 1980s and early 1990s, when it was thought that the use of three-way catalytic converters (which reduce most exhaust emissions of cars by more than 80 per cent) could result in a shortage of platinum, but substitutes became available. Noise pollution is expected to exist for at least the next two to three decades, but in urban areas electric vehicles may reduce noise pollution. On motorways and other roads with speed limits of over 50 km/h, quiet vehicles have little impact because the tyre noise dominates engine and exhaust noise (see section 10.5).

10.9 Conclusions

The most important conclusions of this chapter are:

1. Road, rail, air and waterborne transport contribute to a wide range of environmental impacts, ranging from climate change at the global level to local air pollution and noise at the local level.
2. There are many policy instruments available to reduce the environmental impact of transport, including pricing, land-use planning, infrastructure policies, public transport policies, marketing, and information and communication. These instruments can influence the environmental pressure of transport by reducing transport volumes, changing modal split, and influencing the technologies used, the efficiency of vehicle use and the way vehicles are used. In many ways the options are clear, but the problem is the effective implementation of the measures so that desired environmental outcomes can be achieved.
3. In the future, the dominant environmental problems related to transport will probably be energy use, climate change, health, the use of raw materials and noise nuisance.
4. The most challenging problems for the coming decades are probably CO_2 emissions and oil dependency. Worldwide CO_2 emissions from transport are increasing, despite emission reduction targets at several levels. Oil consumption through transport is increasing and so is oil dependency, despite the possibility of peak oil in the (near?) future.
5. To achieve strong CO_2 emission reductions in the transport sector,

new technological solutions are needed. There is a range of technology options that may deliver large-scale emission reductions over time, but these carry with them a high degree of uncertainty and may have important side-effects.

NOTES

1 This section is based on part of Banister et al. (2011).
2 Arrhenius developed a theory to explain the Ice Ages, and in 1896 he was the first scientist to speculate that changes in the levels of carbon dioxide in the atmosphere could substantially alter the surface temperature through the greenhouse effect.
3 This section is largely based on van Wee (2009).

REFERENCES

Air Resources Board (2011), *Key Events in the History of Air Quality in California*, accessed June 2011 at http://www.arb.ca.gov/html/brochure/history.htm.

Aleklett, K. (2008), 'Peak oil and the evolving strategies of oil importing and exporting countries', in *Oil Dependence: Is Transport Running Out of Affordable Fuel?*, European Conference of Ministers of Transport, ECMT RT 139, Paris: ECMT, pp. 37–114.

Banister, D. (2009), 'Getting in the right lane: low carbon European transport beyond 2050', paper prepared for the Netherlands Environmental Assessment Agency and the Stockholm Resilience Centre, The Hague, July, pp. 1–28.

Banister, D., K. Anderton, D. Bonilla, M. Givoni and T. Schwanen (2011), *Transportation and the Environment*: *Annual Review of Environment and Resources*, 36, 247–270.

Banister, D. and D. Stead (2003), 'Reducing transport intensity', *European Journal of Transport Infrastructure Research*, 2 (2/3), 161–178.

Blok, P. and B. van Wee (1994), 'Het Verkeersvraagstuk', in F. Dietz, W. Hafkamp and J. van der Straaten (eds), *Basisboek milieu-economie*, Amsterdam and Meppel: Boom, pp. 215–234.

Bonilla, D. (2009), 'Fuel demand on UK roads and dieselisation of fuel economy', *Energy Policy*, 37, 3769–3778.

Bovy, P.H.L., J. van der Waard and A. Baanders (1991), 'Substitution of travel demand between car and public transport: a discussion of possibilities', paper presented at the PTRC Summer Annual Meeting.

Browne, M., C. Rizet, S. Anderson, J. Allen and B. Keita (2005), 'Life cycle assessment in the supply chain: a review and case study', *Transport Reviews*, 25, 761–782.

Chapman, L. (2007), 'Transport and climate change: a review', *Journal of Transport Geography*, 15, 354–367.

Chester, M.V. and A. Horvath (2009), 'Environmental assessment of passenger transportation should include infrastructure and supply chains', *Environmental Research Letters*, 4 (2009), 024008 (8 pp.).

Christ, P. (2009), 'Greenhouse gas emissions reduction potential from international shipping', Discussion Paper no. 2009-11, Joint Transport Research Centre of the OECD and International Transport Forum, Paris.

Dargay, J.M. (2007), 'The effect of prices and income on car travel in the UK', *Transportation Research Part A*, 41, 949–960.

Davis, S., S.W. Diegel and R. Boundy (2010), *Transportation Energy Databook*, edition 29, Oak Ridge, TN: Oak Ridge National Laboratory.

Den Boer, L.C., F.P.E. Brouwer and H.P. van Essen (2008), *STREAM: studie naar Transport Emissies van Alle Modaliteiten*, Delft: CE Delft.

Dorland, C. and H.M.A. Jansen (1997), *Externe Transport: The Netherlands. Dutch Case Studies on Transport Externalities*, Amsterdam: Institute for Environmental Studies (IVM), Free University.

EC (2002), *Position Paper on Dose Response Relationships between Transportation Noise and Annoyance*, Luxembourg: Office for Official Publications of the European Communities.

EC (2007), *Limiting Global Climate Change to 2 Degrees Celsius: The Way Ahead for 2020 and Beyond*, COM (2007) 0002 final, Brussels: European Commission.

EC (2011), *White Paper: Roadmap to a Single European Transport Area – Towards a Competitive and Resource Efficient Transport System*, COM(2011) 144 final, Brussels: European Commission.

EEA (2010a), accessed May 2011 at http://www.eea.europa.eu/data-and-maps/indicators/final-energy-consumption-by-sector-1/final-energy-consumption-by-sector-5.

EEA (2010b), *The European Environment, State and Outlook: Synthesis*, Copenhagen: European Environment Agency.

EEA (2010c), *European Union Emission Inventory Report 1990–2008 under the UNECE Convention on Long-Range Transboundary Air Pollution (LRTAP)*, EEA Technical Report no. 7/2010, Copenhagen: European Environment Agency.

El-Shawarby, I., K. Ahn and H. Rakha (2005), 'Comparative field evaluation of vehicle cruise speed and acceleration level impacts on hot stabilized emissions', *Transportation Research Part D*, 10 (1), 13–30.

Ernst, J. (2011), 'Environmental challenges of urban transport: the impacts of motorization', in H. Dimitriou and R. Gakenheimer (eds), *Urban Transport in the Developing World: Perspectives from the First Decade of the New Millennium*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Evans, J.S., S.K. Wolff, K. Phonboon, J.I. Levy and K.R. Smith (2002), 'Exposure efficiency: an idea whose time has come?', *Chemosphere*, 49, 1075–1091.

Eyre, N.J., E. Ozdemiroglu, D.W. Pearce and P. Steele (1997), 'Fuel and location effects on the damage costs of transport emissions', *Journal of Transport Economics and Policy*, 31 (1), 5–24.

Gilbert, R. and A. Perl (2008), *Transport Revolutions: Moving People and Freight without Oil*, London: Earthscan.

Goldemberg, J. and L. Tadeo Prado (2010), 'Decarbonising the world energy mix', *Energy Policy*, 38, 3274–3276.

Goodwin, P.B. (1996), 'Empirical evidence on induced traffic: a review and synthesis', *Transportation*, 23, 35–54.

Goodwin, P., J. Dargay and M. Hanly (2004), 'Elasticities of road traffic and fuel consumption with respect to price and income: a review', *Transport Reviews*, 24, 275–292.

Graham, D.J. and S. Glaister (2004), 'Road traffic demand elasticity estimates: a review', *Transport Reviews*, 24 (3), 261–274.

Grahame, T.J. and R.B. Schlesinger (2010), 'Cardiovascular health and particulate vehicular emissions: a critical evaluation of the evidence', *Air Quality, Atmosphere, and Health*, 3 (1), 3–27.

IEA (2009a), *How the Energy Sector Can Deliver on a Climate Agreement in Copenhagen: Special Excerpt of the World Energy Outlook for the Bangkok UNFCCC Conference*, Paris: IEA.

IEA (2009b), 'Transport, energy and CO_2: Moving toward sustainability – how the world can achieve deep CO_2 reductions in transport by 2050', Press Release (09)15, accessed February 2012 at http://www.iea.org/press/pressdetail.asp?PRESS_REL_ID=293.

IEA (2010), *Key World Energy Statistics*, Paris: IEA.

IPCC (2007), *Climate Change 2007: Synthesis Report. Summary for Policymakers*, Intergovernmental Panel on Climate Change, accessed June 2011 at http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr_spm.pdf.

ITF (2011), International Transport Forum, accessed June 2011 at http://www.internationaltransportforum.org/statistics/CO2/index.html.

Janssen, N.A.H., B. Brunekreef, G. Hoek and P. Keuken (2002), *Traffic-Related Air Pollution and Health*, Utrecht and Apeldoorn: Utrecht University, Institute for Risk Assessment Sciences, Environmental and Occupational Health Division/TNO Milieu, Energie en Procesinnovatie.

Maibach, M., C. Schreyer, D. Sutter, H.P. van Essen, B.H. Boon, R. Smokers, A. Schroten, C. Doll, B. Pawlowska and M. Bak (2007), *Handbook on Estimation of External Cost in the Transport Sector: Internalisation Measures and Policies for All External Costs of Transport (IMPACT)*, Delft: CE Delft.

Marshall, J.D., T.E. McKone, E. Deaking and W.W. Nazaroff (2005), 'Inhalation of motor vehicle emissions: effects of urban population and land area', *Atmospheric Environment*, 39, 283–295.

Marshall, J.D., W.J. Riley, T.E. McKone and W. Nazaroff (2003), 'Intake fraction of primary pollutants: motor vehicle emissions in the South Coast Air Basin', *Atmospheric Environment*, 37 (2003), 3455–3468.

Meadows, D.H., D.L. Meadows, J. Randers and W.W. Behrens III (1972), *The Limits to Growth: A Report for the Club of Rome Project*, New York: Universe Books.

Meadows, D., J. Randers and D. Meadows (2005), *The Limits to Growth: The 30 Year Update*, London: Earthscan.

Netherlands Environmental Assessment Agency (2009), *Getting into the Right Lane for 2050: A Primer for Debate*, The Hague and Bilthoven: Netherlands Environmental Assessment Agency.

Newton, P.N. (ed.) (1997), *Reshaping Cities for a More Sustainable Future: Exploring the Link between Urban Form, Air Quality, Energy and Greenhouse Gas Emissions*, Research Monograph 6, Melbourne: Australian Housing and Urban Research Institute (AHURI).

Pope, C.A., III and D.W. Dockery (2006), 'Health effects of fine particulate air pollution: lines that connect', *Journal of the Air and Waste Management Association*, 56 (6), 709–742.

Rietveld, P. (2002), 'Why railway passengers are more polluting in the peak than in the off-peak: environmental effects of capacity management by railway companies under conditions of fluctuating demand', *Transportation Research Part D*, 7 (5), 347–356.

Simmons, M. (2005), *Twilight in the Desert: The Coming Saudi Oil Shock and the World Economy*, London: John Wiley.

Smith, K.R. (1993a), 'Fuel combustion, air pollution, and health: the situation in developing countries', *Annual Review of Energy and the Environment*, 18, 529–566.

Smith, K.R. (1993b), 'Taking the true measure of air pollution', *EPA Journal*, 19 (4), 6–8.

US Department of Transportation, Federal Highway Administration (2011), accessed June 2011 at http://www.fhwa.dot.gov/environment/noise/noise_barriers/inventory/summary/stable71.cfm.

Van Dender, K. (2009), 'Energy policy in transport and transport policy', *Energy Policy*, 37, 3854–3862.

Van der Toorn, J.D. and T.C. van den Dool (1997), *Geluidemissie door motorvoertuigen: klassieke metingen en analyses met de Syntakan* [Noise emission by motor vehicles: classic measurements and analyses using the Syntaken], Delft: TNO-TPD/TU Delft.

Van Wee, B. (2009), 'Transport policy: what it can and can't do', paper presented at the European Transport Conference, Noordwijkerhout, 4–6 October.

Van Wee, B., P. Janse and R. van den Brink (2005), 'Comparing environmental performance of land transport modes', *Transport Reviews*, 25 (1), 3–24.

Wang, M., H. Huo, L. Johnson and D. He (2006), *Projection of Chinese Motor Vehicle Growth, Oil Demand and CO_2 Emission through 2050,* Oak Ridge, TN: US Department of Energy, Office of Scientific and Technical Information.

WHO (World Health Organization) (2009), *Global Status Report on Road Safety*, Washington, DC: WHO.

WHO and JRC, European Commission (2011), *Burden of Disease from Environmental Noise: Quantification of Healthy Life Years Lost in Europe*, Geneva: World Health Organization.

Woodcock, J., D. Banister, P. Edwards, A.M. Prentice and I. Roberts (2007), 'Energy and transport', *The Lancet*, 370, 1078–1087.

World Commission on Environment and Development (1987), *Our Common Future*, Oxford: Oxford University Press.

11

Traffic safety

Fred Wegman

11.1 Introduction

The number of road crashes, fatalities and injuries is considered unacceptably high in many countries. This can be derived, for example, from the fact that the European Commission (EC, 2010) announced the ambition of halving the overall number of road deaths in the European Union by 2020, starting from 2010. At this stage a European target for reducing serious injuries is not possible, owing to a lack of a common definition of serious and minor injuries. In many highly developed and highly motorized countries the number of fatalities has decreased over the last few decades (OECD/ITF, 2011a). However, this favourable development cannot be observed in low- and middle-income countries so far (Peden et al., 2004).

Risks in road traffic are considerably higher than in other transport modes, and the amount of injuries in road traffic is far higher than the numbers in trains, planes or ferries. Although crashes in these other modes attract a lot of public and media attention, road crashes kill far more people, but in a 'diluted' way, resulting in only limited media coverage and relatively limited attention from the public and politicians. At the same time, serious road crashes are tragedies at a personal level. Road crashes can happen to everybody, anytime, anywhere, and they are unexpected. Often the lives of young people and their families are suddenly changed.

This chapter aims to give a concise introduction to road safety. Using this chapter the reader will be able to explain some basic concepts of road safety, get an insight into some recent traffic safety developments worldwide and be able to talk about a new policy vision and options for how to reduce further crashes and (serious) injuries. The relevance of various technologies was discussed in Chapter 8.

Risk factors in traffic are discussed in section 11.2. Section 11.3 treats the subject of identifying the causes of crashes. In section 11.4 there is an explanation of three important components of road traffic when it comes to risks: transport modes, age of road users and road types. In section 11.5 the difficulties of measuring road safety (danger) are discussed. Some developments in road safety are given in section 11.6. Section 11.7 explains the development in dominant thoughts about traffic safety. This section shows that the amount of knowledge on causes for road accidents and how to implement successful policies has increased dramatically over the years. Still, the next steps for further improvements can be made. Scientific information to support this statement and one such next step, Sustainable Safety, are presented in section 11.8. The chapter's main findings are presented in section 11.9.

11.2 Risk factors in traffic

Taking part in traffic is a dangerous affair in itself. This is due to some fundamental risk factors in traffic (sometimes also denoted as basic factors): the vulnerability of the body of road users in combination with speed levels in traffic as well as the presence of objects with large mass and/or stiffness with which one can collide. In addition, there are also road user factors that increase crash risk, such as alcohol use, fatigue or distraction.

Fundamental risk factors

Fundamental risks are inherent to road traffic and are the basis of the lack of safety in current road traffic. These are a combination of factors such as speed and mass (and the kinetic energy in a crash) and the vulnerability of the human body.

Speed is related to the risk of being involved in a crash (for an overview, see Aarts and van Schagen, 2006). Higher absolute speeds of individual vehicles are related to an exponential increase in risk. If the average speed on a road increases, then the increase in crash risk can be best described as a power function: a 1 per cent increase in average speed corresponds with a 2 per cent increase in injury crashes, a 3 per cent increase in serious injury crashes and a 4 per cent increase in fatal crashes (Nilsson, 2004). With the same increase in speed, for both individual speed and average road section speed, an increase in risk is higher on urban roads than on rural roads and motorways. Speed is also related to crash severity. This is based on the kinetic energy (of which speed is an important component), which is converted into other energy forms and/or bodily damage during a crash. Injury risk (the chance of being injured in a crash) is also determined by (impact) speed level, the relative directions of crash partners, their mass differences and the protection level.

Speed differences are also linked with increases in crash risk. Aarts and van Schagen (2006) show that it has not been proven that vehicles travelling at lower speeds than the traffic flow have a higher risk than vehicles that go with the flow. However, vehicles going faster than the traffic flow have an increased risk (Aarts and van Schagen, 2006). Speed variance at the level of a road section is also linked to increased crash risk.

Mass differences are also fundamental risk factors. In a crash between two incompatible parties, the lighter party is at a disadvantage because this party absorbs more kinetic energy and the vehicle generally offers less protection to its occupants than a heavier vehicle. Mass differences between colliding objects can amount to a factor of more than 300 (a pedestrian weighing 60 kg versus a heavy goods vehicle weighing 20 000 kg). Furthermore, in view of their stiffness and structure, heavier vehicle types generally offer better protection to their occupants in the event of a crash. For occupants of vehicles with a high mass, injury risk is much lower than that of the lighter crash party. If we assume the injury risk for a crash party of an 850 kg passenger car as 1, then the injury risk for an average crash partner is 1.4 if the car weighs 1000 kg, and 1.8 if the vehicle weighs more than 1500 kg (Elvik and Vaa, 2004).

Finally, vulnerability is to be considered a fundamental risk factor. Several methods can be used to protect the human body in a crash, foremost by improving the crashworthiness of a vehicle. Over the years great progress has been made to improve vehicle design to protect car occupants. The most famous example is the use of seat belts in combination with airbags. Glassbrenner and Starnes (2009) estimate that seat belts reduce fatality and injury risks by more than 40 per cent, and in combination with airbags by more than 50 per cent. However, vulnerable road users such as pedestrians and cyclists have almost no possibilities to protect themselves from injury risk in a crash. Only a crash helmet for (motorized) two-wheelers can be considered, and some developments of airbags for motorcyclists can be seen in practice. Furthermore, modern car designers try to incorporate safety features when designing a car front, which aim to be safer for pedestrians and cyclists in the case of an accident.

Risk-increasing factors

Besides these fundamental risk factors, road traffic also has to contend with risk-increasing factors caused by road users:

- **Lack of driving experience.** Lack of driving experience results in higher risks. The effect of (lack of) driving experience on crash risk is strongly linked to age effects. Since driving experience is strongly correlated with age and as both factors are associated with specific characteristics which increase risk, it is difficult to separate the effects of age and experience. About 60 per cent of the (relatively high) crash risk for novice drivers (broadly speaking, people who have driven less than 100 000 kilometres) can be explained by lack of driving experience, and the other 40 per cent is age related (see Wegman and Aarts, 2006). The increased crash risk for novice drivers decreases rapidly within the first year after passing a driving test (Vlakveld, 2005). Male novice drivers especially run an additional risk (a factor of 10) compared to more experienced drivers (male and female) and also compared to novice female drivers (a factor of 2.5).
- **Psycho-active substances: alcohol and drugs.** Alcohol consumption by road users is one of the most important factors that increase risk in traffic. Crash risk increases exponentially with increased blood alcohol content (BAC). Compared to sober drivers, the crash risk is a factor of 1.3 with a BAC between 0.5 and 0.8 per mille, a factor of 6 with a BAC between 0.8 and 1.5 per mille, and a factor of 18 above 1.5 per mille (Blomberg et al., 2005).The crash risk of road users under the influence of psycho-active substances (Walsh et al., 2004) can be about a factor 25 higher with the use of drugs. This risk can even increase up to a factor of 200 with the combined use of alcohol and drugs, relative to sober road users, also depending on the quantity of alcohol consumed; there is cumulative road crash fatality risk when combined with the use of alcohol and drugs. Drugs in traffic is not a very mature area of research and policy-making; however, it has received quite a lot of (political) attention recently.
- **Illnesses and ailments.** Visual limitations or ailments are generally associated with a very small increase in crash risk (on average a factor of 1.1), relative to healthy people (see Vaa, 2003). Further examination indicates that crash risk is higher under two conditions: reduced useful field of view (UFOV) and glare sensitivity. Decreased hearing only results in a slightly increased risk. People with Alzheimer's disease run a risk of crash involvement which is twice as high as that of healthy people. Other psychiatric disorders, such as cognitive disorders and depression, result in a slightly increased risk with a factor 1.6, on average.
- **Emotion and aggression.** During the past few years some have expressed the view that aggression in traffic is a major contributor to road crashes. Several questionnaire studies show the (positive) relationship between self-reported aggressive behaviour (offending behaviour)

and self-reported road crash involvement. However, this does not imply a causal relationship between the two elements. It is also the case that aggressive behaviour coincides with risk-seeking behaviour. This makes it difficult to draw conclusions about the relationship between aggression and road safety. The literature leaves the impression that there is a coherent behavioural pattern of a combination of various aggressive and/or risky behaviour types that result in a dangerous driving style. However, for the time being it is not possible to quantify the risk associated with this risk factor (Mesken, 2006).

- **Fatigue.** Fatigue is most probably a much more frequently occurring factor in increasing risk than data from police reports show. Participating in traffic whilst fatigued is dangerous because, in addition to the risk of actually falling asleep behind the steering wheel, fatigue reduces the general ability to drive (keeping course), reaction time and motivation to comply with traffic rules. Research shows that people suffering from a sleep disorder or an acute lack of sleep have a 3 to 8 times higher risk of injury crash involvement (Connor et al., 2002).
- **Distraction.** Like fatigue, distraction is probably a much more frequent crash cause than reported police data show (Regan et al., 2009). Currently, one of the more common sources of distraction is use of the mobile phone while driving. The hands-free option, permitted in many countries, does not reduce the effect of distraction either (McEvoy et al., 2005). This research indicates that using a mobile phone while driving results in an increase in risk by a factor of 4 relative to non-users. Other activities such as operating route-navigation systems, tuning CD-players and radios and so forth can also be a source of distraction, as can activity such as eating, drinking, smoking and talking with passengers (Young et al., 2003).

11.3 Cause: 'unintentional errors' or 'intentional violations'?

In identifying the cause of crashes in whatever system, 'man' is always quoted as the most important cause of crashes. People make errors, no matter how hard they try. At the same time, people do not always (consciously or otherwise) obey rules and regulations designed to reduce risks. The question arises: how serious are offences actually for road safety and with what frequency do they cause traffic crashes? However, this section will show that no clear picture emerges from the research of the relative contribution to crashes by intentional violations and unintentional errors.

A Canadian study looked into the relationship between violations and crashes as evidenced by driver behaviour (Redelmeier et al., 2003). The research team tracked car drivers who were convicted of causing a fatal crash and recorded the crash involvement of these offenders in the period following the conviction. The first month after the penalty, the chance of being involved in a fatal crash was 35 per cent lower than could be expected on the basis of coincidence. The authors attributed this effect to the fact that there were fewer traffic violations immediately after the period in which the drivers were fined. However, this benefit lessened substantially over time and disappeared after three to four months. Out of the above research, a strong relationship emerges, particularly between violations and crash involvement. It must be emphasized, however, that this type of research cannot prove anything conclusive about causality between the two phenomena.

Thus both errors and (intentional) violations (and related extreme behaviour) play a role in the cause of crashes and therefore deserve a place in road safety policies. How large the share of (unintentional) error and (intentional) violation is exactly cannot be stated, based on current knowledge. The role of (unintentional) error seems to be the more important one. Unfortunately, the information that can be extracted from police registration forms about crash causes cannot be used to identify the underlying causes of crashes. This is not surprising given that the data are gathered primarily with the objective of being able to identify the guilty party, rather than identifying precisely the underlying causes of a crash. It should also be remembered that crashes are always the result of a combination of factors.

On the one hand, it is logical that unintentional errors form the lion's share of crash causes, given that intentional offending in itself never leads directly to a crash. Violations certainly can increase the risk of error and the serious consequences of these errors. On the other hand, there is no evidence to support the widely held opinion that anti-social road hogs are the major perpetrators of crashes. Without doubt they cause part of the road safety problem, if only because other road users cannot always react appropriately to them. However, many crashes are the result of unintentional errors that everybody can make in an unguarded moment, as illustrated by Dingus et al. (2006).

Dingus et al. (2006) concluded that, in nearly 80 per cent of the recorded crashes and in 65 per cent of the near-crashes, driver inattention was involved just prior to the onset of the conflict. In this study drivers (100 cars) were followed for a year by observation systems installed in their cars: a black box and small cameras. The idea was to observe everyday behaviour. The role of driver inattention is rarely found on police registration forms, because who would tell the police that a cigarette fell to the floor just prior to the crash and that in a state of some panic he was trying to retrieve it? Therefore, it is time to rethink the idea many people, including road safety professionals and decision-makers, have of crashes only being caused by the traffic offences that are frequently found on police registration forms.

11.4 Transport modes, age groups and road types

A lot can be learned about the differences in risks by comparing transport modes and road types. These differences relate to the fundamental risk factors (section 11.2), speed, mass and vulnerability, in combination with protection. Users of motorized two-wheelers, for example, have the highest fatality and injury risk in road traffic (Table 11.1), which can largely be explained by a combination of high speed with the relatively low mass of the vehicle in conflict with other motorized traffic, as well as poor crash protection. In relation to crash protection, the relatively high amount of motorcycle deaths in the USA is striking (Table 11.1). The reason is the weakening of state helmet laws in the USA. Two-thirds of fatally injured motorcycle riders were not wearing a helmet in states without universal helmet laws (OECD/ITF, 2011b). On top of these factors, two-wheelers (especially mopeds) are popular with young people. This age group already has a relatively high risk in traffic because of age-specific characteristics and needs, and lack of experience (Table 11.2 and Figure 11.1).

Table 11.1 Road fatalities per 1 billion vehicle kilometres

	All[a)]	Mopeds[a)]	Motorcycles[a)]	Cars[a)]	Heavy goods vehicles[a)]	Bicycles[a)]
Austria	9	113.2	98.1	6.6	1.7	
Belgium	9.6		82.9	6.3	2.8	
Denmark	8.2	41[b)]	38	5	2	
France	7.8	125	135	6	2	
Ireland	4.9		84.3	3.8	2	
Netherlands	5.6	63	64	3	1.3	11
Slovenia	9.6		170	5	3	
Sweden	4.4	39.8	75.7	3.9	1.3	
Switzerland	5.7	60.2	34.4	2.6	0.3	
USA	7.1		227.9	5.8	5.2	

Notes:

Only those countries are shown for which data regarding risk per mode were presented.

[a)] The figures represent fatalities per users of the mode.

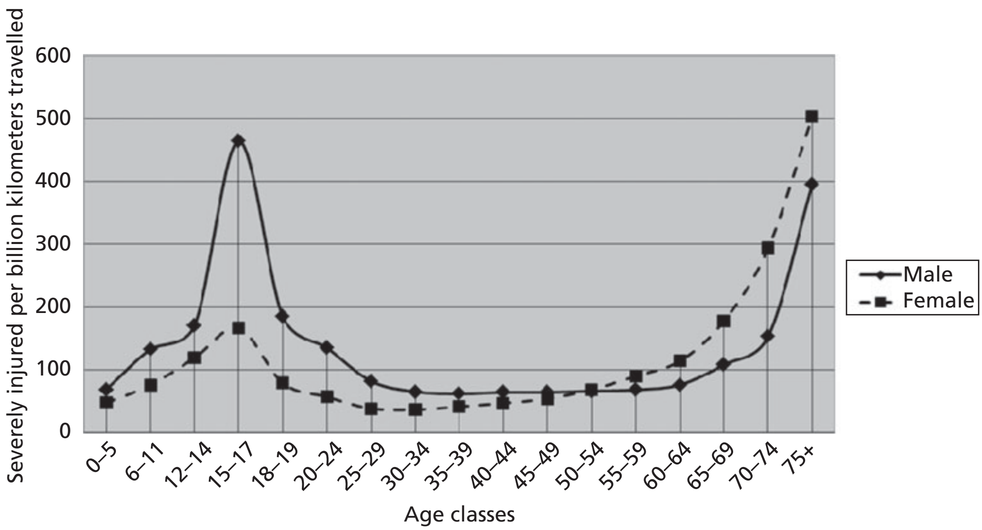
[b)] Bicycles and mopeds.

Source: OECD/ITF (2011a).

Table 11.2 Road user fatalities per 100 000 population of the age group and per 1 billion vehicle kilometres per road type

	Per 100 000 population of the age group							Per 1 billion vehicle kilometres	
	All	0–14	15–17	18–20	21–24	25–64	65+	All roads	Motorways
Australia	6.7	1.7	7.6	15.3	11.1	7.1	7.6	6.7	
Austria	7.6	1.2	9.6	16.5	11.7	7.2	11.0		2.5
Belgium	8.8	1.2	–	–	–	9.5	9.5	9.7	3.9
Canada	6.6	1.0	–	–	–	6.4	8.3	6.5	–
Czech Republic	8.6	1.1	3.5	10.2	16.1	9.4	10.5	15.8	3.1
Denmark	5.5	1.0	8.1	15.1	9.2	5.5	7.0	6.6	1.8
Finland	5.2	0.7	11.5	15.2	8.1	4.5	7.7	5.0	–
France	6.8	1.1	8.3	16.7	15.7	6.9	7.6	7.8	2.4
Germany	5.1	0.8	5.2	14.4	9.6	4.5	6.6	5.9	–
Great Britain	3.7	0.6	5.4	9.2	6.4	3.8	4.2	4.4	1.3
Greece	12.9	2.7	10.9	24.9	29.2	13.1	13.1	–	–
Hungary	8.2	1.4	4.4	7.0	10.5	9.5	10.1	–	–
Iceland	5.3	0.0	7.1	6.7	0.0	7.1	8.1	5.5	–
Ireland	5.4	1.3	5.3	18.9	12.8	5.2	5.3	4.9	–
Israel	4.2	1.2	3.1	7.1	7.4	4.4	8.3	6.4	–
Italy	7.1	0.8	6.8	12.8	13.8	6.8	9.2	–	–
Japan	4.5	0.7	3.7	6.4	4.3	3.0	10.2	7.7	1.7
Korea	12.0	1.9	6.6	7.0	9.8	11.6	35.2	20.0	–
Luxembourg	9.7	6.8	0.0	29.4	21.7	8.2	13.0	–	–
Netherlands	3.9	0.8	4.3	8.8	9.2	3.1	7.6	5.6	–
New Zealand	8.9	2.5	13.1	19.4	15.4	8.8	10.1	9.6	–
Norway	4.4	0.9	8.9	14.2	8.2	4.1	5.3	5.4	–
Poland	12.0	2.2	8.2	22.1	19.6	12.1	15.7	–	–
Portugal	7.9	1.3	5.7	12.2	13.3	8.1	10.8	–	–
Slovenia	8.4	0.7	6.4	15.1	17.6	8.2	11.7	9.6	5.6
Spain	5.9	0.9	5.6	11.0	9.1	6.4	6.6	–	–
Sweden	3.9	0.6	6.5	8.9	5.7	3.6	5.6	4.4	–
Switzerland	4.5	1.8	4.1	11.6	8.6	3.8	6.9	5.7	–
United Kingdom	3.8	0.6	5.4	9.9	6.5	3.8	4.3	4.6	–
USA	11.1	–	–	–	–	12.8	13.4	7.1	–

Source: OECD/ITF (2011b).



Source: Reworking and presentation by the author of data from CBS, Ministry of Infrastructure and the Environment, and DHD, 2011.

Figure 11.1 Number of severely injured people in traffic per 1 billion kilometres travelled per age group and sex for the Netherlands, 1999–2009

Currently, in many highly motorized countries car occupants have the major share of the total number of road fatalities because of the relatively high amount of kilometres travelled in cars. On the one hand, the car is a fast and weighty collision partner in conflicts with two-wheelers and pedestrians, who also include especially vulnerable road users such as children and the elderly. On the other hand, the car is the vulnerable party in terms of weight in conflicts with heavy goods vehicles and not very 'forgiving' roadside obstacles. Young people are an especially high-risk group of those involved in serious crashes because of their lack of driving or riding experience and age-specific characteristics. Elderly road users (of 75 years old or more; see Figure 11.1) are the next most important risk group because of their physical frailty. In many low- and middle-income countries the majority of the casualties are vulnerable road users such as pedestrians and cyclists, most of the time young people.

Differences of safety for different road types can also, to a large extent, be explained by a combination of the fundamental risk factors introduced earlier. For example, serious crashes outside urban areas, and particularly on rural roads, are dominated by single-vehicle conflicts along sections of road, often running off the road. These are usually the result of inappropriate speeds, possibly in combination with other factors which increase risk such as alcohol consumption, distraction and/or fatigue. The fact that many roadsides are not 'forgiving' also results in severe outcomes. On urban roads, transverse conflicts, in particular, predominate. On these streets and roads, in particular, where most people are killed in urban areas, mass differentials and the vulnerability of road users are important factors, combined with comparatively high speeds and the vulnerability of vehicles in transverse conflicts (side impacts). Motorways are the safest roads when it comes to crash risk (see Table 11.2, where for some countries in the two right-hand columns road fatalities are given per billion vehicle kilometres for all roads and motorways). This is due to a combination of high-quality road design and slow-moving traffic not being allowed on these roads. This is appropriate for high driving speed conditions, both physically (separation of driving direction, grade-separated intersections) and psychologically (predictable design). Then high speeds can be managed safely.

11.5 Measuring safety and danger

All countries in the world seem to have the ambition to improve road safety, or at least no country is known to be making public statements that the road toll of today is acceptable. However, measuring road safety is not as simple as measuring a temperature. Researchers or policy-makers cannot read a simple measuring instrument. Additionally, they can even have a discussion about which elements to include in a definition, and which not. The most common measure used to define road safety is the number of road crashes and/or the number of casualties and the associated negative consequences resulting from such crashes. Sometimes subjective feelings related to fears of being involved in a crash are included in the measure as well. In those cases people's perceptions about (lack of) road safety are taken into account in the measure.

The definition of a road traffic crash is a collision or incident on a public road (or private road to which the public has right of access) that results in damage to objects and/or injury to people and that involves at least one vehicle in motion. The international definition of a road death, taken from the UNECE Glossary of Transport Statistics 2009, is someone who dies immediately or dies within 30 days as a result of a road crash, excluding suicides. For countries that do not apply the threshold of 30 days, conversion coefficients are estimated for international comparison purposes. The definition of injury and injury severity can be classified as follows: admitted to hospital, had to take sick leave, had to rehabilitate, suffered permanent injury, remained in a coma, or died from the consequences more than 30 days afterwards. Different countries use different definitions and different procedures to collect data; from this perspective, comparison of data of numbers of injured are difficult to compare, to some extent. However, international efforts are being made to improve sound and meaningful comparisons (OECD/ITF, 2011c).

Crashes can result in more serious or less serious outcomes: fatal injuries, other injuries or damage only to vehicles involved in a crash. Sometimes, damage-only crashes are not considered serious enough to be included in official crash statistics. Data collection is needed to learn how many crashes occurred in a certain time period and in a certain geographical area. The longer the time period or the larger the area, the more crashes. For that reason it is a good habit to normalize the number of crashes for time and space, expressing the road safety level. This normalizing can be done in different ways serving different purposes. If we relate the number of fatalities or injuries to the number of inhabitants (the first ratio) we have the mortality rate (fatalities per 100 000 inhabitants; see also Table 11.2, where mortality rates are presented for different countries) or morbidity rate (injuries per 100 000 inhabitants). These rates are public health indicators, allowing us to compare road injuries with other threats or diseases. Mortality rates are often used in international comparisons. An important reason is that fatal road crashes have a common definition (dead within 30 days) and are well recorded in many countries, as is the case with the number of inhabitants. This is not the case for injuries.

A second ratio is the so-called fatality rate or injury rate. In this case we relate the number of fatalities or injuries to the degree to which people are exposed to traffic or, more precisely, to risks in traffic. Often, the number of kilometres travelled is used to estimate this 'exposure' or, even more often, the number of motorized kilometres (see Table 11.1). We can also use time in traffic as a measure of exposure.

Unfortunately, the measuring of road crashes, and their consequences, and the measuring of exposure suffer from problems related to the use of different definitions, data quality, data completeness and data availability (OECD/ITF, 2011c). In almost all countries the crash registration is carried out by the police. However, crash statistics are always incomplete as a result of underreporting. Furthermore, data collections suffer from certain biases: crashes with motorized vehicles are better registered than crashes with non-motorized transport, such as pedestrians and cyclists (Derriks and Mak, 2007). Another bias in data collection is that the higher the severity of injuries, the lower the underreporting. At present, many initiatives are being developed in the world to improve the quality, comparability and availability of crash data.

An important measure for road crashes is their associated costs. There are two good reasons to estimate road crash costs. Firstly, it allows policy-makers to compare the economic consequences of road crashes with other impacts of traffic and transport, such as environmental impacts and congestion. A second reason is that it allows policy-makers to compare these costs with the costs of other public health issues. For that purpose, public health indicators denoted as 'DALY' (disability adjusted life years) or 'QALY' (quality adjusted life years) are also sometimes used. These are measures for loss of life years and/or quality of life.

In many countries, a growing interest in estimating the costs of road crashes can be observed. The cost estimation methods have improved considerably. However, an internationally accepted 'standard' method does not exist at the time of writing this chapter.

Some convergence on the cost categories to be included in cost estimates can be noted (Alfaro et al., 1994). However, the methods applied still differ in including or excluding certain cost categories (Elvik, 1995):

1. medical costs;
2. production loss;
3. value of statistical life;
4. property damage;
5. settlement costs.

Sometimes costs related to congestion as a consequence of a crash are added. For the cost categories 1, 4 and 5, a method is used called 'restitution cost method' and for 2 the 'human capital method'. In methods 1, 2, 4 and 5, direct financial costs related to crash injuries are estimated, for example the amount of money hospitals have to spend on injury treatment, vehicle repair costs, lost production hours (e.g. lost wages) and so forth.

Cost category 3 (value of statistical life, VOSL) is based on people's willingness to pay for lower risks (or willingness to accept a reward for higher risks). In this cost category the change in traffic casualties due to a measure is valued in monetary terms (e.g. de Blaeij, 2003). A VOSL does not reflect the monetized value of an individual life, which is, naturally, priceless. Instead, the VOSL is based on the relation between changes in risks and willingness to pay for these changes. For example, if someone drives on a road with a risk of 2.5/1 000 000 of death, but is willing to pay 6 minutes by taking a detour to drive on a road with a lower risk of 2/1 000 000, this driver is valuing his or her 'statistical life' at 2 million euros. The reason is that the VOSL is (assuming a value of time of 10 euros/hour, which equals 1 euro/6 minutes; see Chapter 14,):

$$\frac{d(travel\ time)}{d(risk)} = \frac{1\ euro}{\left(\frac{0.5}{1\,000\,000}\right)} = 2\,000\,000\ euros$$

Societal costs of road crashes can be compared with other road traffic societal costs. In the Netherlands, for example, the yearly societal costs of road crashes (around 10 to 14 billion euros) are higher compared to congestion costs (2 to 3 billion euros) and the environmental costs of road traffic (2 to 8 billion euros) (KiM, 2010).

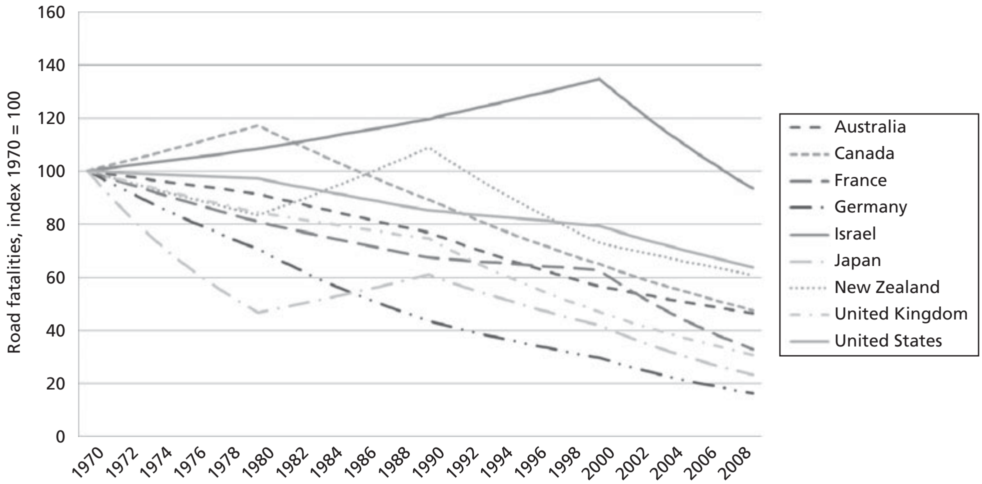
11.6 Developments in road crashes

Each year more than 1.2 million are killed in a road crash, and 20 to 50 million suffer non-fatal injuries worldwide (Peden et al., 2004). Peden et al. (2004) expect the number of road traffic fatalities to increase by 67 per cent over the period 2000 to 2020 to 1.9 million road fatalities a year, without appropriate action. With these numbers, road fatalities are the 11th leading cause of death, resulting in 2.1 per cent of all deaths. By far the majority of all crashes, deaths and injuries occur in low- and middle-income countries: 85 per cent of road traffic deaths (see Table 11.3 for road injury mortality rates per income class worldwide). The majority of these deaths and injuries are vulnerable road users. The economic costs of road crashes and injuries are estimated to be 1 per cent of GNP in low-income countries, 1.5 per cent in middle-income countries and 2 per cent in high-income countries. These developments resulted in a resolution adopted by the United Nations in 2010 (A/64/255) to declare the period 2011–20 a Decade of Action for Road Safety. This DoA started on 11 May 2011.

Table 11.3 Road traffic injury mortality rates (per 100 000 population) in WHO regions, 2002

WHO region	Low- and middle-income	High-income
African region	28.3	–
Region of the Americas	16.2	14.8
South-East Asia region	18.6	–
European region	17.4	11.0
Eastern Mediterranean region	26.4	19.0
Western Pacific region	18.5	12.0

Source: Peden et al. (2004).



Note: Index 1970 = 100.

Source: Based on OECD/ITF (2011a: 7, table 1), which gives this long-term trend in road fatalities for more countries worldwide.

Figure 11.2 Long-term trends in road fatalities for a selection of high-income countries, 1970–2009

Since 1970, many high-income countries have made remarkable progress (see Figure 11.2).

Three countries are the top-scoring countries worldwide, with mortality rates of less than 4.0: Sweden, the United Kingdom and the Netherlands (see Table 11.2). In the so-called SUNflower studies (Koornstra et al., 2002; Wegman et al., 2008) the road safety developments in these three countries were compared, and later six other European countries were added. The main conclusions (Koornstra et al., 2002) were that all three countries have achieved similar levels of safety through continuing planned improvements over recent decades and that policy areas targeted have been similar, but that implementation of policies has differed at a detailed level. It was assumed that differences in focus for safety programmes resulted from both different relative sizes of accident groups and differences in the structure of road safety capability, which influenced their ability to deliver different types of policy. Progress has been achieved through directing improved policies in all three areas – vehicle, road and road user.

An example: the Netherlands

To give more detail on the explanations for the relatively high rate of improvement in some countries, the Netherlands has been chosen as an example. Details for more countries can be found in *Safety Science*, Special

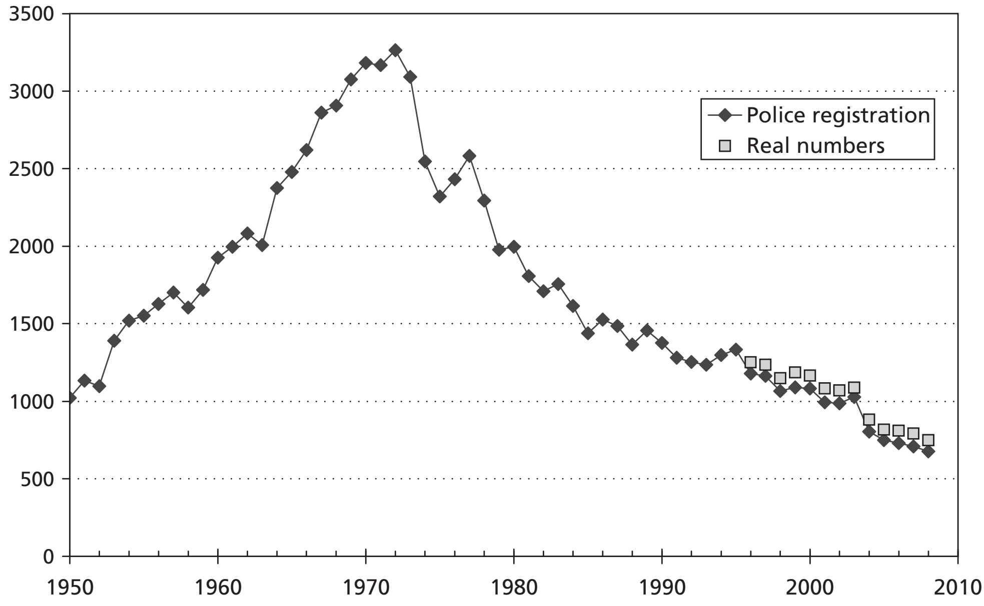


Figure 11.3 The development of the number of traffic deaths in the Netherlands

Issue on Road Safety Management (volume 48, issue 9, November 2010). A 50 per cent reduction in the mortality rate occurred in the period 1995–2007 in the Netherlands, whereas Great Britain and Sweden reached a little bit more than 20 per cent. This is partly due to a 'learning society' or an 'investing society', which has adapted itself to motorized, fast-moving traffic and making substantial safety investments at the same time. Infrastructural adaptation has taken place (such as the construction of relatively safe motorways), passive safety in vehicles has been improved, and there is more safety legislation and enforcement which takes account of factors which increase risk and reduce injury (such as alcohol consumption in traffic and mandatory crash helmet and seat belt use, respectively). These measures have all contributed to reductions in the number of traffic fatalities and injuries, despite increased mobility (Koornstra et al., 2002; Elvik and Vaa, 2004). But, as yet, researchers do not have a totally conclusive explanation for these improvements.

In the period 2008–2010, approximately 700 traffic deaths were lamented each year in the Netherlands. This is just a quarter of the 3264 traffic deaths in the disastrous year 1972. The number of traffic deaths in the Netherlands should not be based exclusively on police registration. Other sources reveal that approximately 8 per cent, in this period, of these deaths are missing from this registration. For this reason, in Figure 11.3, the concept of the 'real' number of traffic deaths during the last few years is used.

SWOV (2007) describes the major changes that occurred during the period 1950–2005 in a report with the striking title 'The summit conquered'. To begin with, there is a rise in the number of traffic fatalities, which is followed by a decline. This report illustrates that, for an understanding of why the annual number of fatalities has decreased, one should not look at the total number of traffic deaths; it is preferable to consider separate components (transport modes, age, road type, etc.), because these components develop differently compared with the totals. This approach shows that different developments take place concurrently. It becomes clear, for instance, that passenger car mobility in terms of vehicle ownership and vehicle use has been increasing steadily during this period. The sales and use of motorized two-wheelers, however, show a less steady picture: they fluctuate strongly and are sometimes popular, sometimes much less popular. This is clearly reflected in the road safety developments.

The quality of roads and vehicles with regard to safety has shown considerable improvement in the past few decades. The structure of the road network in the Netherlands has undergone considerable adaptations to meet the increased mobility. This can be illustrated by the fact that approximately half of all motorized vehicle kilometres travelled are travelled on relatively safe motorways. The separation of different traffic modes, mainly by the construction of safe bicycle facilities, has taken a considerable step forward. Secondary (passive) vehicle safety has been improved considerably.

Three important aspects of safety related human behaviour have also improved: drinking and driving has decreased, the safety belt is worn much more frequently, and the helmet for motorized two-wheelers is also worn much more frequently. These improvements are such that the behaviour problem related to drinking and driving, wearing safety belts and wearing helmets is reduced to 'only' a hard core of offenders. In the Netherlands, the speeds driven have gone down because the speed limits have been lowered on a substantial part of the road network. For driving speeds, it may be observed that, although road users have reduced their speed somewhat, a considerable proportion of road users exceed the limit.

11.7 Shifts in road safety paradigms

Section 11.6 explains that different countries in the world are at a completely different stage of development. Low- and middle-income countries, especially in Asia, face a serious risk of growing numbers of fatalities and injuries in the coming years. At the same time, we see a positive development in many highly motorized and highly developed countries (Figure 11.2). How

Table 11.4 Road safety 'paradigms' as seen over time

Period	Characteristic
1900–1920	Crashes as chance phenomena
1920–1950	Crashes caused by the crash-prone
1940–1960	Crashes as mono-causal
1950–1980	A combination of crash causes fitting within a 'system approach'
1980–2000	The person as the weak link: more behavioural influence
2000–	– Better implementation of existing policies – Safe system approach, e.g. Sustainable Safety (Netherlands) and Vision Zero (Sweden)

Source: Inspired by OECD (1997).

can these improvements be explained, and which road safety problems still remain? This section focuses on highly motorized countries.

Over the years, there have been very many different ways of tracing crash causes and how they can best be avoided. Table 11.4 presents, by means of a few words, what the dominant thoughts in the OECD countries were in the past century (see also OECD, 1997).

In short, one can notice an increase in sophistication in thinking about road safety. The 'crash-prone theory' (1920–50) dates primarily from the phase in which the legal guilt question was the main one: which road user has broken which law and is, thus, both guilty and liable? This question was answered by the police on the registration form of a crash, finally decided inside or outside the court room, and used by insurance companies to determine how to compensate damages. From 1940 to 1960 the idea shifted to the notion that crashes could be explained using a mono-causal model. In-depth studies showed, however, that there are few mono- or single-cause crashes; accidents are usually caused by, and the result of, a combination of circumstances, which led to the so-called 'multi-causal approach' (1950–80). This approach, sometimes also called the system approach, was strongly influenced by the so-called Haddon matrix.

Haddon (1972) designed a matrix using two axes: on the one hand he distinguishes three phases in the crash process: before a crash, during a crash and after a crash. The other axis is filled with the three components of our traffic system: the road user, the road and the vehicle. Consequently, this 3×3 matrix comprises nine cells. The Haddon matrix was used to classify crash factors and to indicate that more action could be taken than just 'pre-crash

– road user related interventions', as was a tradition at the time. As Haddon tried to structure road safety (in nine cells of a matrix), other attempts were made. One came from Sweden (Rumar, 1999) in which the size of the traffic safety problem is explained as the product of three dimensions:

1. exposure (E);
2. accident risk (A/E: number of accidents per exposure);
3. injury risk (I/A: number of people killed or injured per accident).

The additional 'dimension' given by Rumar (and Nilsson, 2004, as well) was the inclusion of exposure as a variable or dimension to be used to improve road safety and to reduce the number of fatalities and injuries.

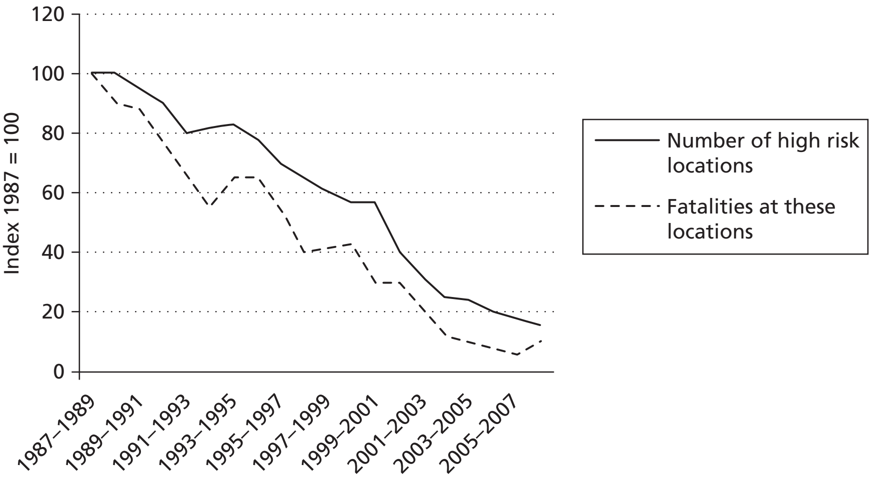
Since 2000 or thereabouts, two new main lines (paradigms) in road safety have appeared. The first one is especially aimed at evidence-based policies implemented in an efficient way. A lot of information has become available about several road safety interventions (see, for example, Elvik et al., 2009), and the idea here is not to develop new policy interventions but to improve the quality of implementing existing ones using evidence-based or research-based information on effects and costs of interventions. Greater effectiveness is considered to be a matter of scale and quality. Improving road safety in such a way that the number of casualties substantially decreases generally requires a considerable effort, given the relatively low frequency of crashes, their low densities in space and the modest effects of most safety interventions. In this first new line of reasoning, since 2000 there has been growing attention given to what is called 'safety culture' and 'cultural change' in the field of decision-making on road safety (Johnston, 2010). In this analysis, road safety progress results from an increased emphasis on strategic planning – comprising the data-driven selection of the major problems to address, the setting of objective but ambitious targets and a focus on effective implementation of programmes and measures through institutional cooperation and coordination: 'evidence-based policies' are the key words. However, despite overwhelming scientific evidence about certain themes, such as reducing speed limits to reduce speed and risks, both politicians and the public are not always convinced about introducing certain measures, even though the evidence supports this.

The second new line of thinking about traffic safety since 2000 is the Safe System Approach (OECD/ITF, 2008), such as the Sustainable Safety vision (Wegman and Aarts, 2006; see also section 11.8). The Safe System Approach recognizes that, prevention efforts notwithstanding, road users will remain fallible and crashes will occur. The approach also stresses that those involved in the design of the road transport system need to accept and share responsibility for the safety of the system and those that use the system need to accept responsibility for complying with the rules and constraints of the system. Furthermore, the Safe System Approach aligns safety management decisions with broader transport and planning decisions that meet wider economic, human and environmental goals, and the approach shapes interventions to meet a long-term goal, rather than relying on 'traditional' interventions to set the limits of any long-term targets.

The Safe System Approach paradigm shift is based on two assumptions: (1) the current traffic system is inherently dangerous, and (2) intensifying current efforts could lead to fewer casualties but not to substantially safer traffic, and the investments are less efficient than in the past and will be even more so in the future. To understand this position, it is useful to analyse the 'remaining' road safety problems in high-income countries.

In very broad terms, two types of problems can be identified in analysing road safety (Wegman, 2010): generic problems and specific problems. Specific problems are those safety problems that are concentrated on specific locations, specific road user groups, specific behaviour or specific vehicles (they relate, among other things, to the fundamental risk factors, as explained in section 11.2). Generic problems are caused by the fact that road traffic is inherently unsafe: ordinary people are killed in crashes under normal circumstances. This means that anybody can be involved in a crash at any particular time and that many people will be involved in a crash at some time in their lifetime because road traffic has not been designed with safety as an important requirement for design and operations.

In road safety policies in many highly motorized countries, for a long time the idea was to identify risk-increasing factors and reduce these specific risks. In public health too, this is a well-known and widely supported approach: cure those who are ill and identify and treat high-risk groups or circumstances. See, for example, vaccination strategies to protect 'high-risk groups' from viruses, such as the H1N1 virus (sometimes called swine flu). As a matter of fact, much of past road safety policy was based on high risks, high numbers and frequent causes, and on well-identified crash patterns. Crash and casualty rates, for example, were determined and divided into age groups, which showed that the young and the elderly had increased risks. The answer that policy-makers have come up with is the effort to reduce these high risks: smoothing the peaks in distributions. Analysis of road safety was aimed at the detection of peaks, explaining them, and finding measures to overcome them.



Source: SWOV (2010).

Figure 11.4 Number of high-risk locations and fatalities at these locations

The specific high-risk approach resulted in successful policy, certainly in the Netherlands, for example (Figure 11.4). Whereas, in the period 1987–89, 10 per cent of the seriously injured were from crashes at locations that could be labelled 'high-risk locations', this decreased to 1.8 per cent in the period 2004–06.

Therefore the least safe locations have successfully been dealt with. However, it is hardly possible for such an approach to have further positive effect in the future. One could say that the approach has become a victim of its own success and will barely make a further contribution to the reduction of the number of road crash casualties in high-income countries with a relatively long history of transport safety policies, such as the Netherlands.

The same case can be used when dealing with crash-prone drivers and for eliminating near wrecks, although the evidence is weaker. In many countries 'peaks in distributions' (hazardous locations, dangerous road users and defective vehicles) still exist and can still be eliminated. However, this approach will increasingly pose practical problems for high-income countries, such as *how* to identify and eliminate these 'relatively small peaks'.

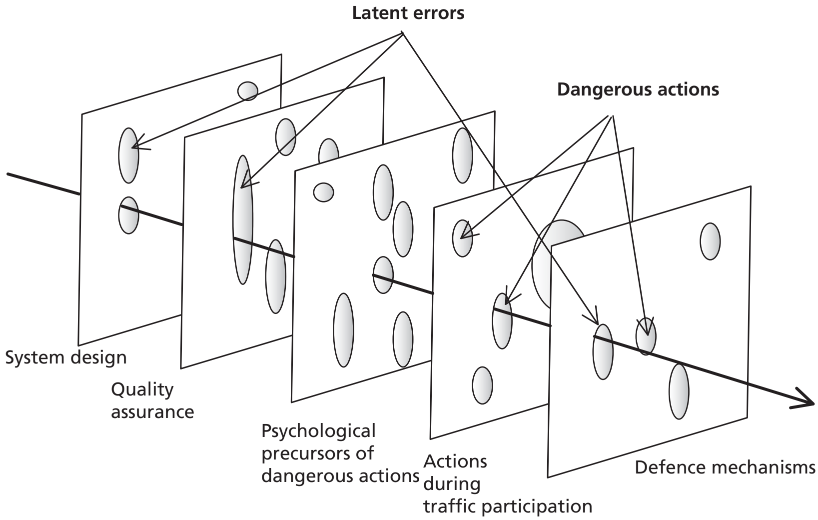
11.8 Sustainable Safety

Therefore the Sustainable Safety vision was developed in the Netherlands because the traditional policies were becoming less effective and less efficient and because the idea was that the Netherlands had not yet found out the core characteristics of its road safety problems. Although, at first glance, the vision seems to be a one-country approach, in this case, for the Netherlands, Sustainable Safety is in fact considered to be an appropriate and general vision for the future and not just for highly motorized and relatively safe countries like the Netherlands. This has been illustrated in the report about road safety by the World Health Organization (WHO) and the World Bank (Peden et al., 2004), in the report *Best Practices in Road Safety* commissioned by the European Commission (KfV, 2007) and in OECD and the International Transport Forum's publication *Towards Zero: Ambitious Road Safety Targets and the Safe System Approach* (OECD/ITF, 2008), for example. Sustainable Safety has also been discussed in the PACTS report about the future of road safety in Great Britain (Crawford, 2007).

The main lines of this vision will be explained here further. For more detail about Sustainable Safety, we refer to Koornstra et al. (1992) and Wegman and Aarts (2006). The vision aims for 'inherently safe' traffic (a concept used in rail and air traffic and also in energy production, for example). The Sustainable Safety approach starts with the idea that the present traffic system is inherently hazardous (that serious crashes can happen anywhere and at any time) and that all possible solutions are considered in an integral and rational manner. There is no a priori preference for improving roads or vehicles or changing behaviour. Furthermore, the rationale should not be restricted to road safety only, but wider deliberations are preferable (congestion, environment, scenery, economic development, health care and so on).

The following key aspects of the Sustainable Safety vision were identified:

1. Ethics:
   a. It is unfair to hand over a traffic system to the next generation with the current casualty levels.
   b. A proactive approach instead of a reactive approach.
2. An integral approach which:
   a. integrates road user, vehicle and road into one safe system;
   b. covers the whole network, all vehicles and all road users;
   c. integrates with other policy areas.
3. Man is the measure of all things:
   a. Human capacities and limitations are the guiding factors.
4. Reduction of latent errors (system gaps) in the system:
   a. In preventing a crash it is better not to be fully dependent on whether or not a road user makes a mistake or error.
5. Use criterion of preventable injuries:
   a. Which interventions are most effective and cost-effective?



Source: Based on Reason (1990).

Figure 11.5 The development of a crash (bold arrow) as a result of latent errors and unsafe actions in the different elements composing road traffic. If the arrow encounters 'resistance' at any moment, no crash will develop

As indicated in section 11.3, intentional or unintentional human errors play a role in nearly every crash. No matter how well educated and motivated people are, they commit errors and do not always abide by the rules. Studies of road traffic crashes invariably indicate that the factors 'road' and 'vehicle' play only a minor role. Present-day road traffic has not been designed with safety in mind. For avoiding crashes, road users now are almost completely dependent on the extent to which they are capable of correcting (and sometimes willing to correct) their own errors. And errors are also made in doing this. Both intentional errors and unintentional errors are made. Intentional errors are committed by the 'unwilling' road user; unintentional errors are committed by the 'incapable' road user.

Additionally, a crash is rarely caused by one single unsafe action; it is usually preceded by a whole chain of poorly attuned occurrences. This means that it is not only one or a series of unsafe road user actions that cause a crash; hiatuses in the traffic system also contribute to the fact that unsafe road user actions can in certain situations result in a crash. These hiatuses are also called latent errors (Reason, 1990) (Figure 11.5). Road crashes occur when latent errors in the traffic system and unsafe actions during traffic participation coincide in a sequence of time and place.

As unsafe actions can never entirely be prevented, the Sustainable Safety vision aims at banishing the latent errors from traffic. The road traffic system must be forgiving with respect to unsafe actions by road users, so that these unsafe actions cannot result in crashes. The sustainable character of measures mainly lies in the fact that actions during traffic participation are made less dependent on momentary and individual choices. Such choices may be less than optimal and can therefore be risk-increasing.

Adjusting the environment to the abilities and limitations of the human being is derived from cognitive ergonomics, which made its entry in the early 1980s, coming from aviation and the processing industry. In all types of transport other than road traffic, this approach has already resulted in a widespread safety culture. Further incorporation of the Sustainable Safety vision should eventually lead to road traffic that can be considered 'inherently safe' as the result of such an approach.

One option to make traffic more inherently safe is simply to ban certain road users from traffic. However, this policy can have very serious consequences for an individual. Therefore a reason for doing this must be well considered. In the current traffic system policy-makers barely make a selection at the start: everybody can participate in traffic; even more strongly, older people are encouraged to participate (independently) in traffic for as long as possible. (Temporary) prohibition of access, keeping someone away from traffic, may be justifiable from a repressive point of view, but under the present conditions it does not make a substantial contribution to a reduction in the number of road crash casualties. As it is, elimination is a rare occurrence and, in addition, the question remains whether those who have been disqualified from driving will not just continue to drive anyway. It remains to be seen how this option can be made more effective, and it seems advisable to work on innovating solutions in this area. Now, and probably also in the near future, only a relatively small proportion of the population is denied access (withdrawal of the licence), and this approach will therefore have limited effect on the further reduction in road casualties.

Another option to make traffic inherently safer is to adjust the environment to the human measure in such a way that people commit fewer errors. Here, environment not only means the physical environment (road and vehicle) but also includes the required 'software' like legislation and the traffic education that is made available. Adjustments can be made along three lines. In the first place, road designers can make potentially dangerous situations less

Table 11.5 The five Sustainable Safety principles

Sustainable Safety principle	Description
Functionality of roads	Mono-functionality of roads as either through roads, distributor roads or access roads in a hierarchical road network
Homogeneity of mass and/or speed and direction	Equality in speed, direction and mass at moderate and high speeds
Forgivingness of the environment and of road users	Injury limitation through a forgiving road environment and anticipation of road user behaviour
Predictability of road course and road user behaviour by a recognizable road design	Road environment and road user behaviour that support road user expectations through consistency and continuity in road design
State awareness by the road user	Ability to assess one's task capability to handle the driving task

frequent so that road users need to make fewer decisions and therefore can commit fewer errors. An example of this is physical direction separation on secondary roads, which prevents head-on collisions. The second possibility is to design the road user environment in such a way that fewer errors are committed and it is easier to make correct and safe decisions; this can, for instance, be done by the construction of a roundabout which makes high speeds at an intersection impossible. Thirdly, a traffic environment may be designed in such a way that *if* errors are still being committed they will not have very serious consequences for the road user. To achieve this, the road user must be presented with an environment which is forgiving of errors that are committed.

Five principles are identified as crucial for a sustainably safe traffic system (see Table 11.5). These are: functionality, homogeneity, forgivingness, predictability and state awareness.

Reduction percentages in traffic deaths in the Netherlands of more than 30 per cent and 40 per cent compared to business-as-usual levels have been estimated for policy interventions coming from or inspired by the Sustainable Safety vision (Weijermars and van Schagen, 2009; Weijermars and Wegman, 2011). Setting the societal cost of the investments alongside the societal benefits of the fatalities, injured and crashes saved shows that these interventions are socially cost-effective. The benefit–cost ratio is highly positive, around 4:1.

11.9 Conclusions

The most important conclusions from this chapter are:

1. Speed, speed and mass differences, and vulnerability are fundamental risk factors for road crashes and injuries. Pedestrians and cyclists are vulnerable road users in collisions with (high-speed) motorized vehicles.
2. Risk-increasing factors from the road users' side are impaired driving (alcohol and drugs), illnesses and ailments, emotions and aggression, fatigue and distraction.
3. Both human errors and (intentional) violations (and related extreme behaviour) are important contributory factors for road crashes.
4. Measuring road safety is not without its problems because of non-harmonized definitions, lack of data quality, data incompleteness and lack of data availability.
5. Low- and middle-income countries, especially in Asia, face a serious risk of a growing number of fatalities and injuries in the coming years. At the same time, a positive development in many highly motorized and highly developed countries can be noted. These highly motorized countries haves made considerable progress by taking all kinds of behavioural, vehicle and infrastructural related measures.
6. However, in the highly motorized countries the effectiveness of these 'traditional' policies will run out. The next step is to move from policies targeted at decreasing specific risks to policies aimed at lowering generic or inherent risks: in other words, to a systems approach.
7. One such generic approach is the so-called Sustainable Safety approach. It starts by using the idea that the present traffic system is inherently hazardous (that serious crashes can happen anywhere and at any time) and that all possible solutions should be considered in an integral and rational manner. Cost–benefit analyses show that such an approach can have high positive benefit-to-cost ratios.

REFERENCES

Aarts, L. and I. van Schagen (2006), 'Driving speeds and the risk of road crashes: a review', *Accident Analysis and Prevention*, 38, 215–224.

Alfaro, J.-L., M. Chapuis and F. Fabre (1994), *Socioeconomic Cost of Road Accidents*, Transport Research COST 313, Brussels and Luxembourg: Commission of the European Communities.

Blomberg, R.D., R.C. Peck, H. Moskowitz, M. Burns and D. Fiorentino (2005), *Crash Risk of Alcohol Involved Driving: A Case-Control Study*, Stamford, CT: Dunlap and Associates.

Connor, J., R. Norton, S. Ameratunga, E. Robinson, I. Civil, R. Dunn, J. Bailey and R. Jackson (2002), 'Driver sleepiness and risk of serious injury to car occupants: population based case control study', *British Medical Journal*, 324 (7346), 1125–1129.

Crawford, E. (2007), *Beyond 2010: A Holistic Approach to Road Safety in Great Britain*, London: Parliamentary Advisory Council for Transport Safety (PACTS).

Derriks, H. and P. Mak (2007), *Underreporting of Road Traffic Casualties*, Paris: Organisation for Economic Co-operation and Development (OECD)/IRTAD.

De Blaeij, A.T. (2003), *The Value of a Statistical Life in Road Safety: Stated Preference Methodologies and Empirical Estimates for the Netherlands*, Research Series, Amsterdam: Vrije Universiteit.

Dingus, T.A., S.G. Klauer, V.L. Neale, A. Petersen, S.E. Lee, J. Sudweeks, M.A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z.R. Doerzaph, J. Jermeland and R.R. Knipling (2006), *The 100-Car Naturalistic Driving Study, Phase II: Results of the 100-Car Field Experiment*, Washington, DC: National Highway Traffic Safety Administration, USDOT.

EC (European Commission) (2010), *Towards a European Road Safety Area: Policy Orientation on Road Safety 2011–2020*, COM (2010) 389 final, Brussels: European Commission.

Elvik, R. (1995), 'An analysis of official economic evaluations of traffic accident fatalities in 20 countries', *Accident Analysis and Prevention*, 27 (2), 237–247.Elvik, R. and T. Vaa (2004), *The Handbook of Road Safety Measures*, Amsterdam: Pergamon Press.

Elvik, R., T. Vaa, A. Hoye, A. Erke and M. Sorenson (2009), *The Handbook of Road Safety Measures*, 2nd revised edition, Bingley: Emerald Group Publishing.

Glassbrenner, D. and M. Starnes (2009), *Lives Saved Calculations for Seat Belts and Frontal Airbags*, Washington, DC: National Highway Traffic Safety Administration.

Haddon, W. (1972), 'A logical framework for categorizing highway safety phenomena and activity', *Journal of Trauma*, 12 (3), 193–207.

Johnston, I. (2010), 'Beyond "best practice" road safety thinking and systems management: a case for culture change research', *Safety Science*, 48 (9), 1175–1181.

KfV (2007), *Best Practices in Road Safety: Handbook for Measures at the Country Level*, final report of the Supreme project, Part C, Brussels: Directorate-General Energy and Transport, European Commission.

KiM (2010), *Mobiliteitsbalans 2010* [Mobility balance 2010], The Hague: Netherlands Institute of Transport Policy Analysis.

Koornstra, M.J., D. Lynam, G. Nilsson, P.C. Noordzij, H.-E. Petterson, F.C.M. Wegman and P.I.J. Wouters (2002), *SUNflower: A Comparative Study of the Development of Road Safety in Sweden, the United Kingdom, and the Netherlands*, Leidschendam, Crowthorne and Linköping: Stichting Wetenschappelijk Onderzoek Verkeersveiligheid (SWOV)/TRL/VTI.

Koornstra, M.J., M.P.M. Mathijssen, J.A.G. Mulder, R. Roszbach and F.C.M. Wegman (eds) (1992), *Naar een duurzaam veilig wegverkeer*: *nationale verkeersveiligheidsverkenning voor de jaren 1990/2010* [Towards sustainably safe road traffic: national road safety survey for 1990/2010], Leidschendam: SWOV.

McEvoy, S.O., M.R. Stevenson, A.T. McCartt, M. Woodward, C. Haworth, P. Palamara and R. Cercarelli (2005), 'Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study', *British Medical Journal*, 331, 428–430.

Mesken, J. (2006), *Determinants and Consequences of Drivers' Emotions*, Leidschendam: SWOV.

Nilsson, G. (2004), *Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety*, Lund Bulletin 221, Lund: Traffic Engineering, Lund Institute for Technology and Society.

OECD (Organisation for Economic Co-operation and Development) (1997), *Road Safety Principles and Models: Review of Descriptive, Predictive, Risk and Accident Consequence Models*, Paris: OECD.

OECD/ITF (International Transport Forum) (2008), *Towards Zero: Ambitious Road Safety Targets and the Safe System Approach*, Paris: OECD/ITF.

OECD/ITF (2011a), *IRTAD Annual Report 2010*, Paris: OECD/ITF.

OECD/ITF (2011b), accessed August 2011 at http://internationaltransportforum.org/irtad/pdf/risk.pdf.

OECD/ITF (2011c), *Reporting on Serious Road Traffic Casualties*, Paris: OECD/ITF.

Peden, M., R. Scurfield, D. Sleet, D. Mohan, A.A. Hyder, E. Jarawan and C. Mathers (2004), *World Report on Road Traffic Injury Prevention*, Geneva: World Health Organization.

Reason, J. (1990), *Human Error*, Cambridge: Cambridge University Press.

Redelmeier, D.A., R.J. Tibshirani and L. Evans (2003), 'Traffic-law enforcement and risk of death from motor-vehicle crashes: case-crossover study', *The Lancet*, 361 (9376), 2177–2182.

Regan, M.A., J.D. Lee and K.L. Young (eds) (2009), *Driver Distraction: Theory, Effects and Mitigation*, Boca Raton, FL: Taylor & Francis.

Rumar, K. (1999), *Transport Safety Visions, Targets and Strategies: Beyond 2000*, Brussels: European Transport Safety Council.

SWOV (2007), *De top bedwongen: balans van de verkeersonveiligheid in Nederland 1950–2005* [The summit conquered: assessment of road safety in the Netherlands 1950–2005], Leidschendam: SWOV.

SWOV (2010), 'The high risk location approach', fact sheet, SWOV, Leidschendam.

UNECE (2009), *Illustrated Glossary for Transport Statistics*, 4th edition, accessed at http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-10-028/EN/KS-RA-10-028-EN.PDF.

Vaa, T. (2003), 'Impairments, diseases, age and their relative risk of accident involvement: results from a meta-analysis', Deliverable R1.1 of the IMMORTAL project, University of Leeds.

Vlakveld, W. (2005), *Jonge beginnende automobilisten, hun ongevalsrisico en maatregelen om dit terug te dringen* [Young novice drivers, their risk of accident and measures to reduce this], R-2005-3, Leidschendam: SWOV.

Walsh, J.M., J.J. de Gier, A.S. Christophersen and A.G. Verstrate (2004), 'Drugs and driving', *Traffic Injury Prevention*, 5, 241–253.

Wegman, F.C.M. (2010), 'De prijs van water bij de wijn' [The price of making a compromise], inaugural lecture, January, Delft University of Technology, TU Delft.

Wegman, F.C.M. and L.T. Aarts (eds) (2006), *Advancing Sustainable Safety: National Road Safety Outlook for 2005–2020*, Leidschendam: SWOV.

Wegman, F., J. Commandeur, E. Doveh, V. Eksler, V. Gitelman, S. Hakkert, D. Lynam and S. Oppe (2008), *SUNflowerNext: Towards a Composite Road Safety Performance Index*, Leidschendam: SWOV.

Weijermars, W.A.M. and I.N.L.G. van Schagen (eds) (2009), *Tien jaar duurzaam veilig: verkeersveiligheidsbalans 1998–2007* [Ten years of sustainable safety: road safety assessment 1998–2007], R-2009-14, Leidschendam: SWOV.

Weijermars, W.A.M. and F.C.M. Wegman (2011), 'Ten years of sustainable safety in the Netherlands: an assessment'. *Transportation Research Record* (2213), 1–8.

Young, K., M. Reagan and M. Hammer (2003), *Driver Distraction: A Review of the Literature*, Melbourne: Accident Research Centre, Monash University.

Part III

Transport policy and research

12

Transport policy

Jan Anne Annema

12.1 Introduction

In Chapters 3 to 6 the different factors explaining mobility were discussed. Broadly speaking, people and companies base the decision to travel or to transport freight by weighing the private benefits and costs of the trip. A trip to a certain location is beneficial to them if it fulfils their needs. The private costs of a trip consist of monetary costs, travel time and effort. This chapter shifts the range of view from the private to the government perspective. Governments also weigh the costs and benefits of mobility but they take a societal perspective. If a government concludes that the transport market from a societal perspective results in undesired outcomes (e.g. congestion, air pollution, traffic casualties) they often interfere with transport policy.

The aim of this chapter is twofold. Firstly, it is to explain the main reasons why governments intervene in the transport market. Secondly, a concise overview is given of dominant transport policies. Transport policy covers many topics, much more than can be discussed in this chapter. For more detail and far more in-depth analysis of transport planning issues, we refer the reader to Banister (2002).

Transport is one of the most regulated sectors in any economy (Button and Gillingwater, 1986). Governments provide and own transport infrastructure, tax car owners, subsidize public transport, decide who has right of way by implementing traffic rules, implement emission standards to make vehicles cleaner and safer, and so forth. There are many reasons why governments intervene with these policies in the transport market. Different political parties and different political systems all over the world will have different considerations. However, generally speaking, all over the world three main reasons can be found for government interference in the transport market:

1. market failure;
2. equity reasons;
3. generating revenues.

Market failure is an economic concept meaning that the market itself will not result in optimal outcomes from a societal perspective. In economic jargon, the market does not result in efficient allocation of resources. The 'transport market' stands for the interaction between, on the one hand, suppliers of transport services such as the infrastructure providers (in most cases governments), public transport companies and vehicle manufacturers and, on the other hand, people and shippers, who demand transport services. Resources are any scarce goods that are needed to satisfy people's needs or wants. Transport related examples of scarce goods are road capacity, cars, aircraft, clean air, silent living areas, land and so forth. If a market does not result in efficient allocation of resources, it means that the market cannot sustain 'desirable' outcomes or stop 'undesirable' outcomes. For example, if car drivers when deciding to make a trip (a desirable outcome for them) do not take into account the fact that they pollute the air (a scarce good), for people living close to the road the result can be poor air quality (an undesirable result). This is a classic example of a transport market failure related to the existence of so-called transport external effects. In section 12.2 we will explain the concept of external effects in more detail. An important policy aim for many governments is to maximize efficiency in the transport market. This will be explained in section 12.3.

Equity reasons can also explain why governments interfere in the transport market. The policy aim with these kinds of policies is to distribute the costs and benefits of transport in a fair way. For example, governments subsidize public transport because they think it is desirable that all people in a country have nearby public transport at their disposal, even in very low density population areas where running a profitable public transport service is impossible. Section 12.4 focuses on 'equity issues' related to transport policy-making.

Finally, governments use the transport market as a source for generating revenue. The most notable example is the transport fuel levy which in Western European countries amounts to 50 to 70 per cent of the total fuel price.

Section 12.5 explains that, as well as improving efficiency and equity, politicians' self-interest is also a reason for implementing transport policies. Section 12.6 discusses criteria that can be used to define healthy transport policies. Section 12.7 gives some examples of transport policy-making. Conclusions can be found in section 12.8.

12.2 External effects of transport

An important reason for government interference is transport market failure in the form of external effects. External effects of transport are effects which people and shippers do not take into account when deciding to make a trip or to transport freight. External effects can be negative (costs) as well as positive (benefits). External costs or benefits accrue by definition to a third party. A third party can be a non-traffic participant, such as someone living close to a road who is exposed to poor air quality. Other traffic participants can also be third parties, such as cyclists, who may be at risk of an accident because a car driver has decided to drive on the same road. Most external costs are related to the environment, safety and accessibility. Chapters 9–11 have introduced the reader to these areas, and Chapter 8 has discussed the importance of technology for these areas. This chapter introduces the reader to the area of external costs from a policy perspective.

External costs

Negative external effects are called external costs. Verhoef (1996) distinguishes three kinds of external transport costs:

1. costs due to the use of transport means such as road vehicles, ships or aeroplanes;
2. costs due to transport means, ownership and availability;
3. costs due to infrastructure.

We will now discuss these three kinds of external costs in more detail.

Costs due to the use of transport means such as road vehicles, ships or aeroplanes

The use of cars, lorries, trains, aeroplanes and ships results in external costs. Table 12.1 gives an overview of external cost items by transport mode for the US (Delucchi and McCubbin, 2010): congestion delay, accidents, air pollution, climate change, noise, water pollution and energy security. As external costs are by definition not included in market transactions, it may be a surprise that the external costs in Table 12.1 are expressed in monetary units. However, different methods have been developed to estimate people's willingness to pay (WTP) to avoid external costs (or people's willingness to accept (WTA) a monetary reward to accept external costs). For more details on this issue, see Chapter 14. These WTPs and WTAs are used in Table 12.1 to estimate the external costs using a monetary value.

Table 12.1 Summary of estimates of external costs by transport mode and cost category

	Passenger (per passenger-mile)				Freight (per ton-mile)			
	Road	Rail	Air	Water	Road	Rail	Air	Water
Congestion delay	0.88–7.5	n.e.	0.35	n.e.	0.54	0.03	n.e.	n.e.
Accident	1.4–14.4	n.e.	n.e.	n.e.	0.11–2.0	0.22	n.e.	n.e.
Air pollution, health	0.09–6.7	0.49	0.01–0.39	1.1	0.10–18.7	0.01–0.35	0.0–1.9	0.08–1.7
Air pollution, other	n.e.	n.e.	n.e.	n.e.	n.e.	n.e.	n.e.	n.e.
Climate change	0.06–4.8	0.02–1.7	0.08	0.16	0.02–5.9	0.01–0.47	0.45	0.00–0.23
Noise	0.0–3.5	0.52–0.89	0.88	n.e.	0.0–5.3	0.05	n.e.	n.e.
Water pollution	0.01–0.05	n.e.	n.e.	n.e.	0.003–0.05	n.e.	n.e.	n.e.
Energy security	0.20–0.84	0.15–0.58	0.18–0.69	n.e.	0.22–0.84	0.02–0.07	n.e.	0.03–0.11

Notes:

n.e. = not estimated.

Year 2006; costs in US cents.

Source: Delucchi and McCubbin (2010).

An interesting item in Table 12.1 is ‘energy security’. This item reflects the heavy use of imported oil by the US transportation sector, which gives rise to several kinds of economic costs that are not reflected in the price of oil. Delucchi and McCubbin (2010) point out the cost of the so-called Strategic Petroleum Reserve in the US, defence expenditures to protect US oil interests, macro-economic disruption and adjustment costs due to price volatility, and pure wealth transfers from US consumers to foreign producers. Not included in Table 12.1 are external costs related to road traffic killing pets and wild animals. Glista et al. (2009) cite research that estimated vertebrate mortality on roads in the United States at 1 million individual animals per day.

Table 12.1 shows that road traffic casualties and injuries are an important external cost item (see also Chapter 11). Delucchi and McCubbin (2010) did not include avoidance costs in their external cost estimate for accidents. Avoidance costs mean that the threat of motor vehicle accidents can make people afraid to walk or cycle. Not including these avoidance costs could be an important omission. For example, Adams (1999) showed that in 1971 around 80 per cent of British children went to school by themselves. In 1990 this share had decreased to only 9 per cent. The most important reason for this was that parents were said to be afraid of traffic accidents. The costs of avoidance relate to freedom loss for children, in terms of children losing the opportunity to develop partly by themselves in a space without parents supervising them and, probably, losing the opportunity to have some ‘easy’ daily physical exercise.

Costs due to transport means, ownership and availability

The second external costs category distinguished by Verhoef (1996) is related to the non-use phase of vehicles. For example, vehicles sometimes use space for parking in the public domain for which the vehicle owner does not have to pay or the payment is not sufficient to cover all costs. Environmental effects related to producing and scrapping vehicles, aircraft or ships are also part of this external cost category.

Costs due to infrastructure

Finally, external costs arise according to Verhoef (1996) because of the construction and existence of transport infrastructure. In a review Geurs et al. (2009) point out that the mere presence of transport infrastructure (roads, railway lines, waterways, etc.) may affect the quality of the physical environment. This applies to visual quality, light pollution and people’s perception of the environment or neighbourhood, aesthetics and quality of life, for example. Furthermore, Geurs et al. (2009) cite research which showed that new or existing transport schemes, such as roads or railways, can have detrimental social impacts on communities (severance). Transport infrastructure can also act as both physical and biological barriers to many wildlife species (Glista et al., 2009). Roads can affect the quality and quantity of available wildlife habitat, most notably through fragmentation. Furthermore, producing asphalt, making or scrapping roads and/or road maintenance can result in environmental impacts or direct negative health impacts to the construction workers which are not sufficiently reflected in their wages. These environmental and health related impacts are also negative external costs of infrastructure.

External benefits

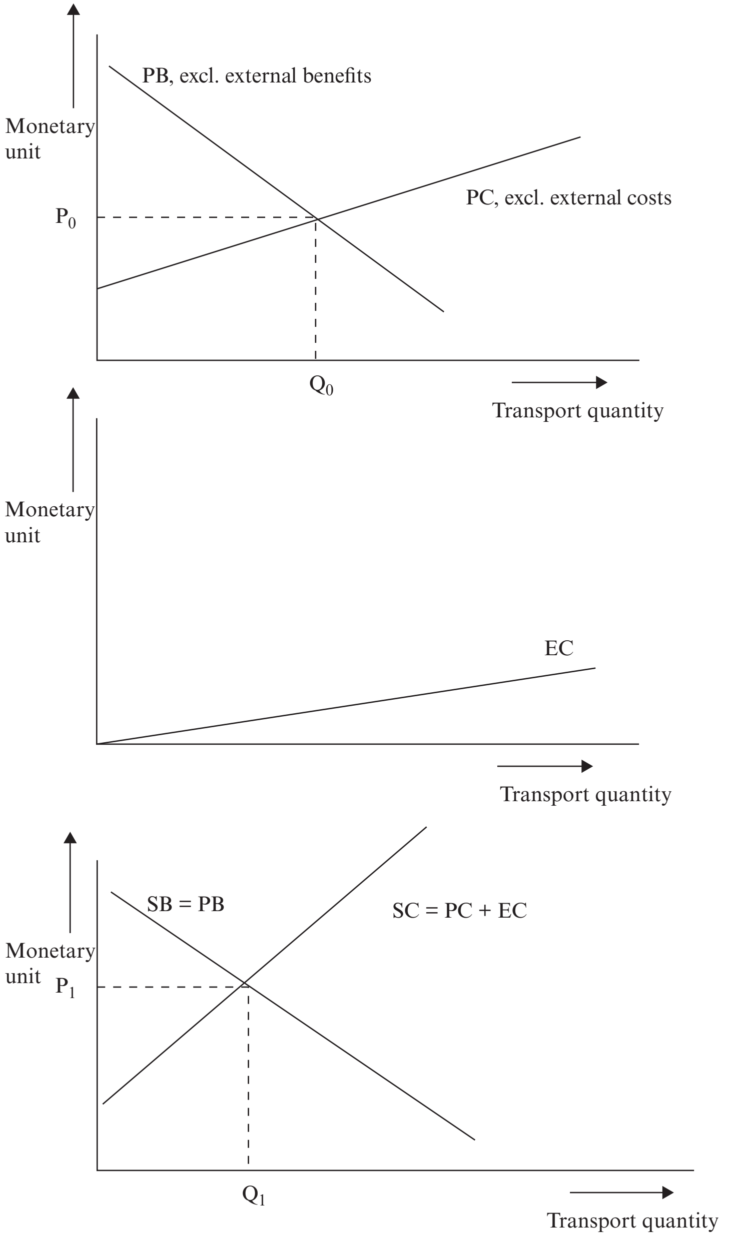
External benefits are the positive external effects of mobility. The existence of external benefits can be a reason for governments to intervene, just as in the case of external costs, for these benefits work by definition outside the 'normal' market: people or shippers who participate in a transport activity do not take these positive impacts into account when they decide to make a trip. This could, as opposed to negative external effects, result in an amount of transport that is too low from a societal perspective. External transport benefits have been discussed for a long time. In most cases there was confusion in these discussions about the concept. For example, some people claim that freight transport results in 'external' impacts such as lower production costs, low consumer prices and a broad product variety. However, these benefits are all monetized impacts in the freight transport market; producers and shippers take all of these impacts into account when deciding to transport freight. Others sometimes claim that mobility leads to external benefits such as enhancing family relations or creating a smaller world, as thanks to cheap aviation people can get acquainted with far-away cultures in a relatively easy way. Nevertheless, these claims are also intended transport impacts when people decide to visit their family or spend their holidays in a far-away country. Thus external benefits of transport are limited to the well-known classic examples such as plane or train spotting. In the remainder of this chapter external transport benefits will not be discussed any further.

12.3 Maximizing welfare

As stated in section 12.1, maximizing welfare (improving efficiency) is an important transport policy aim. The aim is based on the so-called economic welfare theory (a relatively easy introduction to welfare economics can be found in Johansson, 1991). The economic welfare theory states that we should strive in policy-making for Pareto optimal welfare, or Pareto efficiency. (The Italian economist Vilfredo Pareto is one of the founders of this welfare theory.) According to many standard economics textbooks, Pareto efficiency is said to exist when no other improvements can be made in the allocation of resources to one individual without it causing a loss to others. A way of explaining Pareto efficiency in the transport market is to think of a policy (e.g. lowering existing fuel levies) that would result in more transport. Assume that noise nuisance is the only transport externality. The people who can travel more car kilometres because of this policy are better off. However, the extra car kilometres will result in an increase in the road noise load. Perhaps lowering the fuel levy slightly would only result in a small increase in traffic and thus only a small amount of extra noise load very close to the road where nobody is affected by it. By increasingly lowering the fuel levy more traffic arises and at a certain point the first people living close to the road will hear the increased noise load and will perhaps be bothered by it. If this is the case, these people are worse off at that point. Thus lowering fuel levies is a Pareto efficient policy: just before the point that the first people are worse off because of this policy.

The Pareto optimum is a strict criterion. Later in history a more applicable criterion was formulated by Nicholas Kaldor and John Hicks (Hicks, 1939; Kaldor, 1993). They stated what is now called the Kaldor–Hicks efficiency criterion (also called the compensation criterion), which means that an outcome of a policy is efficient if those who are made better off could in theory compensate those who are made worse off. The compensation criterion is the most used criterion in cost–benefit analysis (CBA) (Chapter 14 and below). Related to the previous example, the Kaldor–Hicks optimum of the 'decrease in fuel levy policy' is not reached as long as the benefits of the people who can drive outweigh the costs of the people who suffer from the extra noise load. It is important to note that in Kaldor–Hicks the actual payment of compensation is not required; it is enough that the possibility for compensation exists. Thus, using the Kaldor–Hicks criterion in practice, a more efficient outcome in the case of the decrease in fuel levy policy can in fact leave some people worse off, namely the people who will suffer from the increased noise load and who are only compensated in theory. In contrast, using Pareto efficiency in practice, nobody can become worse off.

Welfare theory can also be explained using Figure 12.1 (Schmidtchen et al., 2009). The figure shows the so-called marginal costs and benefits. 'Marginal' means per extra unit of transport, such as one extra passenger kilometre or one extra ton freight kilometre. The top figure reflects the situation on the



Notes: PB = private benefits; PC = private costs; EC = external costs; SB = social benefits; SC = social costs; Q = quantity; P = price.

Source: Schmidtchen et al. (2009).

Figure 12.1 Transport prices and quantities in two equilibrium situations: without external benefits and costs (top) and with external benefits and costs (bottom). The middle picture shows the marginal external costs transport market where external costs (EC, see middle picture) are not taken into account. In the top figure the marginal private benefits (PB) of transport are sorted in such a way that they decrease according to the amount of transport. The explanation is that people value the kilometres they travel differently. There are highly beneficial trips for them, for example the trips to their work where they can earn money. For these trips they are willing to pay relatively large amounts of money, travel time and/or effort. People also make less important trips for which they are not willing to pay much money, travel time and/or effort. It seems obvious that people prioritize their kilometres from the most beneficial ones to the less beneficial ones. At the top of Figure 12.1 the trips are sorted from highly to less beneficial ones. For the marginal private costs (PC) it is the other way around (top of Figure 12.1): it is assumed in the figure that the marginal private costs increase according to transport quantity. To be clear, the figure is schematic. In reality it is highly probable that the private cost line will stay constant for a long while: the marginal private costs for the first kilometres are more or less the same as for the 20 000th kilometre travelled. However, for the sake of clarity an increasing marginal private cost line is assumed, because then it can be shown more clearly to the reader that there is an equilibrium at quantity Q_0 with price P_0. Of course, there would also be an equilibrium at the intersection, assuming a constant private marginal cost line, but the picture would be messier. No matter what, at Q_0 with price P_0 the private optimum is reached. At the equilibrium the marginal private cost of the last kilometre travelled equals exactly the marginal private benefit of that kilometre. Adding one more kilometre would still result in marginal private benefits but, at the same time, the marginal private cost of that one more kilometre is higher. Thus, from a private perspective, it is not rational to drive that extra kilometre.

As can be derived from the definition, private parties do not take external costs into account when deciding to make a trip. At the bottom of Figure 12.1 is depicted what would happen if marginal external costs were taken into account in the transport price. This 'taking into account' is also called internalization of external costs. The marginal social cost line becomes steeper when external costs are internalized compared to the marginal private cost line (SC = PC + EC). The reason is that increasing quantities of transport often result in higher marginal external costs (see middle picture). For example, the first car or lorry kilometres rarely result in traffic jams. However, above a certain point, traffic jams grow more or less exponentially in proportion as traffic quantities increase without adding new road capacity. Thus, as a result of the steeper SC line, a new equilibrium arises at price P_1 and transport quantity Q_1 (bottom of Figure 12.1): the so-called social optimum. Figure 12.1 shows how to internalize from a theoretical point of

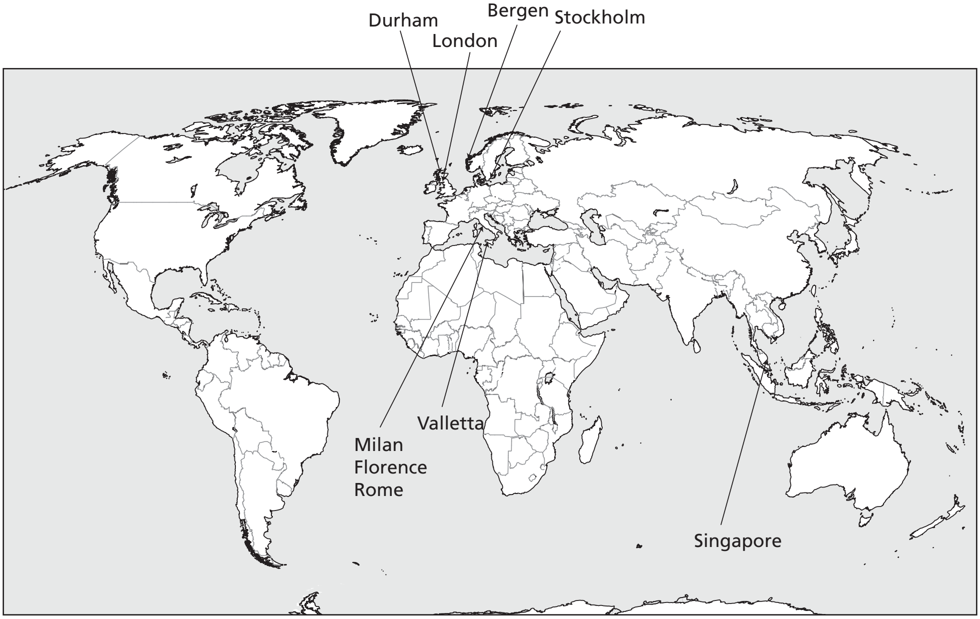


Figure 12.2 Implemented urban congestion charging

view: governments should increase the private transport price (P_0) through a charge equal to P_1 minus P_0. In the process, the social optimum in the transport market will arise, resulting in less transport (Q_1) compared to the situation where external costs are not internalized (Q_0). In other words, in the world without internalizing external costs there is too much traffic (Q_0) from a societal point of view.

Governments carry out transport pricing policies on the basis of welfare theory. For example, as stated in a recent European Union policy document (EC, 2009: 11): 'in transport, like in any other sector, there cannot be economic efficiency unless the prices reflect all costs – internal and external – actually caused by the users'. Another example, based on the notion of external costs all over the world, is that of the implementation or study of urban charging schemes (Figure 12.2). The congestion schemes have different motivations, but one of them is to internalize external costs of road traffic.

As already mentioned, CBA is based on the welfare theory. CBA is applied to evaluate many transport policies (e.g. Deb and Filippini, 2011, on welfare changes from efficient pricing in public transit in India), but it is mostly used as an appraisal tool for new transport infrastructure all over the world. Mackie and Kelly (2007) report systematic CBA use in countries such as Australia, Canada, the Czech Republic, Denmark, France, Germany, Hungary, Japan, Mexico, the Netherlands, New Zealand, Norway, Portugal, South Africa, Sweden, Switzerland, the UK and the USA. A CBA tries to quantify all marginal impacts of a new infrastructure as much as possible and puts a monetary value to these impacts. Impacts are, for example, construction costs, decreased travel times, lower air quality and so forth. In most cases travel time gains are the most important marginal benefit and construction costs the most important marginal costs. The idea is that, if the marginal benefits of new infrastructure outweigh the marginal costs, politicians could decide to build the new infrastructure, as by doing so they increase total welfare in a country or region. For more details on CBA, see Chapter 14.

The practice

The welfare theory may seem elegant and rather straightforward. However, implementing the theory in practice is not particularly easy. One important reason is that it is complicated to monetize external costs. To determine the charge level (P_1 minus P_0) or to carry out a proper CBA, researchers need to know the price people or shippers are willing to pay to avoid one unit of traffic jam, noise nuisance, traffic accident, air pollution and so forth. The problem here is that these impacts are external and thus, by definition, outside a market of supply and demand where prices are determined. Still, based on different kinds of valuation techniques scientists are capable of monetizing external effects. Table 12.1 shows US external cost estimates in US cents per passenger mile and ton mile. Maibach et al. (2008) carried out a large-scale overview of valuation techniques and practical results for transport external costs. Nevertheless, the valuation results are uncertain and sometimes highly debated (see Aldred, 2002, 2006, for example). The valuation of one extra or less ton of carbon dioxide (CO_2) emission or the valuation of damage to nature or landscapes due to new infrastructure is especially often highly controversial (see Niemeyer and Spash, 2001, for a critique on nature valuation).

12.4 Equity

As well as policies aimed at improving efficiency, governments all over the world implement transport policies because they consider them to be fair. We give three main categories of equity considerations in transport policy-making.

Firstly, governments can find it unjust that transport causes air pollution or traffic deaths and therefore implement strict policies to decrease these negative impacts. These could be policies of which the marginal social benefits do not outweigh the marginal social costs (see section 12.3). However, governments may deem damage caused by transport at certain places to be so wrong or unfair that they decide to implement those policies anyhow. Society as a whole loses welfare owing to the policy, but some people to whom important social transport costs (e.g. negative health impacts) accrue become better off.

Secondly, governments can find it important that people in their country or region have a certain amount of public transport or road network quality at their disposal, even if the costs of building and maintaining a certain transport quality do not outweigh the benefits. Governments therefore redistribute the total welfare in a country or region in a way that they judge as being fair. An example of such a policy is public transport subsidies. Parry and Small (2009) estimate that, across the 20 largest transit systems in the United States (ranked by passenger mile), the subsidy, as measured by the difference between operating costs and passenger fare revenues, ranges from 29 to 89 per cent of operating cost for rail and from 57 to 89 per cent for buses. In the European Union total public transport fare reduction subsidies in 2005 amounted to 33 billion euros (at 2005 prices; EEA, 2011). There are various motivations for such a policy (Button, 1993; van Goeverden et al., 2006), but one of them is the so-called 'social function' of public transport, meaning that vulnerable groups such as low income households, people without a driving licence, the elderly and people with a handicap need public transport to avoid problems of social exclusion (e.g. Lucas et al., 2001).

Thirdly, governments can find it fair that certain transport mode users (e.g. all car users) pay for their total social costs, even if this means that, in this way of pricing, marginal social costs outweigh the marginal social benefits. To explain the notion of total social costs we have made a balance sheet of the total costs and benefits of a transport mode (see Table 12.2 with car driving as an example). We can ignore the internal costs and benefits because these items do not belong to social costs and benefits (they are private costs and benefits). Thus the social balance sheet has the external costs caused by car driving and government expenditures for roads and police on one side and, on the other side, the tax revenues paid by the car drivers. So car drivers pay for their total social costs if the government revenue from fixed and variable car taxes equals the external car costs and government expenditures on car driving.

Governments have the choice to charge transport modes their marginal or total social costs dependent on their policy goal. They can aim for an effi-

Table 12.2 Balance sheet of the costs and benefits of car driving – an example

Car driving benefits	Car driving costs
Internal benefits: the car drivers have benefits because they can reach the desired destination; some car drivers will have benefits because they can satisfy certain social-psychological needs (e.g. car driving results in pure joy or status enhancement).	Internal costs: travel time, effort, money costs related to fuel, car purchase, taxes (see Chapter 6).
External benefits: negligible.	External costs: accident risks, environmental damage, traffic jams and so forth (see Table 12.1).
Benefits for the government: fixed and variable tax revenues.	Costs for the government: road construction and maintenance, traffic police and so forth.

cient or fair transportation system. In the scientific literature the tension in this policy choice is sometimes denoted as efficiency versus equity (see Verhoef, 1994, for example). To explain this tension a bit more we give an example of rail infrastructure construction. The social costs in this example are construction costs and external costs such as damage to landscapes and nature. The construction and damage costs to landscapes and nature are not directly related to use. If there are 25 000 or 50 000 train movements per year between locations A and B, the whole track between A and B has to be built because it makes no sense to build only half of the track. When governments aim at equity they will charge the rail track users a price based on covering total social costs. Thus a price per kilometre has to be paid including construction and landscape or nature costs which are not directly related to the amount of users. However, with this charge a price level per kilometre travelled may arise that is higher compared to the price where marginal social costs equal marginal social benefits: point P_1 in Figure 12.1. Consequently, too few rail kilometres will be travelled from a social optimum point of view. In conclusion, the tension is that it may seem fair to confront road or rail users with their total social costs but this may result in an inefficient transport system. On the other hand, if these users are only charged according to their marginal social costs governments cannot cover all their social costs.

As mentioned above, CBA is a much used appraisal method for transport policies. However, the CBA is often criticized for ignoring equity issues such as distribution effects (e.g. Rietveld et al., 2007; Thomopoulos et al., 2009; see also Chapter 14). Another equity related criticism is that rich people matter more than poor people in CBA because, for example, rich people are, ceteris paribus, willing to pay more for travel time gains and environmental benefits compared to poor people. It is important to realize that CBA does not exclude reporting distribution effects, for example over income classes or regions. These criticisms mainly relate to the use of CBA in practice. However, there is also a more fundamental criticism on CBA related to equity or justice. Some point out that it would not be sensible to use highest total utility as the only yardstick in decision-making. In their view it is not wise to disregard the presence of tragic choices in politics (Nussbaum, 2000), as when cost–benefit analysis leads to a choice of course A (many winners) over course B (hardly any winners), but course A leads to uncompensated losers (a potentially small group whose members may suffer from losing their homes, serious illnesses and even death, for example). For more discussion on the limitations of CBA from an ethical perspective, see van Wee (2012). To overcome problems of tragic choices in CBA, appraisal methods such as multi-criteria analysis (MCA), environmental impact assessment and social impact assessment are used. Sometimes these evaluation studies are used in addition to CBA outcomes; sometimes only these appraisal methods are applied. See Chapter 14 for more information on MCA. At the time of writing this book it is interesting to note that the so-called social justice literature has also been developed further in relation to transport (Beyazit, 2011). Based on the works of Sen and Nussbaum (Sen, 2009, is a recent comprehensive source on justice), the capability approach (CA) is suggested as a means to assess the social justice of transport policies. The CA considers notions such as quality of life, opportunities, functioning, alternatives and freedoms of individuals as being important criteria for sound policy-making.

12.5 Public choice theory

Sections 12.3 and 12.4 describe rational considerations or aims which politicians can apply when deciding whether or not to implement a new transport policy. There are others reasons – perhaps more irrational – which explain the actual behaviour of politicians and bureaucrats in the practice of transport policy-making. Here, the core idea is that psychological reasons and politicians' self-interest also explain the implementation of transport policies. The idea of self-interest is rooted in the so-called 'public choice theory' (Buchanan and Tullock, 1962). In this economic theory it is assumed that people acting in the political marketplace act in the same way as in other markets: they are mainly concerned for themselves (and their nearest relatives).

Relatively speaking, there is a lot of scientific literature that points to psychology and 'self-interest' as explanations for the question as to why sometimes new infrastructure that does not meet the expected efficiency and/or equity considerations is actually built. Flyvbjerg et al. (2003a) have built a database of 258 transport infrastructure projects all over the world. Of these projects 86 per cent showed cost overruns, with an average overrun of 28 per cent of the estimated costs. 'Cost overruns' mean that the costs of building the new infrastructure are higher compared to the cost estimation used for making the political decision. Flyvbjerg and colleagues (Flyvbjerg, 2007; Flyvbjerg et al., 2003b, 2006) also investigated 210 projects on demand shortfalls, comparing transport demand the first year after introduction of the new infrastructure with the ex ante estimate. They found large inaccuracies, mainly for rail projects. The literature distinguishes four different types of explanations for these problems occurring: technical, psychological, economic and political explanations (Flyvbjerg, 2005; Cantarelli et al., 2010). Technical explanations explain failure in terms of honest mistakes related to difficulties in predicting the future (Flyvbjerg, 2005). Nevertheless, if only technical reasons explain the mistakes it seems improbable that mainly cost underestimations and benefit overestimations (for rail) would occur. Therefore two additional reasons are proposed. Firstly, it seems probable that politicians and bureaucrats are unintentionally too optimistic about their projects. Psychological explanations state that humans tend to overemphasize their own abilities and to be overly optimistic about the future, rather than rationally weighing gains, losses and probabilities (Lovallo and Kahneman, 2003). Finally, and here self-interest comes into play, it is probable that decision-makers deliberately present wrong numbers (Flyvbjerg et al., 2003a; Flyvbjerg, 2005; Cantarelli et al., 2010). Project funds are scarce and projects that look good on paper can more easily be financed than projects that do not. Politicians, planners and forecasters are said deliberately to underestimate costs while overestimating benefits in order to gain approval and funding for their 'own' (sometimes much loved) project.

12.6 'Healthy' transport policies

An important question is: how is a 'healthy' transport policy to be achieved (van Wee, 2009)? A lot of transport research and analysis is aimed at helping decision-makers to answer this question. Researchers and consultants help by developing and improving tools such as CBA and MCA (Chapter 14), by carrying out future studies (Chapter 13) and all kinds of effectiveness studies. Van Wee (2009) distinguishes six general criteria for policy interventions which should be taken into account in the decision-making process in order to achieve healthy policies: (1) effectiveness, (2) efficiency, (3) equity, (4) ease of implementation, (5) flexibility and (6) long-term robustness:

1. Effectiveness relates to the question: does the policy do what it is supposed to do? For example, if free public transport is implemented for environmental reasons, the question is: will it lead to less environmental pressure? Van Wee notes that not only the question of whether the policy is effective at all is important but also the level of effects.
2. Efficiency can be expressed in terms of cost-effectiveness or cost-to-benefit ratio. Cost-effectiveness is generally a relatively easy indicator in the case of 'simple' policy options, having one dominant effect and only monetary costs. It can, for example, be applied to helping the political choice between subsidizing technology A or technology B, which only differ in technical costs and CO_2 emission reduction potential. Of course, it would be 'healthy' to subsidize the technology with the lowest costs in achieving one kilogram of emission reduction. It is less simple to use cost-effectiveness as an efficiency indicator if a policy option has (1) multiple effects or (2) monetary as well as non-monetary costs. An example of multiple effects is improvements in public transport contributing to accessibility, safety and the environment. An example of non-monetary effects is reductions in speeds on motorways resulting in lower emissions. Additionally, they result in reduced fuel use and costs, which can be expressed in monetary terms, but they also increase travel times and might reduce the fun of driving for some, both being non-monetary costs. For such less simple policy options a (simple) CBA is preferable to estimate efficiency.
3. Equity relates to the questions: are there winners and losers because of the policy instrument and who are the winners and losers? It seems obvious that policy instruments with hardly any equity issues are relatively much easier to implement compared to instruments with a lot of equity issues. For example, road pricing is often criticized for not being fair: those with a high value of time, generally people with high incomes, benefit (Gunn, 2001); those with a low value of time, that is, people with low incomes, lose.
4. The fourth criterion to be analysed is the ease of implementation. It is an advantage if a policy option is easy to implement. But van Wee (2009) stresses that this criterion is not added to suggest that only easy-to-implement policy options are 'healthy'. A policy option should be considered as an important candidate option particularly if it could have major effects and is cost-effective. For example, some road pricing designs could belong to this category. However, it is worth trying to understand the major barriers to implementation and to learn from successful implementations elsewhere. For example, the equity barrier (see above) can be reduced or even solved by carefully selecting options for revenue use. It could be an option to reduce income tax for people with

low incomes or to reduce tax on fuel efficient cars, which tend to be owned by people with low incomes.

5. Flexibility relates to the ease of adapting the policy because of the ease or difficulty in foreseeing changes. For example, once introduced, levies on fuels and cars and emission regulations can be changed relatively easily. In Chapter 13 (exploring the future) the importance of the flexibility criterion will be discussed in more detail.
6. Long-term robustness, the final criterion, relates to the question of whether a policy is 'no-regret' under uncertain long-term developments that could have a major impact on society. This criterion is strongly related to flexibility (see further Chapter 13). Here, the term 'flexibility' is used for any foreseen or unforeseen changes, and also short-term changes and changes with relatively low impacts. Long-term robustness relates to major changes. A check on robustness is recommended by van Wee (2009) at least in cases of expensive land-use or transport infrastructure policies. Are these policies no-regret when there are major changes?

12.7 Current transport policy

The previous chapters contained many topics for which government policies are developed. Some transport policies aim at decreasing transport resistance factors, other policies try to influence the needs and location of activities or try to improve the environmental performance of vehicles, and so forth. It is impossible to give a complete overview of all policies at all levels. Therefore this section will only outline some of the main policies. Extensive overviews of economic policies in road transport and of policy instruments for sustainable road transport can be found in Santos et al. (2010a, 2010b).

There are different ways to classify transport policies:

1. According to policy goal, for example policies to improve accessibility or policies to improve transport safety.
2. According to the kind of instrument, for example pricing policies or policies providing new infrastructure.
3. According to the policy body responsible for implementing the policy. These bodies are: national governments, regional governments (federal states or provinces), city councils, supranational bodies (such as the International Civil Aviation Organization (ICAO) and the International Maritime Organization (IMO) of the United Nations) and international economic and political blocs (e.g. the European Union, EU).

In Table 12.3 we have chosen to give a transport policy overview according to goal and dominant instruments. In the column remarks, we give some examples of bodies responsible for the policy. Note that this table is not complete.

Accessibility

In government expenditure terms the most important policy is providing and maintaining transport infrastructure. These policies are often especially aimed at improving the accessibility of a region and to a lesser extent to improve equity. In 2008 investment in inland transport infrastructure as a percentage of GDP amounted to 0.6 per cent for the US and 0.8 per cent for Western European countries (International Transport Forum, 2011), which means that these regions together spent roughly USD 250 billion on inland transport infrastructure in 2008. In most countries, governments provide and maintain transport infrastructure. However, since the 1990s so-called public–private partnerships (PPPs) have become increasingly popular. Australia is seen as a forerunner, where in 1985 discussions had already begun with private partners to build the Sydney Harbour Tunnel (Czerwinski and Geddes, 2010). These authors also give a formal definition (Czerwinski and Geddes, 2010: 5): 'PPP has become a catch-all term for a range of contractual agreements between a public sector project sponsor and private sector partners who are able to provide the design and construction of a new transportation facility, or the operation, renovation, and expansion of an existing facility.' One argument for PPP is that this approach results in better projects compared to projects which are solely provided by public parties. Neo-classical economic theory states that private companies are more efficient and innovative than public companies because they act more rationally and they strive for maximum profit. 'Act more rationally' may sound strange, but when the idea is set against public choice theory (section 12.5) one can imagine that for private parties it is impossible systematically to underestimate the costs and overestimate the benefits of their investments because by doing so they would not be in business for very long. For a broad overview on all kinds of PPP related issues we refer to the special issue on PPP in *Research in Transportation Economics* (2010).

Environment and accidents

For policy goals related to improving the environment and transport safety, the dominant instruments are standards and regulations which try to improve the technical characteristics of vehicles and fuels (Table 12.3). For example, vehicle emission and fuel standards aim to make new vehicles, respectively,

Table 12.3 Much used transport policy instruments

Goal	Important instruments	Remarks
Improving accessibility	– Providing and maintaining road, port and rail infrastructure. – Subsidizing public transport fares. – Road pricing. – Providing traffic management. – Implementing land-use policies.	These policies are mostly implemented by national governments. Road pricing is the exception, as almost all current road pricing schemes are implemented in specific cities.
Improving the environment and liveability	– Setting vehicle emission and noise standards. – Setting fuel standards. – Providing noise barriers along railways and roads, providing eco-tunnels and tunnels in urban areas.	National governments implement standards in most cases. In Europe the EU sets standards, which are transferred to national laws accordingly. For aviation and shipping the ICAO and IMO sign international agreements to make aircraft less noisy and aircraft and ships less polluting and more fuel efficient, amongst other things. Adapting infrastructure is national and/or local/regional policy.
Improving transport safety	– Setting safety standards for vehicles. – Implementing rules such as making the wearing of seatbelts and crash helmets mandatory. – Adapting infrastructure to make traffic situations safer (e.g. constructing roundabouts in place of junctions).	National governments implement standards in most cases. Adapting infrastructure is national and/or regional/local policy.
Improving equity	– Subsidizing public transport fares. – Providing railways, ports and roads in poor and/or less densely populated areas.	National, regional and local governments subsidize public transport fares. National governments, but also international economic blocs (such as the EU), sometimes decide to subsidize new infrastructure being built in poorer or slow-developing regions or countries (the aim is often to improve accessibility as well as social inclusion). Additionally, the World Bank[a)] offers low-interest loans to poorer countries, for

Table 12.3 (continued)

Goal	Important instruments	Remarks
		example in order to make it possible for them to build new transport infrastructure. One of the aims of the World Bank is to fight poverty.
Generating government revenues	– Implementing taxes on vehicles and fuels.	Mostly national and regional governments tax transport vehicles and fuels.

Note: [a] The World Bank is like a cooperative, where the 187 member countries are shareholders (for more information, see http://web.worldbank.org).

cleaner (less polluting) per kilometre driven and more fuel efficient (e.g. in miles per gallon driven). Chapter 8 gives a detailed overview of these technology oriented policies and their effects.

12.8 Conclusions

The main conclusions are:

1. Governments implement transport policies from a societal perspective. They weigh up all the social costs and benefits of transport, including so-called external effects.
2. External effects of transport are 'real' effects which people and shippers do not take into account when deciding to make a trip or to transport freight. External effects can be negative (costs) as well as positive (benefits). Nearly all policies aimed at external effects relate to decreasing external costs such as congestion delays, air pollution, climate change and accident risk.
3. Governments implement policies because they aim to improve efficiency: with this aim they want to increase total welfare. Another main reason for transport policies is because governments consider them to be fair: with this aim they want to distribute welfare more fairly.
4. Psychological reasons and politicians' self-interest can also explain the implementation of policy.
5. The policy goal to improve accessibility is reached mainly by providing new infrastructure. The improvement of equity is fulfilled mainly by providing new infrastructure and also by subsidizing public transport fares. For policy goals such as improving the environment and liveability, and transport safety, dominant policy instruments are vehicle and fuel standards and regulations which try to improve the technical characteristics of vehicles and fuels.

REFERENCES

Adams, J. (1999), 'The social implications of hypermobility: speculations about the social consequences of the OECD scenarios for environmentally sustainable transport and business-as-usual trend projections', in Project on Environmentally Sustainable Transport (EST), *The Economic Implications of Sustainable Transportation*, Ottawa Workshop, 20–21 October, 1998, Paris: OECD.

Aldred, J. (2002), 'Cost–benefit analysis, incommensurability and rough equality', *Environmental Values*, 11 (1), 27–47.

Aldred, J. (2006), 'Incommensurability and monetary valuation', *Land Economics*, 82 (2), 141–161.

Banister, D. (2002), *Transport Planning*, 2nd edition, London and New York: Spon Press.

Beyazit, E. (2011), 'Evaluating social justice in transport: lessons to be learned from the capability approach', *Transport Reviews*, 31 (1), 117–134.

Buchanan, J.M. and G. Tullock (1962), *The Calculus of Consent*, Ann Arbor: University of Michigan Press.

Button, K.J. (1993), *Transport Economics*, 2nd edition, Aldershot, UK and Brookfield, VT, USA: Edward Elgar.

Button, K.J. and D. Gillingwater (1986), *Future Transport Policy*, London: Routledge.

Cantarelli, C.C., B. Flyvbjerg, E.J.E. Molin and B. van Wee (2010), 'Cost overruns in large-scale transportation infrastructure projects: explanations and their theoretical embeddedness', *European Journal of Transport Infrastructure Research*, 10 (1), 5–18.

Czerwinski, D. and R.R. Geddes (2010), *Policy Issues in U.S. Transportation Public–Private Partnerships: Lessons from Australia*, San José, CA: Mineta Transportation Institute, College of Business, San José State University.

Deb, K. and M. Filippini (2011), 'Estimating welfare changes from efficient pricing in public bus transit in India', *Transport Policy*, 18, 23–31.

Delucchi, M.A. and D.R. McCubbin (2010), *External Costs of Transport in the U.S.*, Davis, CA: Institute of Transportation Studies, UC Davis, accessed January 2011 at http://escholarship.org/uc/item/13n8v8gq.

EC (2009), *A Sustainable Future for Transport: Towards an Integrated, Technology-Led and User Friendly System*, COM(2009) 279 final, Brussels: EC.

Flyvbjerg, B. (2005), 'Policy and planning for large infrastructure projects: problems, causes, cures', Policy Research Working Paper, World Bank, Washington, DC.

Flyvbjerg, B. (2007), 'Megaproject policy and planning: problems, causes, cures', Ph.D. thesis, Aalborg University.

Flyvbjerg, B., N. Bruzelius and W. Rothengatter (2003a), *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge: Cambridge University Press.

Flyvbjerg, B., M.K. Skamris Holm and S.L. Buhl (2003b), 'How common and how large are cost overruns in transport infrastructure projects?', *Transport Reviews*, 23 (1), 71–88.

Flyvbjerg, B., M.K. Skamris Holm and S.L. Buhl (2006), 'Inaccuracy in traffic forecasts', *Transport Reviews*, 26 (1), 1–24.

Geurs, K.T., W. Boon and B. van Wee (2009), 'Social impacts of transport: literature review and the state of the practice of transport appraisal in the Netherlands and the United Kingdom', *Transport Reviews*, 29 (1), 69–90.

Glista, D.J., T.L. DeVault and J.A. DeWoody (2009), 'A review of mitigation measures for reducing wildlife mortality on roadways', *Landscape and Urban Planning*, 91, 1–7.

Gunn, H. (2001), 'Spatial and temporal transferability of relationships between travel demand, trip cost and travel time', *Transportation Research*, 37, 163–189.

Hicks, J.R (1939), 'The foundations of welfare economics', *Economic Journal*, 49 (196), 696–712.

International Transport Forum (2011), accessed February 2011 at http://www.international-transportforum.org/statistics/GlobalTrends/Investment.pdf.

Johansson, P.-O. (1991), *An Introduction to Modern Welfare Economics*, Cambridge: Cambridge University Press.

Kaldor, N. (1993), 'Welfare propositions of economics and interpersonal comparisons of utility', *Economic Journal*, 49 (195), 549–552.

Lovallo, D. and D. Kahneman (2003), 'Delusions of success: how optimism undermines executives' decisions', *Harvard Business Review*, 81 (7), 56–63.

Lucas, K., T. Grosvenor and R. Simpson (2001), *Transport, the Environment and Social Exclusion*, York: Joseph Rowntree Foundation/York Publishing.

Mackie, P.J. and C. Kelly (2007), *Transport Appraisal in Other Countries: Lessons for the NATA Refresh*, Leeds: University of Leeds.

Maibach, M., C. Schreyer, D. Sutter, H.P. van Essen, B.H. Boon, R. Smokers, A. Schroten, C. Doll, B. Pawlowska and M. Bak (2008), *Handbook on Estimation of External Costs in the Transport Sector*, Delft: CE Delft.

Niemeyer, S. and C.L. Spash (2001), 'Environmental valuation analysis, public deliberation, and their pragmatic syntheses: a critical appraisal', *Environment and Planning C: Government and Policy*, 19, 567–585.

Nussbaum, M.C. (2000), 'The costs of tragedy: some moral limits of cost–benefit analysis', *Journal of Legal Studies*, XXIX (June), 1005–1036.

Parry, I. and K. Small (2009), 'Should urban transit subsidies be reduced?', *American Economic Review*, 99, 700–724.

Rietveld, P., J. Rouwendal and A.J. van der Vlist (2007), 'Equity issues in the evaluation of transport policies and transport infrastructure projects', in M. van Geenhuizen, A. Reggiani and P. Rietveld (eds), *Policy Analysis of Transport Networks*, Aldershot: Ashgate, pp. 19–36.

Santos, G., H. Behrendt, L. Maconi, T. Shirvani and A. Teytelboym (2010a), 'Part I: Externalities and economic policies in road transport', *Research in Transportation Economics*, 28, pp. 2–45.

Santos, G., H. Behrendt and A. Teytelboym (2010b), 'Part II: Policy instruments for sustainable road transport', *Research in Transportation Economics*, 28, 46–91.

Schmidtchen, D., C. Koboldt, J. Helstroffer, B. Will, G. Haas and S. Witte (2009), *Transport, Welfare and Externalities*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Sen, A. (2009), *The Idea of Justice*, London, Penguin Books.

Thomopoulos, N., S. Grant-Muller and M.R. Tight (2009), 'Incorporating equity considerations in transport infrastructure evaluation: current practice and a proposed methodology', *Evaluation and Program Planning*, 32 (4), 351–359.

Van Goeverden, C., P. Rietveld, J. Koelemeijer and P. Peeters (2006), 'Subsidies in public transport', *European Transport/Trasporti Europei*, 32, 5–25.

Van Wee, B. (2009), *Transport Policy: What Can It and What Can't It Do?*, European Transport Conference 2009, accessed February 2011 at http://etcproceedings.org/paper/transport-policy-what-it-can-and-what-it-can-t-do.

Van Wee, B. (2012), 'How suitable is CBA for the ex-ante evaluation of transport projects and policies? A discussion from the perspective of ethics', *Transport Policy*, 19 (1), 1–7.

Verhoef, E.T. (1994), 'Efficiency and equity in externalities: a partial equilibrium analysis', *Environment and Planning*, 26A, 361–382.

Verhoef, E.T. (1996), *Economic Efficiency and Social Feasibility in the Regulation of Road Transport Externalities*, Tinbergen Research Series 108, Amsterdam: Thesis Publishers.

13

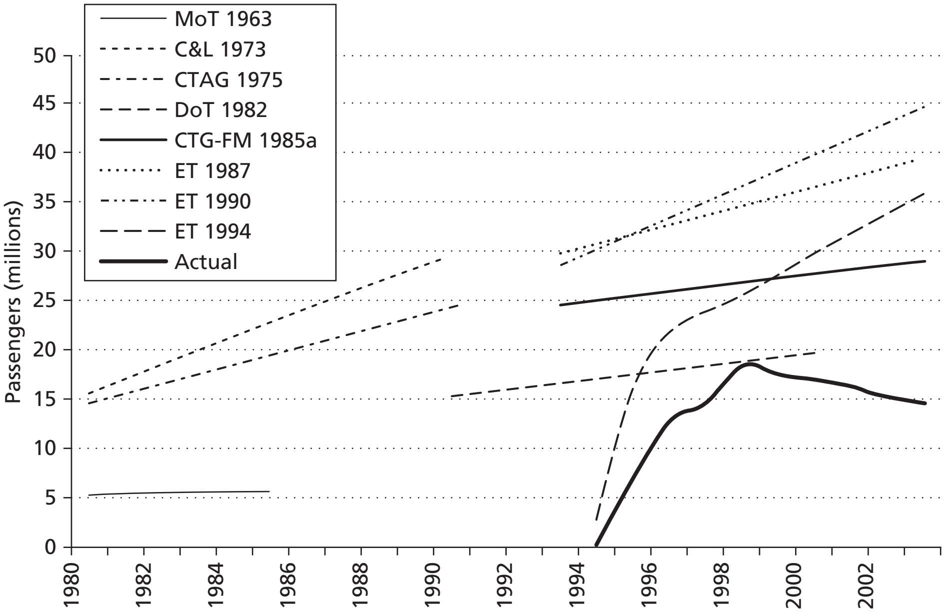
Transport futures research

Vincent Marchau, Jan Anne Annema, Warren Walker and Jan van der Waard

13.1 Introduction

It has often been said that 'to govern means to foresee'. This adage is also valid for transport policy making. For example, an increasing transport demand in the future could lead to politically unacceptable future levels of congestion, air pollution, and traffic casualties if no additional policies were implemented. Thus, it is important for governments to know about possible future transport expectations in their region in order to be able to implement new policies in time. For several decades, governments have been developing policies to reduce the negative impacts of transport (see Chapters 9–12). Chapter 12 discusses criteria for 'sound' policies. But, how are the future impacts of candidate transport policy options to be explored? To do this we enter the area of transport futures research – this area is the scope of this chapter. Chapter 14 will then discuss how to evaluate *ex ante* the impacts of alternative policies via a cost–benefit analysis or multi-criteria analysis. Chapter 15 discusses the use of transport (impact) models.

Research can map plausible futures and transport policy making strategies. Here, it is very important to note the term 'plausible'. In fact, the future is unknown and is largely determined by non-predictable developments (Taleb, 2007). This implies that future outcomes are surrounded by a lot of uncertainty. In our view, in proper transport futures research this uncertainty is adequately taken into account and clearly communicated to the decision makers. In poor transport futures research, the opposite is true. The analysts in poor quality transport futures research often seem to think that they are able to predict the future, which is, of course, impossible. An example of poor transport futures research is shown in Figure 13.1. This figure shows the forecasted and actual number of passengers for the Eurostar in different forecast studies of the past, which are abbreviated MoT, C&L, DoT, and so forth (see Figure 13.1, top left). The competition from low cost air carriers and



Source: Anguera (2006).

Figure 13.1 Channel Tunnel passengers traffic forecasts and actual results

the price reactions of ferries, amongst others, were not taken into account in most studies for the rail tunnel under the English Channel. This resulted in a significant overestimation of the tunnel's revenues and market position (Anguera, 2006), with devastating consequences for the project.

This chapter is written from the perspective of transport policy analysis. The aim of the chapter is to specify research approaches to study transport futures and to explain the role of futures research related to transport policy analysis. Generally speaking, futures research in transport policy analysis is carried out for two reasons: (1) to identify the types and magnitudes of future transport problems, and (2) to identify ways to reduce these transport problems.

With respect to future transport problems, an increase in future transport could result in societally unacceptable levels of air pollution, congestion, and traffic casualties, for example (see Chapters 9–11). Here, transport futures research supports estimating the future levels of these problems. By doing so, policy makers may decide to implement new policies, for example. On the other hand, futures research on economic and demographic developments, for example, might show that transport demand will decrease. Thus problems could be solved, or at least reduced, without intervention.

Next, if future transport problems are identified, policy makers often desire to know the policy options that could help solve the problems, their effects, and their costs and benefits (see Chapter 14). The specification and analysis of current and future options are not trivial tasks. For instance, some options might currently be under development (e.g. new vehicle technologies) or even unknown. In addition, most transport policy options have a long-term character. For instance, the benefits from new infrastructure will be realized over decades after the moment of opening. Furthermore, it takes years before the full impacts of pricing and technical measures for new vehicles are reached. Transport futures research can help to specify current and future options and to estimate their long-term impacts.

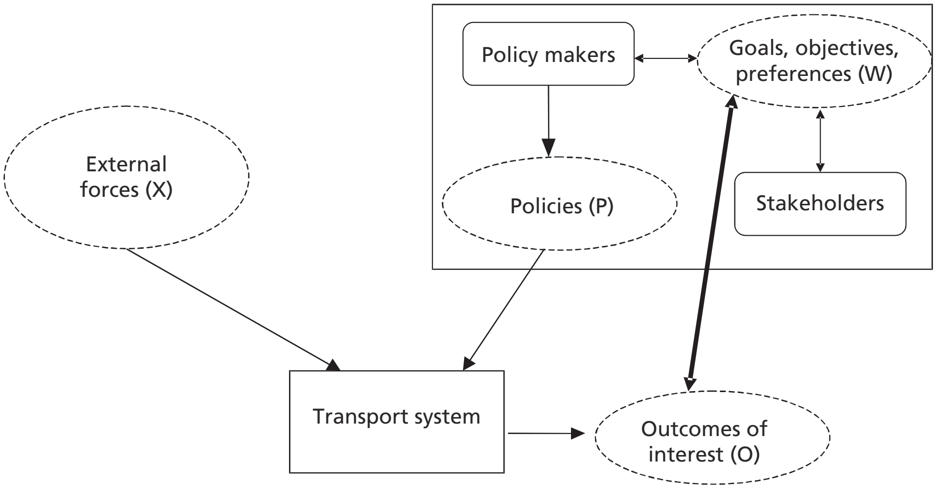
In section 13.2, futures research in relation to the transport policy domain is explained, and a framework for futures uncertainty is discussed. In section 13.3 the traditional scenario planning approach and the exploratory modelling and analysis (EMA) approach are explained. In section 13.4, the 'backcasting' method is treated. The role of research in designing dynamic adaptive policies (DAP) is presented in somewhat more detail in section 13.5. Finally, section 13.6 contains the conclusions.

13.2 Futures research and transport policy analysis

Futures research to help public decision making starts with an understanding of the policy domain. As detailed by Walker (2000a), a common approach to a rational-style policy analysis is to create a model of the system of interest (in this book, the transport system) that defines the boundaries of the system and describes its structure and operations, that is, the elements, and the links, flows, and relationships among these elements (see Figure 13.2).

In Figure 13.2 different elements and links can be distinguished:

1. **External forces** (X) are forces that work from outside the transport system, that is, are not under the control of the problem owner. These forces have influence on the demand for transport and the supply of transport (e.g. technical developments for vehicles and infrastructure, oil price, cultural changes, and so forth).
2. The **transport system** box represents the elements of the transport system (e.g. drivers, operators, vehicles, infrastructure) and their interactions. These elements and their interactions are affected by the external forces and future policies (see section 13.3, 'Forward-looking scenarios') and result in intermediate transport system outputs (such as the amount



Source: Based on Walker (2000a).

Figure 13.2 Schematic view of the transport system from a policy analysis perspective

of transport in passenger kilometres and/or ton kilometres per transport mode in a future year).

3. The **outcomes of interest** (O) box represents policy relevant outcomes from the transport system – output such as traffic congestion, air pollution, and traffic casualties. In an explanation of a future situation, these are the transport system output levels as estimated by the transportation model (see Chapter 15 on transport modelling).
4. The estimated outcomes of interest for the future may not be in accordance with policy goals or preferences. This produces a need for new future transport **policies** (P), such as new infrastructure, road or fuel pricing, stricter vehicle emission standards, and so forth, which can be fed into a new estimation of outcomes via the link from policies to the transport system.
5. The **valuation of outcomes (weights** (W)) involves the relative, subjective importance given to the outcomes of interest by crucial stakeholders. It involves how stakeholders value the results of the changes in the transport system, such as improved traffic efficiency, fewer fatalities, reduced emissions, and so on (see Chapter 14 on evaluation methods).

To be most useful (and to increase the chances of the results of a policy analysis actually being used) a policy analysis study should be carried out as a partnership between the policy makers and the researchers. The steps in doing a policy analysis study are, in short:

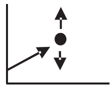
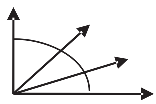
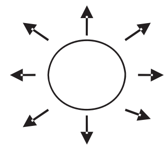
1. Formulate the transport problem: problem definition, setting goals and specifying options (often done in close partnership between the researchers and the problem owners or policy makers).
2. Estimate the impacts of the various policy options using transport system models, for example, and analyse the options (done mainly by the researchers).
3. Compare the options (done mainly by the problem owners or policy makers, but often supported by the researchers through quantitative, analytical tools, e.g. cost–benefit analysis or multi-criteria analysis).
4. Choose and implement the chosen option (done mainly by the problem owners or policy makers).

This chapter focuses on methodological approaches for estimating the future impacts of policies and external forces on the transport system. An essential criterion for choosing an approach is how uncertain the future is assumed to be, that is, the level of future uncertainty assumed. In general, uncertainty can be defined as limited knowledge about future, past, or current events. Formally, as defined by Walker et al. (2003), we will consider uncertainty in this chapter to be 'any departure from the (unachievable) ideal of complete determinism'. Or, in mathematical terms:

Let Y be some event. If Probability(Y) $\neq$ 0 or 1, then the event Y is uncertain.

This abstract formula can be illustrated with an example of a future transport outcome of interest: suppose Y is the estimate produced by a model of the carbon dioxide (CO_2) emitted by road transport in 2030 in some country X. The model estimates that Y = 25 billion kilograms. The probability of this event actually happening in 2030 is not 0 or 1. In fact the probability is unknown. Thus the estimate of 25 billion kilograms of CO_2 emitted by road transport in 2030 is uncertain.

Based on the policy analysis framework (Figure 13.2), a classification of uncertainties with respect to policy making can be made. Such a classification was developed by Walker et al. (2003). For the purposes of this chapter, we do not need to elaborate on the issue of uncertainty (for this, we refer the reader to Funtowicz and Ravetz, 1990; van der Sluijs, 1997; van Asselt and Rotmans, 2002; Walker et al., 2003). Here, the most important notion is to realize that transport policy analysis problems can be characterized by different levels of uncertainty about the external forces (X), the transport system and transport system models, the outcomes of interest (O), and valuations of the outcomes (W) (e.g. Courtney, 2001; Walker et al., 2003; Makridakis et al., 2009; Kwakkel et al., 2010a).

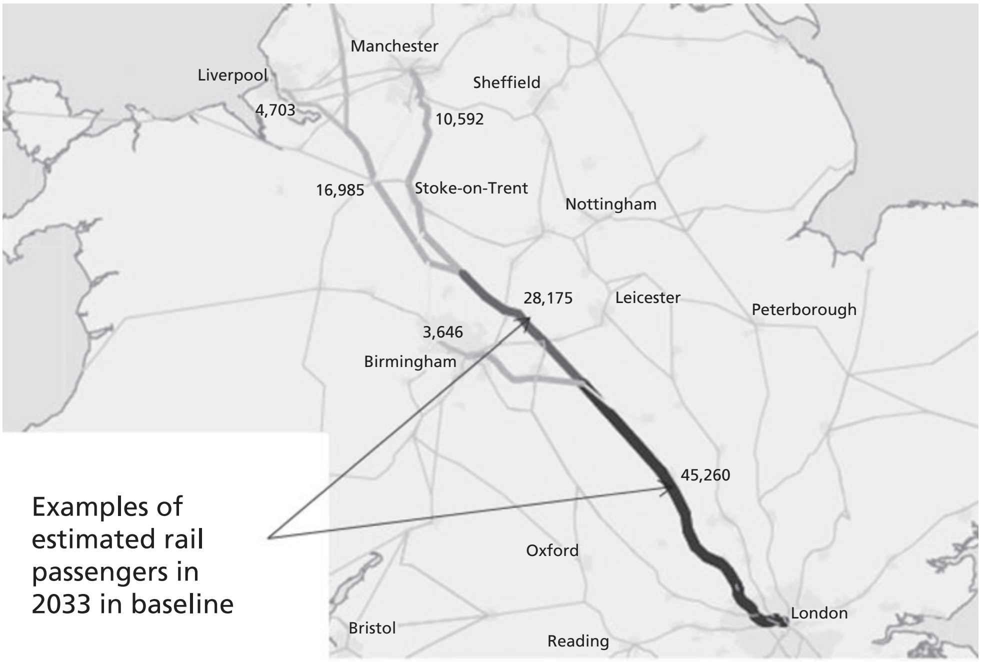
		LEVEL					
		Complete certainty	Level 1	Level 2	Level 3	Level 4	Total ignorance
LOCATION	**Context**		A clear enough future 	Alternative futures (with probabilities) 	A multiplicity of plausible futures 	Unknown future 	
	System model		A single (deterministic) system model	A single (stochastic) system model	Several system models, with different structures	Unknown system model; know we don't know	
	System outcomes		A point estimate for each outcome	A confidence interval for each outcome	A known range of outcomes	Unknown outcomes; know we don't know	
	Weights on outcomes		A single set of weights	Several sets of weights, with a probability attached to each set	A known range of weights	Unknown weights; know we don't know	

Source: Walker et al. (2003).

Figure 13.3 The progressive transition of levels of uncertainty from complete certainty to total ignorance

A way of representing different levels of uncertainty is shown in Figure 13.3. Level 1 uncertainty is often treated through a simple sensitivity analysis of transport model parameters, where the impacts of small perturbations of model input parameters on the outcomes of a model are assessed. Level 2 uncertainty is any uncertainty that can be described adequately in statistical terms. In the case of uncertainty about the future, Level 2 uncertainty is often captured in the form of either a (single) forecast (usually trend based) with a confidence interval, or multiple forecasts ('scenarios') with associated probabilities.

Many (quantitative) analytical approaches for transport policy analysis deal with future uncertainties as being Level 1 and Level 2 uncertainties, which is highly questionable because the future is in almost all cases 'not clear enough' (as is assumed by Level 1), or it is hardly possible to attach probabilities to different possible futures (as is assumed by Level 2). As an example of a Level 1 mistaken attribution, Figure 13.4 gives figures for the expected number of rail passengers in 2033 on the West Coast Main Line in the UK. In this case, the analysts have informed decision makers incorrectly about their assumptions about the future (only some sensitivity analysis was carried out).



Source: High Speed Two (HS2) Limited (2010).

Figure 13.4 Baseline prognosis of West Coast Main Line rail passengers in 2033 (treated as a Level 1 uncertainty), as used in the benefits study for building a high speed rail link from London to the West Midlands and beyond. The number of future daily long-distance trips is presented. The baseline in this study is one single future estimate. In the report, sensitivity analysis is used to evaluate, amongst other things, the impact of other future rail demand forecasts

In the view of the authors of this chapter, all long-term transport policy analysis problems are characterized by higher levels of uncertainty (i.e. Levels 3 and 4; see Figure 13.3). Only relatively short-term 'predictions' (for example, forecasts for one day or one week ahead of congestion levels on certain road stretches) can be characterized as Level 1 and Level 2 uncertainties.

The long-term related Level 3 and Level 4 uncertainties cannot be dealt with through the use of probabilities and cannot be reduced by gathering more information, but are basically unknowable and unpredictable at the present time. And these higher levels of uncertainty can involve uncertainties about all aspects of a transport policy analysis problem – external or internal developments, the appropriate (future) system model, the parameterization of the model, the model outcomes, and the valuation of the outcomes by (future) stakeholders. Many of the negative consequences from policy decisions in the past were due to the use of approaches that did not take into account the fact that they were facing conditions of Level 3 and Level 4 uncertainty. For example, Figure 13.1 shows the result of some poor reference forecasts (Anguera, 2006), which turned out to be far too optimistic about future demand, because, among other things, future uncertainty about competitors' changes (e.g. low budget aircraft started to flourish, low ferry prices were introduced) were ignored. Another example of not taking into account the right level of uncertainty in transport futures research is shown by Flyvbjerg et al. (2003, 2006). They investigated 210 projects on traffic demand shortfalls, comparing the first year after introduction with the *ex ante* estimate. Large inaccuracies in the prognoses were observed, mainly for rail projects.

13.3 Level 3 approaches: forward-looking scenarios and exploratory modelling

Forward-looking scenarios

When faced with Level 3 uncertainties, transport policy analysts will generally use scenarios. The core of this approach is the assumption that the future can be specified well enough to identify policies that will produce favourable outcomes in one or more specific plausible future worlds. The future worlds are called scenarios. Börjeson et al. (2006) call these 'explorative scenarios' to differentiate them from 'predictive scenarios', which some analysts think they can use to deal with Level 1 and Level 2 uncertainties (which is not the case for long-term transport planning, in our view), and 'normative scenarios', which use backcasting (see, for example, Quist, 2007, and section 13.4) to determine how a specific desired target can be reached.

Scenarios are 'stories' of possible futures, based upon logical, consistent sets of assumptions, and fleshed out in sufficient detail to provide a useful context for engaging planners and stakeholders. A forward-looking scenario includes assumptions about developments within the system being studied and developments outside the system that affect the system, but exclude the policy options to be examined (see also Figure 13.2). Because the only sure thing about a future scenario is that it will not be exactly what happens, different scenarios, spanning a range of developments, are constructed to span a range of plausible futures. No probabilities are attached to the futures represented by each of the scenarios. They have a qualitative function, not a quantitative function. Scenarios do not tell us what will happen in the future; rather they tell us what can (plausibly) happen. They are used to prepare for the future: to identify possible future problems and to identify robust (static) policies for dealing with the problems.

In transport policy analysis, best estimate models are often used (based on the most up-to-date scientific knowledge; see Chapter 15) to examine the consequences that would follow from the implementation of each of several possible policies. This 'impact assessment' is done for each of the scenarios. The 'best' policy is the one that produces the most favourable outcomes across the scenarios. Such a policy is called a robust (static) policy.

There is no general theory that allows us to assess scenario adequacy or quality. There are, however, a number of criteria that are often mentioned in the literature as being important. Schwarz (1988) gives a brief summary of them. The most important of these are consistency, plausibility, credibility, and relevance.

1. **Consistency:** the assumptions made are not self-contradictory; a sequence of events could be constructed, leading from the present world to the future world.
2. **Plausibility:** the posited chain of events can happen.
3. **Credibility:** each change in the chain can be explained (causality).
4. **Relevance:** changes in the values of each of the scenario variables are likely to have a great effect on at least one outcome of interest.

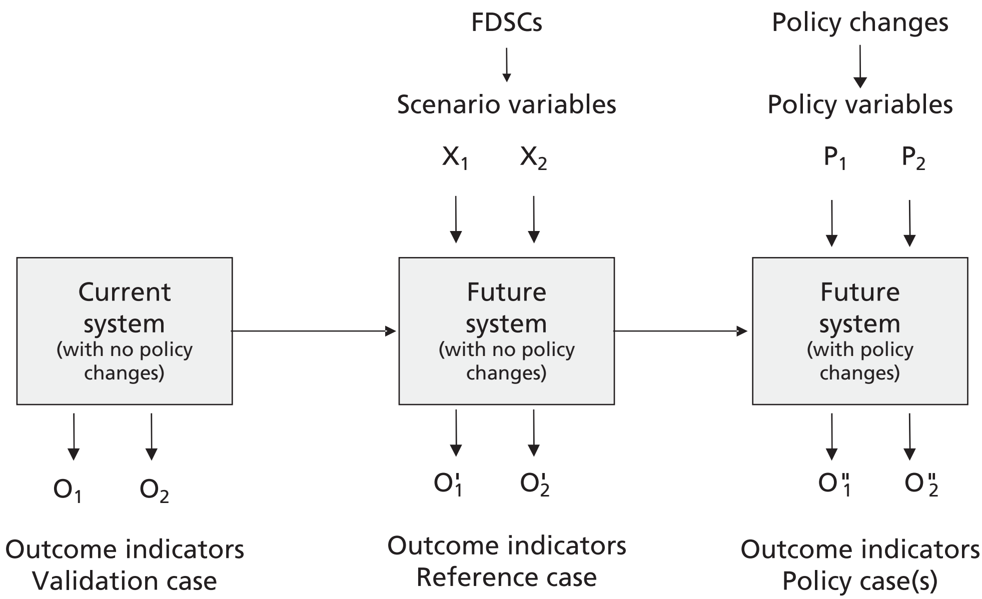
A structured process for developing forward-looking scenarios, consisting of a number of explicit steps, has been used in several policy analysis studies. The steps, summarized by Thissen (1999), and based on the more detailed specifications of Schwartz (1996), RAND Europe (1997), and van der Heijden, et al. (2002), are (see also Figure 13.2):

1. **Specify the system, its outcomes of interest, and the relevant time horizon.** A system diagram can be used to identify what is considered inside and outside the system, the system elements that affect or influence the outcomes of interest, and their interrelationships.
2. **Identify external forces (X) driving changes in the system and thereby producing changes in the outcomes of interest (O).** Whether or not a particular external force is potentially relevant depends on the magnitude of the change in the system and its implications for the outcomes of interest. There are many judgements involved in defining the system under consideration, the relationships among the subsystems, and the definition of what is relevant. Thus the determination of relevant forces and changes is necessarily subjective. Potentially relevant forces and changes are often best identified by conducting a series of interactive brainstorming or focus group sessions involving experts and/or stakeholders.

Table 13.1 Selecting relevant forces or system changes for forward-looking scenarios

	Change would lead to a low impact (for all outcomes of interest)	Change would lead to a high impact (on at least one outcome of interest)
Force or change is uncertain	These forces or changes can be left out of the scenarios (or left in for colour).	These forces or changes are candidates for scenarios.
Force or change is fairly certain	These forces or changes can be left out of the scenarios (or left in for colour).	These forces or changes are included in all the scenarios as 'autonomous developments'.

3. **Categorize forces and resulting system changes as fairly certain or uncertain.** The forces or system changes from step 2 are placed into one of two categories – fairly certain or uncertain (see Table 13.1). Those forces or system changes about which the researcher is fairly certain are placed into this category. The remaining forces or changes are placed into the uncertain category. The forces or system changes in the fairly certain category are included in all the scenarios. The uncertain forces or system changes are used to identify the most important and relevant uncertainties that have to be taken into account.
4. **Assess the relevance of the uncertain forces or system changes.** The analyses should focus on the uncertain forces or system changes that have the greatest effects on the outcomes of interest. To identify them, the impact of each uncertain force or system change is considered with respect to each of the outcomes of interest. Based on the estimated impact that the resulting system change has on the outcomes of interest, the force or system change is placed in either a high or a low impact category (see Table 13.1). The uncertain forces and system changes in the low impact category are dropped from further consideration (or can be left in for 'colour'). The uncertain forces and system changes in the high impact category (those that have a high impact on at least one of the outcomes of interest), along with the fairly certain elements, form the basis for the scenarios.
5. **Design several future scenarios based on combinations of different developments in the driving forces.** These should provide strikingly different images of the future that span the space of what is plausible. A brief but imaginative description of the essential characteristics of the future depicted by each of the scenarios should then be provided. Once the specific scenarios are identified, the assumptions underlying them are converted into inputs that can be used by the system models. This forms the basis for the subsequent assessment of policy options.

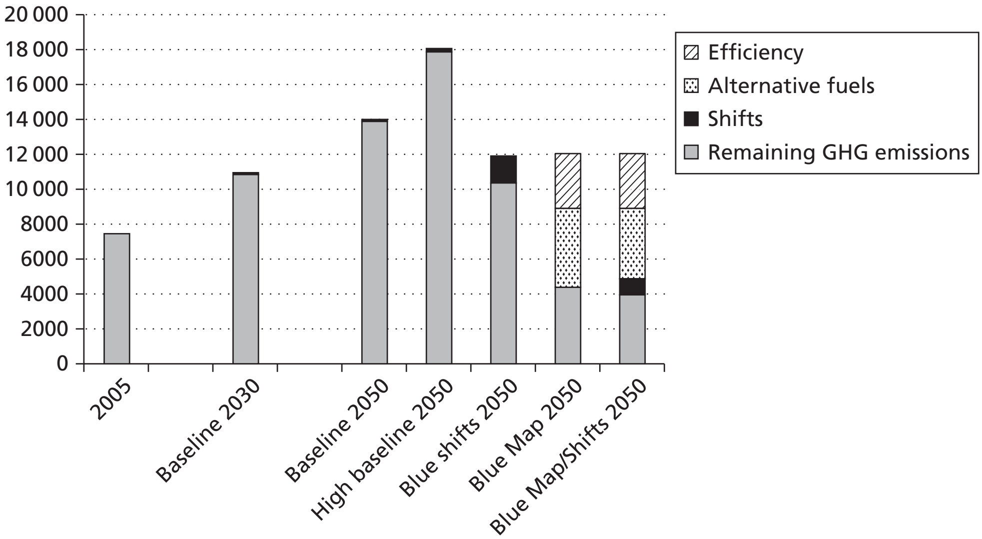


Note: FDSCs: forces driving structural change.

Figure 13.5 Evaluating policy options using scenarios

After constructing scenarios using steps 1 to 5, these are first used to specify the (magnitude of the) future problem if no additional action is undertaken. For example, these 'reference scenarios' might imply high congestion levels, a high increase in carbon dioxide emissions, and so forth. In other words, related to Figure 13.2, the scenario outcomes of interest may not be in accordance with the goals. The idea is that these reference scenarios assume that only the already-existing and/or agreed-upon policies will be implemented. Thus the need for additional policies can be identified. In practice, these reference scenarios are sometimes given other names, for example business-as-usual scenarios, baseline scenarios (e.g. Figure 13.4), or background scenarios.

In evaluating the impacts of policy options, the researcher should always include the reference scenario (see Figure 13.5; see Chapter 14). If a transportation model (see Chapter 15) was used to estimate the reference scenario outcomes, the same model should be used again with exactly the same input, except for the input parameter changes or model changes due to the policy option (or options) under study. In the 'validation case', the current system is used in the model to make sure that the outcomes are reasonably close to the real world outcomes. More on model validation can be found in Chapter 15. For example, if the policy option under study is to increase petrol fuel levies, in the model runs for the policy cases only the petrol fuel price inputs are



Note: Emission reductions in Mt CO_2 eq.

Source: IEA/OECD (2009).

Figure 13.6 Worldwide transport GHG emission reductions by different policies

changed. Consequently, the policy maker can see the impacts of the fuel levy increase on, for example, petrol car use, petrol fuel-efficiency, a shift to more diesel car usage (possibly in Europe) and so forth. To be clear, the impact assessment shows the differences between the validation case (O_1 and O_2), the reference outcomes of interest (O_1' and O_2') and the outcomes of interest due to the policy measure (O_1'' and O_2''). These differences – the final pros and cons of the policy option – can be weighed up by the policy makers using different methods. For example, they can ask for a multi-criteria analysis (MCA), a cost-effectiveness study, or a cost–benefit analysis (see Chapter 14).

The policy option or options to be studied and the relevant outcomes of interest are dependent on the policy question. Impacts of new roads compared to the reference scenarios can be studied, or extra investments in public transport, or new vehicle emission standards, or kilometre charging, and so forth. It is also possible to evaluate the impacts of policy packages (combinations of policies). Figure 13.6 gives an example of an evaluation of the impact of different policy packages on worldwide greenhouse gas (GHG) emissions by transport in the long term (2050) (IEA/OECD, 2009). The figure shows that, for 2050, two reference scenarios were constructed: baseline and high baseline. Furthermore, the impacts of three policy packages (called 'policy scenarios') on GHG emissions compared to the reference scenarios are shown. One policy scenario is called 'Blue Shifts' and contains only policy measures aimed to shift people and freight from relatively low energy-efficient transport modes to higher energy-efficient ones. 'Blue Map' focuses only on policy measures to improve vehicle technologies: the introduction of high efficiency technologies in conventional fossil fuel-based vehicles and the introduction of alternative transport fuels, such as hydrogen, biofuels, and electricity (Chapters 8 and 10). Finally, 'Blue Map/Shifts' shows the impact of the two policy packages combined.

Pros and cons of the scenario approach

The benefits from using scenarios in policy analysis are threefold. First, scenarios help analysts and policy makers deal with situations in which there are many sources of uncertainty. Second, scenarios allow analysts to examine the 'what ifs' related to external uncertainties. They suggest ways in which the system could change in the future and facilitate the examination of the implications of these changes. Finally, scenarios provide a way to explore the implications of Level 3 uncertainties for policy making (prepare for the future) by identifying possible future problems and identifying (static) robust policies for dealing with the problems.

However, from an analytic perspective the scenario approach has several problems. The first problem is deciding which assumptions about future external developments to include in the scenarios. Typically, these assumptions are decided upon by experts (collectively and individually). However, in the face of uncertainty, no one is in a position to make this judgement. A second problem is that the researcher has little idea about whether the range of futures provided by the scenarios covers all, 95 per cent, or some other percentage of the possible futures. A third problem with the scenario approach has to do with range in the performance estimates generated by the scenarios. If this range is large, policy makers often tend to fall back on a single, 'most likely' scenario (implicitly assuming Level 2 uncertainty) or the do-nothing approach, arguing that 'We do not have sufficient information to make a decision at this time.' The latter is probably the worst possible outcome – when the level of uncertainty is high and the potential consequences are great, it would probably be better if policy makers acted rather than waited.

The exploratory modelling and analysis approach

An approach to overcome some of the disadvantages of scenarios is the EMA approach (Agusdinata, 2008), or the closely related robust decision making (RDM) approach (Lempert et al., 2006).

EMA begins by acknowledging the fact that a validatable predictive policy analysis model cannot be built (since the future is unknown). It then asks the question 'In that case, what can the researcher do with the system model?' Rather than attempting to predict system behaviour, EMA aims to analyse and find reasons for the system's behaviour (Bankes, 1993). Under conditions of Level 3 (and even Level 4) uncertainty, even a model that cannot be validated can still be useful (Hodges, 1991). One use is as a hypothesis generator, to get insight into a system's possible behaviours. A combination of input and system variables can be established as a hypothesis about the system. One can then ask what the system behaviour and the resulting outcomes would be if this hypothesis were correct.

EMA supports this process of researching a broad range of assumptions and circumstances. In particular, EMA involves exploring a wide variety of scenarios, alternative model structures, and alternative value systems. The exploration is carried out using computational experiments. A computational experiment is a single run of a given model structure and a given parameterization of that structure. It reveals how the real world would behave if the various hypotheses presented by the structure and the parameterization were correct. By exploring a large number of these hypotheses, one can get insights into how the system would behave under a large variety of assumptions. To support the exploration of these hypotheses, data mining techniques for analysis and visualization are employed. EMA aims to 'cover the space' of possibilities, which can be described as the space being created by the uncertainty surrounding the many variables.

A potential problem in exploratory analysis is that, since the number of uncertainties is large, the number of model runs will be large. The database containing the outcomes from these runs will, therefore, be very large. It is very difficult for anyone to scan this large database and interpret the results in order to identify a preferred policy for each of the plausible scenarios. Therefore, software that offers graphical tools to summarize the results of an exploratory analysis is required. Given complete sets of external forces, policies, system models, outcomes, and their weights, software is available that is able to determine the values of the (uncertain) parameters that would lead to preferences for the different policy options (i.e. it is able to map the decision space).

Agusdinata (2008), Kwakkel et al. (2010b), and van der Pas et al. (2010) supply examples of how EMA can be applied to transport policy analysis problems involving Level 4 uncertainty.

13.4 The backcasting approach

Backcasting is quite different from the forward-looking scenario approach described in section 13.3. Here, a normative target in the future – a desired set of outcomes – is chosen as the starting point of the future analysis; then appropriate paths towards this desired outcome are searched for. In general, in backcasting, first an image is found that might be a future solution for the societal problem at hand. If such an image can be made explicit, the next step is to identify and assess a path between today and that future image. If no path can be found, the image will be redeveloped and adjusted (Hojer, 1998).

Within transport, several studies on sustainable development, or specifically on reducing CO_2 emissions in future transport, have used the backcasting method – for example the OECD project Environmentally Sustainable Transport, the EU project POSSUM, and the UK project VIBAT (Banister et al., 2000; Geurs and van Wee, 2000; Hickman and Banister, 2005).

Four different steps in the backcasting process related to policy making can be distinguished:

1. **The definition of a future target or targets.** For example, the objective is a 60 per cent reduction in CO_2 emissions for long-distance transport in Europe in 2047 compared to the 2005 level. This is one of the normative targets in a backcasting study on the future of European long-distance transport (Schippl et al., 2008).
2. **The construction of a reference (forward-looking) scenario.** This reference case should illustrate the scale of change required to meet the target or targets. For example, a reference scenario can show that, without additional policies, the CO_2 emissions in the long-distance transportation sector in Europe will increase by 30 per cent in 2047 compared to the current level. If the objective is a 60 per cent reduction, this reference scenario points to a huge gap that has to be bridged by implementing additional policies.
3. **The design of 'images of the future'.** Images are descriptions of the future that (from today's point of view) seem to meet the targets (Schippl and Leisner, 2009). Table 13.2 gives the three images in the backcasting study on the future of European long-distance transport. The specific content of this table is not of particular interest for this book. What is more interesting is to see that three different future images were created. Banister et al. (2008) have suggested criteria for future images. These are (Schippl and Leisner, 2009): (1) the images should meet the targets; (2) each image should be plausible, but can be relatively extreme; (3) the

Table 13.2 Three images in the backcasting study on the future of European long-distance transport

2047	Image 1 Strong and rich high-tech Europe	Image 2 Slow and reflexive life-styles	Image 3 (contrast image): economic pressure + very expensive energy
Governance	EU is cohesive and has a leading role in the world	Strong UN has established successful climate instruments	Weak EU, weak UN, limited international cooperation
Economy/GDP growth	Roughly 2.4%	Roughly 1.7%	Roughly 0.7%
Lifestyles	Consumption-oriented, fast	Focus on health and quality of life	
Means of accessibility	Air and high-speed rail	Virtual mobility and comfortable rail (and slow air)	Air and virtual mobility
Main LDT fuels	Electricity, hydrogen, biofuels, CNG, kerosene	Electricity, biofuels, CNG, diesel, kerosene	Biofuels, CNG, diesel, kerosene
Biofuels share 2047	30%	25%	15%
Improvement of carbon intensity for aviation (2005–47)	64.3%	58.3%	58.3%
Improvement of carbon intensity for trucks (2005–47)	57.2%	44.1%	40.1%
Transport volume 2047 compared to baseline[5]	–30%	–45%	–60%

Source: Schippl et al. (2008).

images should be clearly different from each other, in order to give an idea of the huge variety of possible futures; and (4) the images should cover a sufficiently wide range of possibilities. However, to keep the research manageable, a small number of images should be selected.

4. **The specification of potential policies.** Policies that might help meet the images are specified. They are then analysed and assessed by identifying the trajectories leading from the future images back to the present state and vice versa. Note that, in some backcasting studies, the trajectories proposed are also called scenarios, which can be confusing. The assessment could, for example, result in the notion that, whichever of the 2047 images will have to come true to meet the strict CO_2 target (Table 13.2), strong government policies are required, such as policies that result in improvement of carbon intensities for trucks.

Related to the uncertainty framework used in this chapter (Figure 13.3), we can identify two potential weak points in state-of-the-art backcasting methods. First, it seems highly risky to assume the future will turn out to be as forecasted by the reference scenario (step 2), as was assumed in the Schippl et al. (2008) example. With this assumption, the future is treated as a Level 1 uncertainty, which is, of course, untrue. Here, the risk is that the policy makers are given a false sense of certainty about the predictability of future outcomes without additional policies, which may result in wrong policy actions (too many or too few changes to the system, or changes made too early or too late). Second, in step 4 it is (implicitly) assumed that, if implemented, the specified policies will actually lead to the desired future, which is an incorrect assumption.

13.5 Level 4 approaches: flexible and adaptive approaches

The previous sections focused on approaches to handle Level 3 uncertainties. However, transport policy problems increasingly emerge in which the uncertainty can be characterized as Level 4. In this case, what is known is only that the researcher does not know the future situation. Level 4 uncertainty is also called 'deep uncertainty'; it is defined as a condition in which analysts do not know (and/or the parties to a decision cannot agree upon) (1) the appropriate models to describe interactions among the system's variables, (2) the probability distributions to represent uncertainty about key parameters in the models, and/or (3) how to value the desirability of alternative outcomes (Lempert et al., 2003).

In most policy analysis studies involving lower levels of uncertainty, the study ends with the researcher presenting the impacts of alternative policies, leaving the choice and implementation of a preferred policy to the policy maker(s) (although the analyst and policy maker(s) should be working closely during the course of the study, as stated in section 13.2). In the case of deep uncertainty, the implementation step of a policy analysis is explicitly addressed by the researcher. This 'implementation research' focuses on how the chosen policy could fail and ways to protect it from failing.

Some scientists have thought about analytical approaches to making policies in the face of deep uncertainty. The key idea is not to develop a static plan that will work well for one or more specific futures, but rather to construct a dynamic plan that is flexible and adaptable, which will perform well across the full range of plausible futures (including surprises) (Walker et al., 2001; Makridakis et al., 2009; Wright and Goodwin, 2009).

For instance, van Wee (2009) suggests using flexibility and robustness as criteria when designing new transport policies (see Chapter 12). Flexibility of a policy relates to the ease with which the policy can be adapted. For example, once introduced, policies such as emissions regulations and levies on fuels and cars can be changed relatively easily. Another example of flexibility he gives relates to new infrastructure. He examines a region with planned urban development at several locations. Based on long-term urban planning, a tram line could be considered. But what if, owing to unexpected demographic changes, a housing crisis, or a lack of firms interested in new employment areas, the original plan cannot be realized? In that case a (free) bus line could be an option, at least for the first 10 years. Depending on real world developments, the track could later be converted into a tram line.

His second criterion concerns the long-term robustness of a policy. This criterion relates to the question of whether a policy is 'no-regret' under uncertain long-term developments that could have a major impact on society. Examples of uncertain but major impact long-term developments are climate change, the depletion of fossil fuels, very high or unstable oil prices, strict CO_2 policies, major shifts in goods flows strongly affecting the position of harbours and freight transport networks, or a breakthrough in battery or hydrogen technologies or in solar or wind power generation, resulting in the availability of large amounts of cheap and clean energy. Van Wee recommends a check on robustness in cases of expensive land-use or transport infrastructure policies.

In general, the literature offers three (overlapping, not mutually exclusive) ways for dealing with deep uncertainty in making policies, although there are differences in definitions, and ambiguities in terminology (see, for example, Leusink and Zanting, 2009):

1. **Resistance:** plan for the worst conceivable case or future situation (e.g. over-dimensioning of infrastructure).
2. **Resilience:** whatever happens in the future, make sure to have a policy that will result in the system recovering quickly (e.g. floating roads, traffic incident management).
3. **Adaptive robustness:** prepare to change the policy if conditions change (i.e. use Dynamic Adaptive Policies (DAP) – see below).

The first approach is likely to be very costly and might not produce a policy that works well, because of 'black swans'. The black swans metaphor is used by Taleb (2007) to explain that many events in the world are a surprise (to the observer) and can have major unforeseen impacts on world development. The second approach accepts short-term pain (negative system performance), but focuses on recovery. The third approach appears to be the most robust and efficacious way of dealing with deep uncertainties (Kwakkel et al., 2010b). It will be explained briefly below.

The basic concept of a dynamic adaptive policy is easy to explain (Walker, 2000b). It is analogous to the approach used in guiding a ship through a long ocean voyage. The goal – the end point – is set at the beginning of the journey. But, along the way, unpredictable storms and other traffic may interfere with the original trajectory. So the policy – the specific route – is changed along the way. It is understood before the ship leaves port that some changes are likely to take place – and contingency plans may have already been formulated for some of the unpredictable events. The important thing is that the ultimate goal remains unchanged and the policy actions implemented over time remain directed toward that goal. (If the goal is changed, an entirely new plan must be developed.) An adaptive policy would include a systematic method for monitoring the environment, gathering information, implementing pieces of the policy over time, and adjusting and readjusting to new circumstances. The policies themselves would be designed to be incremental, adaptive, and conditional.

We now discuss the steps to operationalize DAP using an example. The example concerns strategic planning for a large airport close to a built-up area. The design of the adaptive policy consists of four steps (from Marchau et al., 2010):

- **Step 1: Specification of problem, objectives, the definition of success, and constraints**
  In the past decade, the rate of growth in air traffic has been twice as large as the growth of the world economy. It is expected that, owing to the increase of the world population, economic growth, and globalization, air traffic will continue to grow. Hence, an objective of an airport operator might be to improve the airport's capacity to handle increased demands. The related definition of success is that future capacity will meet future demand. Success means having a good match between supply and demand – not too much capacity, which would mean a lot of unused capacity, but not too little capacity, which would lead to delays in take-offs and landings. The constraints on policy options include costs, safety, life quality, spatial restrictions and public acceptance.
- **Step 2: Specification of a basic policy and its conditions for success**
  A basic policy might be to expand the physical capacity of the airport (add a runway). Conditions for the success of this basic policy include that demand continues to grow and that the extra aircraft noise generated

does not bring strong protests. Traditional policy analysis tools are available for identifying a basic policy (Findeisen and Quade, 1985).

- **Step 3: Identifying the vulnerabilities of the basic policy and anticipatory actions to protect it**
  In step 3 of the DAP process, the actions to be taken immediately to enhance the chances of success of the basic policy are specified. This step is based on identifying, in advance, the vulnerabilities associated with the basic policy, and specifying actions to be taken in anticipation. Vulnerabilities are external developments that could degrade the performance of the policy so that it is no longer successful. In short, the question is asked 'How can the basic policy fail?', and then actions are designed to prevent it from failing.

  Scenarios are used in this step and in Step 4; but they are used in a different way from the way they are used in dealing with Level 3 uncertainty. They are used to identify the ways in which the basic policy could go wrong (i.e. not lead to success). In DAP, since the researcher is looking for changes in the world that can make the basic policy fail, the scenarios should differ from the present in major ways. For example, there should be some very negative scenarios. People tend to view very negative scenarios as implausible and reject them out of hand. Nevertheless, they are crucial to an adaptive policy; having thought about a situation (no matter how implausible) in advance allows contingency plans to be formulated so that they are ready to be implemented in the (however unlikely) event they are needed. So, as many 'black swans' as possible should be identified in order to be prepared in case one of them actually occurs. In the airport case, demand for air transport is one of the key scenario variables. There could be a sharp decrease in demand, for example because of a financial crisis. This would make the policy fail. But there could be a sharp increase in demand, which could lead to unacceptable delays in take-offs and landings, which would also make the policy fail. We deal with this vulnerability in step 4.

  Another vulnerability of the basic policy is resistance from people living around the airport because of the noise from the anticipated additional flights. This vulnerability is fairly certain. So, at the same time as the new runway is agreed upon, it would be wise to offer financial compensation to residents in the high noise zone to enhance the chances of success of the basic policy.
- **Step 4: Setting up a monitoring system and preparing to adapt the policy**
  After the basic policy and anticipatory actions are implemented, there is still a need to monitor changes in the world and the performance of the

policy, and to take actions, if needed, to guarantee the policy's progress and success. Similar to the approach in step 3, scenarios or even EMA can be used to identify what to monitor, and when to trigger responsive actions, and the specific actions to take. In this step, actions that might be taken to guarantee the basic policy's progress and success are prepared. In addition, signposts are identified that specify information that should be tracked, and critical values of signpost variables (called 'triggers') are specified beyond which actions to change the policy should be implemented to ensure that the resulting policy keeps moving the system in the right direction and at a proper speed. The starting point for the identification of signposts is the set of vulnerabilities specified in step 3.

In the airport case, it is possible that the increases in demand are much greater than expected. This would lead to unacceptable delays, and airlines might decide to shift flights (or even their hubs) to other airports, which would lead to failure of the plan. In preparation, plans could be made to shift specific types of flights to surrounding airports (e.g. all-cargo flights or flights by low cost carriers). Making these plans would not be expensive, and they might never be needed. But, if the conditions warranted them, the plans would be there and could be implemented quickly at the appropriate time (specified by the trigger), thus saving the basic policy.

Although they are promising, adaptive policies have not yet become commonplace in public policy making. More research is required before this will happen. First, their validity and efficacy need to be established. Evidence is being gathered through a variety of methods, including gaming and computational experiments. The costs and benefits of dynamic adaptation measures compared to traditional policy making approaches also need to be studied. Finally, the implementation of dynamic adaptation will require significant institutional and governance changes, since some aspects of these policies are currently not supported by laws and regulations (e.g. the implementation of a policy triggered by an external event).

13.6 Conclusions

The most important conclusions of this chapter are:

1. Futures research often plays an important role in transport policy making. However, it is very important to note that the future is unknown, which makes future research outcomes (highly) uncertain, by definition.

Uncertainty in this chapter is defined as being any departure from the (unachievable) ideal of complete determinism.

2. That uncertainties exist in practically all long-term transport policy making situations is generally understood by most policy makers, as well as by most policy analysts. But there is little appreciation of the fact that there are many different dimensions of uncertainty, and there is a lack of understanding about their different characteristics, their relative magnitudes, and the available approaches and tools for dealing with them.
3. A much used approach in transport policy planning is the scenario approach.
4. An important advantage of using scenarios in futures research is that scenarios provide a way to explore the implications of deep uncertainty for policy making (prepare for the future) by identifying possible future problems and identifying potential policies for dealing with the problems.
5. An important disadvantage of the use of scenarios is that the scenario results are often used as 'certain' predictions, while they should be interpreted as 'what if' estimates for some plausible futures, and it is unknown (and unknowable) whether the actual future is covered by them.
6. In the backcasting method, a normative target in the future – a desired outcome – is chosen as the starting point for the futures analysis. Images of the future have to be designed that meet the specified targets. They should be clearly different from each other, in order to give an idea of the huge variety of possible futures, all of which meet the specified targets. In the backcasting method, (1) it is important to avoid using forecasted futures as certainties, and (2) it is incorrect to assume that specified policies will actually lead to the desired future.
7. Some scientists are now thinking about policies that take uncertainty into account. The key idea is not to specify an 'optimal' policy for a single best estimate future, but rather to design a policy that is flexible and adaptable.

REFERENCES

Agusdinata, D.B. (2008), *Exploratory Modeling and Analysis: A Promising Method to Deal with Deep Uncertainty*, Delft: Next Generation Infrastructures Foundation.

Agusdinata (2010), 'Exploratory MCDA for handling deep uncertainties: the case of intelligent speed adaptation implementation', *Journal of Multi-Criteria Decision Analysis*, 17 (1–2), 1–23.

Anguera, R. (2006), 'The Channel Tunnel: an ex post economic evaluation', *Transportation Research Part A*, 40, 291–315.

Banister, D., D. Stead and R. Hickman (2008), 'Looking over the horizon: visioning and backcasting', in A. Perrels, V. Himanen and M. Lee-Gosselin (eds), *Building Blocks for Sustainable Transport: Obstacles, Trends, Solutions*, Helsinki: VATT and Emerald, pp. 25–54.

Banister, D., D. Stead, P. Steen, J. Åkerman, K. Dreborg, P. Nijkamp and R. Schleicher-Tappeser (2000), *European Transport Policy and Sustainable Mobility*, London and New York: Spon Press.

Bankes, S. (1993), 'Exploratory modeling for policy analysis', *Operations Research*, 43 (3), 435–444.

Börjeson, L., M. Höjera, K.-H. Dreborg, T. Ekvall and G. Finnvedena (2006), 'Scenario types and techniques: towards a user's guide', *Futures*, 38 (7), 723–739.

Courtney, H. (2001), *20/20 Foresight: Crafting Strategy in an Uncertain World*, Boston, MA: Harvard Business School Press.

Findeisen, W. and E.S. Quade (1985), 'The methodology of systems analysis: an introduction and overview', in H.J. Miser and E.S. Quade (eds), *Handbook of Systems Analysis: Overview of Uses, Procedures, Applications, and Practice*, New York: Elsevier Science, pp. 117–149.

Flyvbjerg, B., N. Bruzelius and W. Rothengatter (2003), *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge: Cambridge University Press.

Flyvbjerg, B., M.K. Skamris Holm and S.L. Buhl (2006), 'Inaccuracy in traffic forecasts', *Transport Reviews*, 26 (1), 1–24.

Funtowicz, S.O. and J.R. Ravetz (1990), *Uncertainty and Quality in Science for Policy*, Dordrecht: Kluwer Academic Publishers.

Geurs, K.T. and G.P. van Wee (2000), *Environmentally Sustainable Transport: Implementation and Impacts for the Netherlands for 2030. Phase 3 Report of the OECD Project 'Environmentally Sustainable Transport'*, RIVM report 773002013, Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu.

Hickman, R. and D. Banister (2005), *Visioning and Backcasting for UK Transport Policy*, London: Department for Transport.

Hodges, J.S. (1991), 'Six (or so) things you can do with a bad model', *Operations Research*, 39, 355–365.

Hojer, M. (1998) 'Transport telematics in urban systems: a backcasting Delphi study', *Transportation Research D*, 3 (6), 445–463.

IEA (International Energy Agency)/OECD (Organisation for Economic Co-operation and Development) (2009), *Transport, Energy and CO_2: Moving toward Sustainability*, Paris: IEA/OECD.

Kwakkel, J.H., W.E. Walker and V.A.W.J. Marchau (2010a), 'Classifying and communicating uncertainties in model-based policy analysis', *International Journal of Technology, Policy and Management*, 10 (4), 299–315.

Kwakkel, J.H., W.E. Walker and V.A.W.J. Marchau (2010b), 'Assessing the efficacy of adaptive airport strategic planning: results from computational experiments', Environment and Planning B: Planning and Design, 39, 533–550.

Lempert, R.J., D.G. Groves, S.W. Popper and S.C. Bankes (2006), 'A general, analytic method for generating robust strategies and narrative scenarios', *Management Science*, 52 (4), 514–528.

Lempert, R.J., S.W. Popper and S.C. Bankes (2003), *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis*, MR-1626-RPC, Santa Monica, CA: RAND.

Leusink, A. and H.A. Zanting (2009), *Naar een afwegingskader voor een klimaatbestendig Nederland, met ervaringen uit 4 case studies: samenvatting voor bestuurders* [Towards a trade-off framework for climate-proofing the Netherlands, with experiences from 4 case studies: executive summary], accessed 4 August 2012 at http://edepot.wur.nl/15219.

Makridakis, S., R.M. Hogarth and A. Gaba (2009), 'Forecasting and uncertainty in the economic and business world', *International Journal of Forecasting*, 25, 794–812.

Marchau, V.A.W.J., W.E. Walker and G.P. van Wee (2010), 'Dynamic adaptive transport policies for handling deep uncertainty', *Technological Forecasting and Social Change*, 77 (6), 940–950.

Quist, J. (2007), *Backcasting for a Sustainable Future: The Impact after 10 Years*, Delft: Eburon Academic Publishers.

RAND Europe (1997), *Scenarios for Examining Civil Aviation Infrastructure Options in the Netherlands*, DRU-1513-VW/VROM/EZ, Santa Monica, CA: RAND.

Schippl, J. and I. Leisner (2009), 'Backcasting in transport scenarios for the future of European long-distance transport', *Technikfolgenabschätzung: Theorie und Praxis*, 2 (18), 63–69.

Schippl, J., I. Leisner, P. Kaspersen and A.K. Madsen (2008), *The Future of European Long-Distance Transport: Scenario Report*, Brussels: European Parliament.

Schwarz, B. (1988), 'Forecasting and scenarios', in H.J. Miser and E.S. Quade (eds), *Handbook of Systems Analysis: Craft Issues and Procedural Choices*, New York: Elsevier Science, pp. 327–367.

Schwartz, P. (1996), *The Art of the Long View: Paths to Strategic Insight for Yourself and Your Company*, New York: Currency Doubleday.

Taleb, N.N. (2007), *The Black Swan: The Impact of the Highly Improbable*, New York: Random House.

Thissen, W.A.H. (1999), 'A scenario approach for identification of research topics', in M.P.C. Weijnen and E.F. ten Heuvelhof (eds), *The Infrastructure Playing Field in 2030*: *Design and Management of Infrastructures*, Delft: Delft University Press, pp. 5–10.

Van Asselt, M.B.A. and J. Rotmans (2002), 'Uncertainty in integrated assessment modelling: from positivism to pluralism', *Climatic Change*, 54, 75–105.

Van der Heijden, K., R. Bradfield, G. Burt and G. Cairns (2002), *The Sixth Sense: Accelerating Organisational Learning with Scenarios*, Chichester: John Wiley.

Van der Pas, J.W.G.M., W.E. Walker, V.A.W.J. Marchau, G.P. van Wee and D.B. Agusdinata (2010), 'Exploratory MCDA for handling deep uncertainties: the case of intelligent speed adaptation implementation', *Journal of Multi-Criteria Decision Analysis*, 17, 1–23.

Van der Sluijs, J.P. (1997), 'Anchoring amid uncertainty: on the management of uncertainties in risk assessment of anthropogenic climate change', Ph.D. dissertation, University of Utrecht.

Van Wee, B. (2009), 'Transport policy: what can it and what can't it do?', European Transport Conference 2009, accessed 4 August 2012 at http://etcproceedings.org/paper/transport-policy-what-it-can-and-what-it-can-t-do.

Walker, W.E. (2000a), 'Policy analysis: a systematic approach to supporting policymaking in the public sector', *Journal of Multi-Criteria Decision Analysis*, 9 (1–3), 11–27.

Walker, W.E. (2000b), *Uncertainty: The Challenge for Policy Analysis in the 21st Century*, P-8051, Santa Monica, CA: RAND.

Walker, W.E., J. Cave and S.A. Rahman (2001), 'Adaptive policies, policy analysis, and policymaking', *European Journal of Operational Research*, 128 (2), 282–289.

Walker, W.E., P. Harremoës, J. Rotmans, J.P. van der Sluijs, M.B.A. van Asselt, P. Janssen and M.P. Krayer von Krauss (2003), 'Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support', *Integrated Assessment*, 4 (1), 5–17.

Wright, G. and P. Goodwin (2009), 'Decision making and planning under low levels of predictability: enhancing the scenario method', *International Journal of Forecasting*, 25, 813–825.

14

Appraisal methods for transport policy

Piet Rietveld

14.1 Introduction

Policy makers in the field of transport face five fundamental uncertainties:

1. uncertainties on the effects of policy measures, for example the effects of a shift in automobile taxation, from taxes on car ownership to fuel taxes;
2. uncertainties on autonomous developments such as the level of economic growth;
3. uncertainties on the transport policies of other main actors, such as the construction of airport infrastructures in neighbouring countries;
4. uncertainties on how to trade off the various effects of the measures, such as environmental effects versus employment effects;
5. uncertainty on social acceptance of the measures, such as the acceptance of intelligent speed adaptors by the general public.

Most of these uncertainties have been addressed in the other chapters in this volume (see Chapter 13 for a discussion on uncertainty). This chapter focuses on the type 4 uncertainty: how to trade off the various effects of the construction of a road on road users (less congestion), firms (improved accessibility), residents (increased noise) and taxpayers.

An important question is: to what extent are policy analytical tools needed to support the authorities that are responsible for decision making? One position is that analysts have the task to forecast the various effects of policy alternatives. Then in a second step the policy makers responsible carry out the various trade-offs involved and select the best alternative (for example, Lichfield, 1988). In this case there is a clear division of tasks between policy makers and analysts. An example of this approach is environmental impact assessment.

In this chapter we consider the case where analysts not only predict the effects of policy alternatives, but also are involved in supporting decision makers in the further process of decision making. We will focus on two particular alternatives: social cost benefit analysis (sections 14.2–14.5) and multi-criteria analysis (section 14.6). Section 14.7 summarizes the main conclusions of this chapter.

14.2 Cost benefit analysis

Social cost benefit analysis (SCBA) has become a major tool for policy analysis in many countries (see, for example, SACTRA, 1999; Berechman, 2009; see also Chapter 12). The aim of SCBA is to derive a summary indicator of the costs and benefits for *all* the actors affected. The term 'social' is used to indicate that the interests of all groups are incorporated. This is one of the main differences from the notion of a 'business case', which exclusively focuses on the interests of one particular actor. The basic notion behind SCBA is that the preferences of consumers form the basis for trade-offs in transport policies. When markets function well, prices of goods and services will provide the necessary information about these preferences. However, in most real world cases, some of the relevant markets may fail and some prices may not be readily available. This holds true for all external effects, such as the environmental effects of transport (see Chapter 12). Another important element in many transport policies is the length of travel time. Most policies will affect the time spent in travelling, but it is not immediately obvious how to value changes in travel time. One of the challenges of transportation research for the past few decades has been to provide estimates of these elements of benefits or costs for transport policies.

With the choice of consumer preferences as the starting point for SCBA, the consumers' willingness to pay is an important element of the valuation of costs and benefits (Small and Verhoef, 2007). The consumers' preference can be described by means of a demand function (see Figure 14.1). A demand function describes how much a group of consumers is prepared to pay for a certain product or service. For example, in Figure 14.1, when p_0 is the price, N_0 consumers will use the service provided. When the price decreases to p_1 the number of consumers increases to N_1. In the context of transport the 'transport service' may mean: making a bus trip from A to B, or making use of the road to drive from C to D. Consider now the group of N_0 consumers who make use of a certain transport service who are prepared to pay the price p_0 for this service. Note that their willingness to pay is *at least* p_0: most of these consumers are prepared to pay more than just p_0. The difference between what a consumer actually pays and what he would be prepared to pay is called

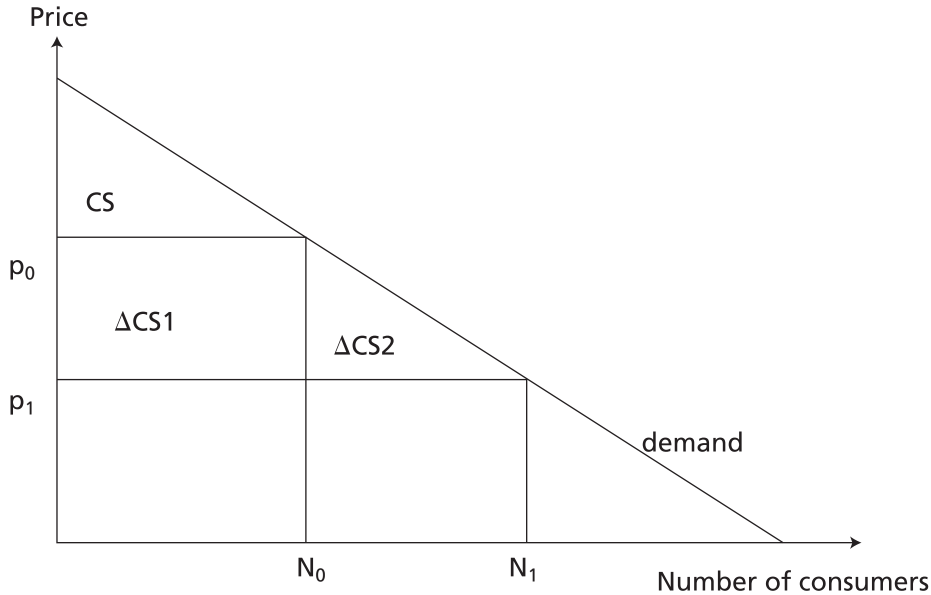


Figure 14.1 Illustration of the consumer surplus (CS) concept

the 'consumer surplus'. In Figure 14.1 the consumer surplus of all users of the service when the price is p_0 is represented by the triangle CS.

The term consumer surplus can be useful in discussions on 'the economic importance of transport'. The total amount of money actually paid by consumers ($N_0 \times p_0$) provides an underestimate of what these consumers might be willing to pay for this service. Note that in Figure 14.1 when the price is p_0 the consumer surplus is considerably smaller than the actual expenditure for this transport service, but for lower prices this no longer applies.

Consider now an improvement of a certain connection so that the costs of using it decrease from p_0 to p_1. For the present users of the transport service the consumer surplus changes from $N_0 \times p_0$ to $N_0 \times p_1$, which implies an increase of $N_0 \times (p_0 - p_1)$. Another consequence of the price decrease is that it generates new users $N_1 - N_0$. The change in consumer surplus of this group is represented by the triangle $\Delta CS1$ in Figure 14.1. This change is equal to the increase in the number of users $N_1 - N_0$ multiplied by half of the price decrease $p_0 - p_1$: $\Delta CS1 = 0.5 \times (N_1 - N_0) \times (p_0 - p_1)$. This is the so-called 'rule of half'. It means that the benefits of a price decrease for new consumers are 50 per cent of the benefits of incumbent consumers. That the result is 50 per cent can be made intuitively clear by first considering a new consumer with a willingness to pay very close to p_0. This consumer benefits almost fully from the price decrease. But a consumer with a willingness to pay close to p_1 hardly benefits from the price decrease. Then the average new consumer's benefit will be 50 per cent of the benefit of incumbent consumers. The rule of half is exact when the demand function is linear, but in real world cases the demand function is probably not linear. In that case the rule of half is only an approximation. When price changes are modest, the error from using the rule of half will be small. Note that the total consumer surplus will always increase when the price of a service decreases. This does not necessarily hold true for the total expenditures: these might decrease when prices go down, that is, when the relative increase in users is less than the relative decrease in the price.

The notion of consumer surplus can of course also be used in the case of price increases. The analysis is entirely symmetric. Note that in this case the disadvantage of a price decrease is largest for the consumers who continue to use the service and smaller for the consumers who stop using the service because they apparently have another alternative (another transport mode, another destination, staying at home) which is sufficiently attractive for them to decide to change their behaviour. Thus the economic approach implies that people who do not change their behaviour as a consequence of a price increase are hit harder than people who do change their behaviour. It is worth noting that this result may be rather different from a political economy perspective, where consumers who do change their behaviour are considered to be the ones hurt most by a price increase. These consumers cannot afford to pay the high price and hence should be considered as the greatest victims. This may well lead to a gap between cost benefit analysis where consumers who change their behaviour get a low weighting and political debates where these consumers get a high weighting.[1] This is an illustration of the difficulties that may emerge when SCBA is applied in situations where equity concerns are high. We will address this subject later in this chapter.

It may be worth noting that the group of potential consumers who are willing to pay less than p_1 do not play a role in the welfare calculations presented above, simply because they don't take part in the actual use of the service as long as the price does not decrease further. In some cases, this group may also be of some relevance. An example is that people may not use public transport but may nevertheless be prepared to pay a certain amount of money to ensure that they have an option to make use of public transport under particular circumstances. For this purpose the term 'option value' is sometimes used. For example, people may be prepared to pay a certain amount of money to live in a municipality with a railway station compared with a municipality without a railway station. See Geurs et al. (2006) for estimates of option values.

To be able to assess the benefits of a change in transport services leading to a decrease in the price, information is needed on the price in the reference situation and the new situation (see Chapter 13) and the number of consumers in both cases. The reference situation may be the current situation, but in many practical applications of SCBA the alternatives refer to infrastructure projects that may take a long time before they are completed. Hence predictions are needed of prices and numbers of users in the future, with and without the policy alternative. For this purpose one needs projections on how the economy and the transport system will look in the future and also on how transport demand will be affected by a price change. This involves the need for information on price elasticities, as mentioned in Chapters 3 and 6. This also means that SCBA, being a method to value benefits and costs of policies (uncertainty type 4 mentioned in the introduction to this chapter), is closely linked to uncertainties of types 1 and 2.

The final step of SCBA is that an overall assessment of alternatives is carried out by comparing the alternatives with a reference alternative as mentioned above. The changes in costs and benefits for all actors are determined for each alternative (compared with the reference alternative) and then the net balance in the change in costs and benefits can be computed. The alternative with the highest positive net balance is the best candidate to be implemented according to SCBA. Since the results of SCBA depend on many uncertain inputs, sensitivity analysis is recommended. For an example of the results of SCBA, see Table 14.1.[2] These values are expressed in terms of net present values. This means that the effects that take place from year to year are added in such a way that they are computed as the aggregate value for the first year. This is done by means of weighting factors based on an interest rate (discount rate) which represents the social time preference. In most countries this discount rate is about 3 or 4 per cent per year. The consequence of discounting future effects at such levels is that effects that take place in say 30 years after completion of the project get a rather low weighting, and effects that will take place in 100 years are negligible. Note that by discounting one provides a common ground for effects that take place every year and effects that take place only once (such as the construction costs).

There are various ways to use the outcomes of a cost benefit analysis. One is to focus on the difference between benefits and costs (B–C); an alternative is to consider the ratio of benefits and costs (B/C). When there is only one alternative as well as the reference alternative, it does not matter which of the two is used: the essence is that a project is recommended when B–C is positive or when B/C is larger than 1 and the two criteria are consistent. However, when there are more alternatives, the ranking of the alternatives

Table 14.1 Example of an SCBA table for a road upgrade involving separation of motorized and non-motorized transport

	Valuation
Benefits:	
Decrease of travel time for present road users	200
Welfare improvement of new road users	60
Improvement in traffic safety	20
Costs:	
Environmental costs (higher emissions because of higher speeds and greater traffic volumes)	15
Extra noise nuisance	20
Construction costs	160
Additional maintenance costs	10
Deterioration of landscape	PM (pro memory)
Benefits – costs	75 – PM

Note: Figures in net present values in million euro.

according to the B–C criterion may be different from that according to the B/C criterion. For example, some alternatives are 'smaller' in scale than other alternatives, and that would mean that alternatives with a comparable B/C result may have very different B–C outcomes. A disadvantage of the B/C ratio is that it may be somewhat sensitive to manipulation, because costs might be classified as negative benefits and that would affect the B/C ratio. An alternative criterion sometimes used is B–C divided by the level of funds available, which may be useful when one wants to make use of a clearly constrained fund in the best possible way.

Finally one can also consider the return on investment resulting from costs and benefits. The return on investment is the rate of return that can be computed on the basis of the monetary values of costs and benefits in each year. This rate of return can then be compared with the social rate of discount to be used for projects: when the rate of return on investment is higher than the social rate of discount this means that the project can be recommended. This is equivalent to the result that B–C is positive. The advantage of the return on investment is that it can be used to compare alternatives of very different size so that the scale problem mentioned above can be overcome.

14.3 Issues in SCBA

The use of SCBA as a decision support tool for the public sector is not without problems. One of the first issues is the equity problem: that the distribution of costs and benefits may be very uneven so that a dichotomy of winners and losers may emerge (see also Chapter 12). For example, users of the road presented in Table 14.1 experience safer and faster traffic, but residents around the road may not like the noise. The final outcome of SCBA is the net result of winners and losers so that the distribution is not directly visible. The main idea behind SCBA is that the gains should be high enough so that the winners can compensate the losers (this is the so-called Hicks–Kaldor principle) and still be better off in the end. In that case everybody would be better off (this is known as the Pareto principle). However, the above only involves a hypothetical possibility to compensate, and there is no guarantee at all that an actual compensation will take place. Hence almost every policy alternative will lead to the situation that there is a group of losers who may see a reason to protest against the proposed plan.

Thus the strict use of the hypothetical compensation notion in SCBA may lead to the situation that alternatives are selected with a very unequal distribution of costs and benefits. Suppose now that there is another alternative with a somewhat lower net outcome but with a more balanced distribution of costs and benefits. Taking equity considerations into account, policy makers might prefer the second alternative above the one with the higher benefit–cost result. A possible way to address this point is to apply equity weights in order to give a stronger weight to some of the actors (see also Rietveld, 2003). Note the link with the discussion above on who are most disadvantaged by a more expensive service: those who continue to use the service or those who stop using it. When the reason for stopping using it is that these are users with low incomes so that they cannot afford the more expensive service, this might give them a higher weight for equity reasons.

There is also a clear link between the equity notion and social acceptance (see section 14.1, uncertainty type 5). Further, note that, when certain transport policies are proposed for equity reasons, such as the provision of special transport services for older or handicapped people, a standard cost benefit analysis is not the proper instrument because it addresses overall efficiency rather than the equity concerns of certain disadvantaged groups. In that case other evaluation methods like multi-criteria analysis (MCA) or cost effectiveness analysis would be more suitable tools.

A second problem with SCBA is that there are no monetary valuations available for certain policy effects. This may be the case with some environmental effects and effects on nature, landscape, safety and health. When such estimates are not available, one can still express in an SCBA table that there is an effect by adding a PM (pro memory) indicator, as done in Table 14.1. This then leaves the challenge to trade off the PM effect with the monetary effects to the decision maker.

A third potential problem with SCBA is that there are broader effects of infrastructure than what one can measure via the use of transport links. Some analysts and policy makers claim that this leads to an underestimate of the economic benefits of transport infrastructure. This is a rather complex issue. It is indeed correct that changes in transport infrastructure may have long run effects on locational patterns of firms and households. When an SCBA is based on a correct transport model these long run adjustments have already been accommodated in the forecast of travel demand. It is clear that it is not easy to give the long run predictions of the effects of infrastructure changes, so part of the discussion on the broader, economy-wide effects of transport infrastructure is not so much a discussion on SCBA itself, but on the underlying models. An important result of economic analysis is that, when markets are not distorted, the total effects of an infrastructure improvement in a network can be measured via the flows on the network. The improvement may also express itself by higher land prices in some areas or by lower prices of certain goods, the price of which depends on transport costs, but it can be shown that there is an immediate equivalence of benefits on the transport markets and the benefits in other markets (SACTRA, 1999; Small and Verhoef, 2007). Thus, adding benefits in the markets of consumer goods to benefits on the transport market would lead to double counting. In the real world, market imperfections mean that this easy result cannot always be used. Market imperfections may imply, for example, that changes in the costs of using transport networks are not fully transferred by producers or traders to the consumers. In that case a gap may emerge between transport benefits and benefits measured on other markets. This is a field where research has not yet led to a final conclusion in terms of how to deal with it. Based on a limited number of studies it appears that in most cases market imperfections lead to a top-up of transport related benefits up to 20 per cent (see for example Zhu et al., 2009).

Problems may also arise when part of the beneficiaries of a transport investment are outside the country where the investment takes place. The usual approach is that a national government that invests in an infrastructure project only considers the costs and benefits that accrue to citizens and firms within their own country. The rationale behind this is that the government has been elected to foster the interests of its own citizens, and also that the project is paid for by the taxpayers in the country. The problem with this approach is that positive spill-over effects to other countries are not considered. Such spill-overs mean that residents of other countries benefit from

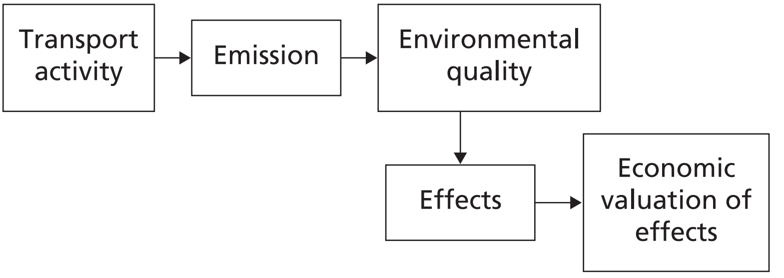


Figure 14.2 Chain from transport to economic valuation of emission

the improvements of infrastructure in another country when they visit it. These benefits may be substantial in the case of investments in international links. This is probably one of the reasons why international links are underdeveloped in infrastructure networks. There are three ways to address this issue in order to overcome a bias towards domestic projects in SCBA. The first one is that a sensitivity analysis is given of a certain project: one with domestic effects only and one where welfare effects on foreign actors are also presented. The second is an extension of this approach: an international project is defined as a joint project of two countries where costs and benefits are added in both countries. The third way is that a supranational party like the European Union becomes involved and provides a subsidy to account for the positive externality in the welfare effects.

14.4 Valuation of environmental effects

In SCBA monetary values have to be obtained for the environmental effects of transport policies. Such a valuation is often not simple since it addresses the final results of a whole chain of effects from transport via emissions to impacts on environmental quality (see Figure 14.2). This chain will vary according to type of environmental effect (see also Chapter 10). In the case of noise, the effect on environmental quality will be restricted to the local environment. The effects concerned are, among others, annoyance, disturbance of sleep, stress and heart diseases for those living or working close by. In the case of SO_2 emissions the spread is much broader and the effects will be in terms of damage to health and ecosystems, among other things. The valuation problems of environmental effects of transport are in essence no different from those of other economic activities. According to the chain described in Figure 14.2 it does not matter what the origin of the emission is. A difference is, however, that the distance between the source of the emissions and the people being exposed is often shorter than with other sources like power plants and the manufacturing industry. Therefore the damage resulting from a kilogram of NO_x emissions in transport will on average be higher than the damage from the same amount emitted by a power plant (Janssen et al., 1997) (see Chapter 10).

There are various methods for arriving at monetary values for the environmental effects of transport. For a survey we refer to Johansson (1991) and Perman et al. (2003). Two different approaches can be distinguished. The first approach focuses on environmental effects, the values of which can be observed via markets. The second approach addresses the subjective expressions of consumers on the values of environmental effects. We will discuss two examples of both approaches in some detail.

An example of the valuation of environmental effects via actual behaviour is that houses in noisy areas will most probably have a lower value. Via a so-called hedonic price approach one can isolate the effect of noise from the effects of many other factors that have an impact on the value of dwellings. The outcome of a hedonic price analysis is the percentage change in the value of a house due to exposure to a certain noise level when all other features of a house are kept constant. This econometric approach is often applied to the valuation of noise near airports (see, for example, Schipper, 2001; Dekkers and van der Straaten, 2009). Note that consumers will be heterogeneous with respect to the noise annoyance: people who strongly prefer a quiet environment will usually not live in very noisy places, since people who give less priority to a quiet place will bid higher prices. Thus the phenomenon of self-selection will reduce the burden of noise (Nijland et al., 2003). A sudden increase in noise levels due to the opening of a new runway, for example, will then affect people with a strong dislike of noise. In the long run a new equilibrium will emerge via spatial sorting of households. A basic assumption behind hedonic price analyses is that the housing market is in equilibrium. Thus this may lead to an underestimate of the effects in the short run when a sudden unanticipated increase in noise levels occurs.

A related method that is close to actual market oriented behaviour concerns the costs that people incur to visit nature areas. These costs appear from the travel costs they incur and the entrance fee, if there is one, among others. This so-called travel cost method (see Perman et al., 2003) can be used to determine the use value of nature areas. This method also has its limitations, since it is not so clear how to value the travel time of these leisure trips. In addition, the travel cost method does not address so-called non-use values of a nature area. An implication is that, when there are two identical nature areas, one in a region with high population density and the other one in a sparsely populated area, the former one will attract more visitors who express their willingness to pay via their travel cost than the latter. So the use value of the former will be larger than that of the latter.

The second approach in the field of valuations is to base them on the subjective valuations of consumers. This approach is particularly recommended in the case of non-use values. A well-known example of this approach is the so-called conjoint analysis method. People are asked to consider a certain hypothetical choice situation with two or more alternatives and they are then asked to indicate which of them they like most. An example may be that respondents are asked to consider alternatives to reduce the number of animals killed by transport. One alternative could be a situation without specific measures; another alternative may then be a case with specific measures such as ecoducts and other protective facilities for animals, accompanied by a certain financial contribution. The second method (contingent valuation) is related to conjoint analysis (see Hoevenagel, 1994). It involves asking respondents how much they are prepared to pay for measures to avoid animals being killed by traffic. A more detailed account of these methods is given in section 14.5.

The disadvantage of both conjoint analysis and contingent valuation is that they address hypothetical situations. In reality car drivers are not asked for a financial contribution, and it is not guaranteed that drivers will really behave according to their statements. Another issue is that the responses will be influenced by the information that respondents have on the subject, possibly provided by the interviewers. One of the risks of research in this area is that respondents will behave like free riders, which means that they assume that other people will pay for the environmental problems. This would lead to an underestimate of their valuation. Alternatively, respondents may respond with a warm glow attitude implying that they suggest a high willingness to pay to overcome certain environmental problems, but their actual willingness may in the end be much lower. Context is also important: respondents may express a high willingness to pay to save a specific vulnerable nature area, but, when they realize that there are many similar areas, their valuation may decrease. A general finding in literature is that people's responses may also be asymmetric: their willingness to pay for a noise decrease from 65 to 60 dB(A) may be smaller than their willingness to accept a noise increase in the opposite direction. This implies a non-linear valuation of environmental quality, where the present level may function as an anchor point. Apparently, respondents find it difficult to adjust consumptive expenditures when they would have to pay more for a clean environment. In general the willingness to pay is considered to be more appropriate than the willingness to accept (Hanemann, 1991).

Table 14.2 Assessment of evaluation methods for environmental damage

Method	Scientific basis	Information needed	Reliability	Applicable to
Travel cost method	Good	Observed choice behaviour	Good	Use value of nature and landscape
Hedonic price method	Good	Outcomes of market processes	Good	Noise nuisance, local emissions, safety
Stated preference	Reasonable	Hypothetical choices	Reasonable	Broad range of applications; also for global effects
Contingent valuation	Reasonable	Hypothetical choices	Difficult to verify	Broad range of applications; also for global effects
Shadow price method	Dubious	Prevention costs	Good	Broad range of applications; also for global effects

A third method to determine prices when there is no market is the so-called shadow price method. This method indicates how much it would cost to prevent a certain target being exceeded. For example, how much would it cost to reduce CO_2 emissions such that they comply with international agreements? This method is rather different from the methods discussed above because it is formulated in terms of prevention costs, so that the damage cost is not estimated. The shadow price method can be used in cases where no damage cost estimates are available, in particular when a clear target has been formulated in the policy domain.[3] A concise comparison of the methods discussed here can be found in Table 14.2.

14.5 Valuation of travel time

The value of travel time has already been addressed briefly in Chapter 3 of this book. In this chapter we give a more in-depth discussion. The valuation of travel time plays a very large role in the evaluation of many policy alternatives in transport. The main effect of much public sector investment is that traffic times are reduced. This holds for both passenger and freight transport. It is important to note that travel time gains owing to transport infrastructure do not express themselves fully in increases in GDP. Some of the gains take place in the business sector, where faster transport indeed implies lower costs and greater opportunities to exploit economies of scale in production, leading to higher GDP. But most of the gains will be realized outside the business sector. The effect of time savings is that people get more leisure time or that

Table 14.3 Travel time valuation for various income categories and trip purposes, the Netherlands (2008)

Net income (euros/month)	Commuting trips (euros/hour)	Business trips (euros/hour)	Other trips (euros/hour)
<1365	0.07	11.05	4.83
1365–2275	7.00	17.21	5.39
2275–3410	7.56	22.18	5.95
>3410	12.59	46.52	8.11

they use the time gains by making trips to more attractive destinations located further away. This is indeed what is found in the literature on the constancy of travel times: as also indicated in Chapter 3, increased speeds lead to changes in travel patterns such that the total travel time remains fairly constant at an aggregate level (see, for example, Zahavi, 1979, and van Wee et al., 2006; see Chapter 6). The fact that most of the travel time gains are not incorporated in GDP reflects the rather limited scope of GDP, which incorporates the value of the goods and services produced, but not the leisure time of consumers.

An important question is how to value travel time gains. Two approaches can be followed that are similar to the first two approaches adopted in the estimation of the values of environmental effects discussed in section 14.4. The first approach considers actual travel behaviour; the second one addresses travel behaviour in hypothetical choice situations. The first approach (revealed preference) is useful when consumers face choice situations where they trade off travel time and monetary expenditure. In aviation, this may, for example, be the choice between a direct flight and a cheaper indirect flight. Or it may be fuelling in one's own municipality versus a cross-border fuelling trip where fuel is cheaper. The second approach (stated preference) means that travellers are confronted with hypothetical choice situations between a fast, expensive alternative and a cheap one.

An example of the results of studies on the value of travel time savings can be found in Table 14.3. The values presented in euros per hour are the result of estimations on the basis of stated preference interviews. The table shows that the value of travel time varies considerably between travel purposes. People value travel time more highly when they travel for business compared with when they travel for other purposes. It also appears that people with high incomes have a higher value of time compared with people with low incomes. This makes sense, since they have more budget to 'buy' fast transport, and the time they save might be used to increase their income even more.

The range of values found in Table 14.3 is sometimes even larger; for example, in the case of aviation or high speed rail one may find values for the value of time like 75 euros per hour in the case of business travellers. The literature shows that travel time gains are valued lower than travel time losses, so the willingness to pay for a travel time gain of five minutes is lower than the willingness to accept a loss of five minutes. The difference between both is that, with a longer travel time, a rescheduling of activities has to be carried out by shortening a certain activity. In the case of a shorter travel time, this problem does not exist. Of course, the mirror image of this result is that in the latter case the money available for other activities becomes less, but since money is so fluid and can be saved, it appears to be easier to make a cut in monetary expenditures in some activities than to make a cut in the time budget of a certain activity prior to or after a trip. This result is similar to the findings on the willingness to pay and the willingness to accept in the case of environmental damage.[4]

The result that travel time gains are valued lower than travel time losses suggests a non-constant valuation of travel time changes (see Blayac and Causse, 2001; Redmond and Mokhtarian, 2001). The valuation of a five minute increase in travel time is therefore not necessarily equal to 50 per cent of the value of a travel time increase of 10 minutes. When the differences in travel time between the various alternatives are bigger, the need to consider varying valuations of changes in travel time according to their length becomes more prominent.

This finding is consistent with a more general result in valuation studies. The models and the valuation methods used are often suitable for the analysis of limited changes in attributes (for example, a five eurocent increase in fuel prices). In the case of large changes like a dramatic price increase due to fuel shortage or an investment in high speed rail compared to conventional rail leading to a strong reduction in travel time, the estimation models do not sufficiently address these large changes. This holds true both for the estimation of the effects (uncertainty type 1 in the first section of this chapter) and for the valuation of the effects.

Another finding from the research on travel time valuation is that travel time should be distinguished in various (partly overlapping) components, such as in-vehicle time, per hour of vehicle time, transfer time and waiting time on platforms (Wardman, 1998; see Chapter 6). For car trips the distinction between time in congested and non-congested traffic is important. Usually the in-vehicle time is valued lower than the out-of-vehicle time, which has important implications. For example, for public transport operators it means

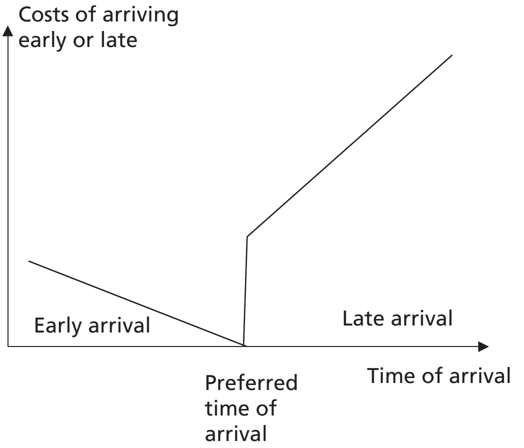


Figure 14.3 Structure of scheduling costs in transport

that less waiting time at platforms is valued higher than fast transport, as such. And for car users it means that they may care more about the possibility of parking near their destination than about the speed of their trip.

There is still another travel time component: the so-called scheduling costs (Small and Verhoef, 2007). This means that, when someone wants to arrive at a certain point of time T, while transport services with low frequencies force him to arrive early or late, this traveller experiences an extra disadvantage. The elements of these costs are illustrated in Figure 14.3. Arriving early (left from the intended time T) has the disadvantage that the time between the actual arrival and T cannot be used in the best possible way, as one would do at home. Arriving late, that is, after T, will most probably have larger disadvantages. In some cases the penalty can be very large, as when one travels to an airport with the risk of missing a flight or when the trip destination is a job interview. But even in more day-to-day activities like going to school or work, the penalty of arriving late may be substantial.

In addition to scheduling costs, dimensions such as comfort, reliability and information also play a role in the valuation of travel time. From stated preference research it appears for example that travellers assign high importance to the problem of travel time reliability (Rietveld et al., 2001). Furthermore, information provision helps to reduce the annoyance of travellers when they have to wait because of disturbances in travel. The possibility of using travel time for alternative purposes is also important. A clear advantage of travelling by high speed train compared to air transport is that a substantial part of a train trip can be used in a rather useful way, which is not the case with air

Table 14.4 Example of an evaluation matrix for MCA

	Criterion 1	Criterion 2	Criterion 3
Reference alternative	100	0.70	+
Alternative 1	300	0.90	+++
Alternative 2	250	1.30	+
Alternative 3	125	0.40	–

transport, where the actual part of the trip that one is seated in the aircraft is limited for short distance trips. Hence the introduction of high speed rail would be beneficial for business travellers even if it did not lead to a reduction of total travel time. The increasing possibilities of using travel time in a productive way owing to developments in information and communication technology also deserve attention (Lyons et al., 2007).

14.6 Multi-criteria analysis

In MCA, policy alternatives are evaluated according to the effects of a number of criteria. These criteria will usually be closely related to the welfare positions of the actors considered in SCBA. Criteria are often classified according to clusters, such as the contribution to accessibility, employment effects, effects on safety and the environment, construction costs and so on. The effects of policy alternatives are represented in an evaluation table. An example of such a table is given as Table 14.4, showing the case of three criteria and four alternatives. This example makes it clear that criteria may be qualitative instead of quantitative (criterion 3).

A second ingredient of MCA is that the importance of criteria is indicated. This may be done by using quantitative or qualitative weights. These weights may be explicitly expressed by policy makers, but they can also be derived by researchers based on decisions that have been made in the past. Applying the weights to the outcomes of the evaluation matrix leads to a certain ranking of alternatives. A wide variety of methods exists to arrive at these rankings. A survey can be found in Nijkamp et al. (1990) and De Brucker et al. (2004), among others. Sensitivity analyses are an important ingredient of MCA applications. Some of the MCA methods also involve the formulation of minimum or maximum requirements for criteria. An example of an evaluation matrix in MCA is given in Table 14.5.

An important difference between MCA and SCBA is that in the case of SCBA the weights depend on the preferences of consumers, whereas in the

Table 14.5 Example of MCA on the construction of a railway line in urban areas to reduce congestion in the road network

Criterion	Reference alternative (unchanged policy)	Construction railway line 1	Construction railway line 2
Total travel time per day	2.0 million hours/day	1.9 million hours/day	1.8 million hours/day
Safety (expected number of casualties in transport system)	500 casualties per year	492 casualties per year	490 casualties per year
Number of dwellings above a certain noise level	40 000 dwellings	39 800 dwellings	40 500 dwellings
Construction costs	0	EUR 1.3 billion	EUR 2 billion
Compatibility with physical planning	0	++	+
Effects on landscape	0	–	–
Possibility of developing new locations for residential development	0	++	+

case of MCA the weights express political priorities. SCBA may therefore be interpreted as a market oriented approach of policy makers, where it is important to note that the outcome has been corrected for market imperfections. MCA is oriented towards the priorities of policy makers or various actors. For example, MCA can be used to arrive at a ranking of alternatives from the perspective of each individual actor so that the sensitivity of the overall outcome can be analysed. For example, it may well be that a certain alternative never achieves the first rank for any of the actors considered but it may nevertheless be a good number two for most of them. In that case this particular alternative might be a good compromise alternative.

The political nature of MCA implies that a policy maker who has fallen in love with a certain alternative may try to legitimize that preference by choosing specific values for the weights involved. This may make MCA to some extent vulnerable to abuse, but on the other hand it can still be claimed that MCA is transparent, since weights have to be made explicit so that they can become part of the political debate. A property shared by both MCA and SCBA is that they involve a systematic presentation of alternatives and their consequences which potentially improves the quality of the political decision making processes. In the case of MCA, participants in the political process are challenged to express their priorities in an explicit manner. In the case of SCBA, this element is missing, but participants in the political process are challenged to indicate to what extent SCBA has overlooked relevant effects and to what extent they feel that equity considerations should lead to the choice of an alternative that did not achieve the highest ranking.

The advantage of SCBA is that it helps to avoid an 'implausible' fixation of political weights. By 'implausible' I mean that policy makers might choose weights that strongly differ from the preferences of citizens. For example, a certain weight associated with noise nuisance used by the government might strongly deviate from the value associated with noise by citizens as can be observed by their own efforts at insulation against noise nuisance. Depending on the level of the weight attached to noise, this may lead to both underinvestment and overinvestment in noise prevention measures by public authorities. In domains like this it is difficult to find good reasons why the government would value noise differently from the valuation given by citizens. In other domains, however, governments may have good reasons for departing from citizens' valuations, for example when citizens might have myopic valuations of safety. In such a case governments may take responsibility by following their own valuation schemes.

Another advantage of SCBA compared with MCA is that double counting can be avoided. Consider as an example the construction of a railway line leading to a reduction in the cost of importing goods. Various parties benefit from this investment: the railway companies, the firms processing the goods transported, the clients of these firms and finally the consumer. As explained above, in the case of a perfectly functioning market, these effects are essentially various aspects of the same thing. In SCBA this double counting is well addressed; in MCA, methods to avoid double counting are less developed. Table 14.6 summarizes the differences and similarities between SCBA and MCA.

The government objectives in the field of transport are a useful framework for comparing the usefulness of MCA and SCBA. These objectives usually relate to four domains: the economy, the environment, traffic safety and social equity (see also Chapter 12). SCBA covers economic aspects and the environment as far as this can be valued in economic terms. The contribution of SCBA with respect to social equity is less evident. One of the main reasons why governments subsidize public transport is that this is a basic facility that should be available to people who do not have sufficient travel alternatives. In this field, SCBA has less to offer, although of course economic analysis can still be useful for analysing which approach is most cost effective to reach certain equity objectives. Because MCA is less demanding in terms of eco-

Table 14.6 Comparison of SCBA and MCA

	SCBA	MCA
Systematic comparison of alternatives	Yes	Yes
Explicit formulation of weights in trade-offs	Yes	Yes
Basis for weights of various effects	Valuation by consumer	Political valuation
Opportunities for abuse by policy makers	By manipulation of inputs	By manipulation of inputs and by manipulation of weights
Degree of compensation between various attributes of alternatives	Every unfavourable attribute can in principle be compensated by a favourable outcome for another attribute	Various degrees of compensation are possible through the possibility of incorporating minimum requirements
Risk of double counting	Limited[1)]	Yes
Opportunities to take into account attributes that cannot be valued in monetary terms	No	Yes
Possibility of attaching weights to the interests of specific actors	Not in the standard form of SCBA	Yes

Note: [1)] It is often stated that the risk of double counting is not present in the case of SCBA. This holds true for the case of perfect markets. When there is a possibly imperfect market, indirect effects are incorporated and there is some risk of double counting.

nomic valuation of environmental effects, and because equity considerations can be given due attention in MCA, it is a useful tool for problems where these dimensions dominate.

An interesting possibility is to use SCBA and MCA in combination: first, SCBA is used for those effects where economic valuations are available, whereas MCA is used for the remaining aspects, where the result of SCBA is an important input of MCA. Examples can be found in van Pelt (1993) and De Brucker et al. (2004).

An important question is: to what extent does SCBA or MCA have an impact on the final outcomes of political decision making? The effects are rather mixed. For example, Fridstrom and Elvik (1997) find for Norway that policy alternatives that perform excellently according to cost benefit analysis may nevertheless not be selected in reality. A possible reason is that voting in parliament is strongly based on regional loyalties. When rural constituencies are better represented than urban ones, this may well lead to a bias against urban projects; note that it is in urban regions that congestion problems are usually most severe, which contributes to favourable outcomes of cost benefit analysis for urban projects. Cadot et al. (2007) arrive at similar results for France: they find that public choice related factors such as the political colour of the region, the congruence between national and regional political colour and the presence of lobby groups indeed play a significant role in infrastructure decision making. In the Netherlands a different outcome can be observed. During the last decade social cost benefit analysis has been imposed as a mandatory tool in policy analysis for large infrastructure projects. Annema et al. (2007) conclude that the use of SCBA has, in a considerable number of the cases, led to a change of original policy intentions, for example by postponing or downsizing projects.

SCBA and MCA are carried out in a context where analysis is very close to politics. The implication is that policy makers, politicians and civil servants may try to influence work on SCBA and MCA in such a way that desired outcomes are produced. It would mean that these tools are not used to critically examine the pros and cons of alternatives but to justify the choices policy makers intend to take. There is some literature giving evidence that systematic misrepresentation of policy effects indeed takes place (Pickrell, 1989; Flyvberg et al., 2003; Button, 2010): costs are often underestimated and demand is often overestimated. Such tendencies would of course reduce the usefulness of SCBA and MCA. There are at least two ways to reduce the risk of abuse of these tools. First, it is important that an independent institution tests the quality of an SCBA or MCA study and the underlying modelling work. Second, it helps to develop protocols containing a description of how the various steps in these decision support tools can be best carried out, and where reference values are provided for key parameters. An example at the EU level is the HEATCO project (http://heatco.ier.uni-stuttgart.de/), where existing practices in the assessment of the value of time and congestion, accident risk reduction, health impacts and nuisance from pollutant and noise emissions are compared for various countries and a common platform is developed.

14.7 Conclusions

The most important conclusions of this chapter are:

1. The use of SCBA has become standard in many countries, but its pros and cons are still the subject of many debates. It is important to realize that SCBA takes place at the end of an analytical chain of effects. Improving the outcomes of SCBA is a matter of improving not only the valuation

methodology as such, but also the various models providing the inputs of SCBA.

2. In several domains, substantial progress has been made in measuring the non-market consequences of transport policies in monetary terms. As a consequence, the coverage of relevant effects by SCBA has increased during the last few decades. Aspects where further progress would be welcome concern the contribution of transport to global environmental problems such as climate change and the impact of transport on fragmentation of landscapes, both outside and within urban areas. Other important themes where further contributions would be most welcome concern the valuation of non-use values and option values, discussed in this chapter. It is important to realize that, as the world population gets richer, the importance attached to these aspects may be expected to increase. The theme of 'indirect effects' or 'economy-wide benefits' also deserves more attention in the future in order to clarify claims that transport benefits are systematically underestimated.
3. A similar conclusion holds for the quality aspects of transport services. As passengers become richer they tend to attach higher values to dimensions such as comfort, reliability, and information provision.
4. The choice between SCBA and MCA depends on whether the public authority wants to use consumer preferences as a benchmark: SCBA is essentially a market oriented approach –including corrections for market failures – to policy problems. MCA on the other hand is oriented towards the priorities of policy makers or various actors. An advantage of SCBA is that it is less susceptible to the risk of manipulation of inputs in order to serve as a justification of an a priori chosen result. As has been outlined in Table 14.6, the differences between SCBA and MCA are smaller than is sometimes thought.
5. A main feature of standard SCBA analysis is its focus on overall efficiency, but it ignores distributions of gains and losses, for example over regions, or groups of people (e.g. by income). Nevertheless there are ways to incorporate distributional dimensions, for example by using equity weights in order to give higher priority to low income groups. This may be interpreted as an incorporation of elements of MCA into SCBA.
6. An important function of both SCBA and MCA is that they provide a structure to decision making processes in terms of the explicit formulation of objectives and alternatives. Their function should preferably be not only that they help in identifying 'the best' alternative (or a group of good alternatives), but also that they stimulate the development of promising alternatives in earlier phases of the preparation of policy making. This means that these methods are preferably applied in an iterative setting where promising alternatives are gradually added. Such an

approach of course means that in the earlier phase of decision making the level of detail and precision of the analysis will be smaller than in later phases.

7. In real world applications, both SCBA and MCA will be subject to tendencies of manipulation by their users in order to achieve outcomes that please policy makers. Therefore it is essential that work underlying these applications is made public and that their quality is assessed by independent bodies. Systematic reviews of applications of SCBA in the past have shown that there are considerable biases in terms of underestimates of costs and overestimates of benefits.
8. A final observation is that it is no exception that, as policy making comes to completion, the main stakeholder who took the initiative for the transport policy tries to get the agreement of important disadvantaged actors by adding particular compensatory measures. This is one of the reasons why transport projects sometimes become much more expensive than originally intended. Expensive compensations added to the most efficient alternative may well make it less attractive in the end compared with some other alternatives. An implication is that SCBA should preferably focus not only on the most efficient alternative, but also on other good alternatives which perform better in terms of distributional effects.

NOTES

1 Of course, the range of effects will usually be broader, since price policies may also lead to changes in congestion and hence to travel times, but the essence of the gap between cost benefit analysis and policy weights remains.

2 The meaning of the PM outcome will be discussed in more detail in section 14.3.

3 The shadow price method and the conjoint analysis method clearly share the property that a price is obtained that cannot directly be observed in a market. The distinguishing feature of the shadow price method is that it follows from the costs to be made to satisfy a policy constraint, and as a consequence the price is based on prevention costs instead of damage costs.

4 Note that in our discussion of the demand curve we did not go into the issues of asymmetric responses with respect to changes in prices.

REFERENCES

Annema, J.A., C. Koopmans and B. van Wee (2007), 'Evaluating transport infrastructure investments: the Dutch experience with a standardized approach', *Transport Reviews*, 27, 125–150.

Berechman, Y. (2009), *The Evaluation of Transportation Investment Projects*, London: Routledge.

Blayac, T. and A. Causse (2001), 'Value of travel time: a theoretical legitimization of some nonlinear representative utility in discrete choice models', *Transportation Research B*, 35, 391–400.

Button, K.J. (2010), *Transport Economics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Cadot, O., L.H. Röller and A. Stephan (2007), 'Contribution to productivity or pork barrel? The two faces of infrastructure investment', *Journal of Public Economics*, 90, 1133–1153.

De Brucker, K., A. Verbeke and C. Macharis (2004), 'The applicability of multicriteria-analysis to the evaluation of intelligent transport systems (ITS)', in Y.I. Nakanishi and E. Bekiaris (eds), *Economic Impacts of Intelligent Transportation Systems: Innovations and Case Studies*, Amsterdam: Elsevier, pp. 151–179.

Dekkers, J. and J.W. van der Straaten (2009), 'Monetary valuation of aircraft noise: a hedonic analysis around Schiphol airport', *Ecological Economics*, 68 (11), 2850–2858.

Flyvbjerg, B., N. Bruzelius and W. Rothengatter (2003), *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge: Cambridge University Press.

Fridstrom, L. and R. Elvik (1997), 'The barely revealed preference behind road investment policies', *Public Choice*, 92, 145–168.

Geurs, K., R. Hajer and G.P. van Wee (2006), 'Option value of public transport: methodology for measurement and case study for regional rail links in the Netherlands', *Transport Reviews*, 26, 613–643.

Hanemann, M. (1991), 'Willingness to pay and willingness to accept: how much can they differ?', *American Economic Review*, 81 (3), 635–647.

Hoevenagel, R. (1994), 'The contingent valuation method: scope and validity', dissertation, Vrije Universiteit, Amsterdam.

Janssen, N.A.H., D.F.M. van Mansom, K. van der Jagt, H. Harssema and G. Hoek (1997), 'Concentration and elemental composition of airborne particulate matter at street and background locations', *Atmospheric Environment*, 31, 1185–1193.

Johansson, P.O. (1991), *The Economic Theory and Measurement of Environmental Benefits*, Cambridge: Cambridge University Press.

Lichfield, N. (1988), *Economics of Urban Conservation*, Cambridge: Cambridge University Press.

Lyons, G., J. Jain and D. Holley (2007), 'The use of travel time by rail passengers in Great Britain', *Transportation Research A*, 41, 107–120.

Nijkamp, P., P. Rietveld and H. Voogd (1990), *Multicriteria Evaluation in Physical Planning*, Amsterdam: North Holland.

Nijland, H.A., E.E.M.M. van Kempen, G.P. van Wee and J. Jabben (2003), 'Costs and benefits of noise abatement measures', *Transport Policy*, 10, 131–140.

Perman, R., Y. Ma, J. McGilvray and M. Common (2003), *Natural Resource and Environmental Economics*, Harlow: Pearson Education.

Pickrell, D.H. (1989), *Urban Rail Transit Projects: Forecast versus Actual Ridership and Costs*, Washington, DC: US Department of Transportation.

Redmond, L.S. and P.M. Mokhtarian (2001), 'The positive utility of the commute: modeling ideal commute time and relative desired commute amount', *Transportation*, 28, 179–205.

Rietveld, P. (2003), 'Winners and losers in transport policy: on efficiency, equity, and compensation', in D.A. Hensher and K.J. Button (eds), *Handbook of Transport and the Environment*, Handbooks in Transport 4, Amsterdam: Elsevier, pp. 585–602.

Rietveld, P., F. Bruinsma and D. van Vuuren (2001), 'Coping with unreliability in public transport chains: a case study for Netherlands', *Transportation Research A*, 35, 539–559.

SACTRA (Standing Advisory Committee on Trunk Road Assessment) (1999), *Transport and the Economy*, London: SACTRA.

Schipper, Y. (2001), *Environmental Costs and Liberalization in European Air Transport: A Welfare Economic Analysis*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Small, K.A. and E.T. Verhoef (2007), *The Economics of Urban Transportation*, London: Routledge.

Van Pelt, M.J.F. (1993), *Ecological Sustainability and Project Appraisal*, Aldershot: Avebury.

Van Wee, B., P. Rietveld and H. Meurs (2006), 'Is average daily travel time expenditure constant? In search of explanations for an increase in average travel time', *Journal of Transport Geography*, 14, 109–122.

Wardman, M. (1998), 'The value of travel time: a review of British evidence', *Journal of Transport Economics and Policy*, 32, 285–306.
Zahavi, Y. (1979), *The UMOT Project*, Report DOT-RSPA-DPB-2-79-3, Washington, DC: US Department of Transportation.
Zhu, X., J.N. van Ommeren and P. Rietveld (2009), 'Indirect benefits of infrastructure improvement in the case of an imperfect labor market', *Transportation Research Part B*, 43, 57–72.

15

Transportation models and their applications

Toon van der Hoorn and Bert van Wee

15.1 Introduction

Models have been defined in many different ways and there are various kinds of models. In this chapter, a model is defined as a simplified image of part of reality. The major form of application for models is to fully compute one or more of the expected, possible, desirable or undesirable situations (scenarios – see Chapter 13) in a specific forecasting year. Possible results of these computations may be, for example, the number of passenger kilometres or total mobility in a country or region in a certain year, the number of truck kilometres or the ton kilometres transported by inland barges in that country or region in a certain year or the load on the trunk-road network in a certain region in the morning rush hour. Models are often used to get an insight into the effects of transportation on what is sometimes called 'autonomous developments' and/or to gain an insight into the effects of proposed policy measures on transportation. Usually not one but several futures are predicted. In other words, the effects on transportation are computed for several scenarios concerning the socio-economic development and policy variants (see Chapters 13 and 14). So a model should not be viewed as a crystal ball for predicting the future but rather as a means for roughly estimating the effects of developments and policy measures.

The aim of this chapter is to provide readers with an insight into the way in which models work and to show what they can be used for as well as what they cannot be used for. This knowledge should enable the reader to form an opinion about the question as to whether or not a specific model can be used to answer certain questions. This chapter limits itself to what are often called 'strategic models'. Models for traffic flows on highways intended for real-time traffic management are presented in Chapter 7. For a further insight into the literature on model building, see Ortúzar and Willumsen (2011).

Section 15.2 provides an overview of various kinds of models. Section 15.3 addresses the concept of elasticities, which is a concept used in many models. Section 15.4 discusses the so-called traditional aggregated models, while the so-called disaggregated models are discussed in section 15.5. Section 15.6 deals with validation of models: which criteria are used to judge whether the model is correct? Section 15.7 presents some examples of models currently in use. Section 15.8 addresses the question of what can and what cannot be done by using models. The main conclusions are presented in section 15.9.

15.2 Kinds of models

Models can be classified in various ways. Some of these classifications are presented in this section, but we don't claim to have covered them all. Before discussing the model categories, what all models have in common, as indicated before, is that they are a simplified image of part of reality. A transportation model describes human behaviour on the basis of a theory. This theory describes the connections between the variables in the model. Exogenous (independent or explanatory) variables are variables the size of which is determined externally, outside of the model. Endogenous (dependent or 'to be explained') variables are variables the size of which is determined by the model. For example, in a model determining car ownership as a function of (among other things) economic growth, economic growth is exogenous and car ownership is endogenous. The correlation between variables is represented in mathematical formulas.

The mathematical formulas in a model often contain so-called coefficients, which quantitatively indicate the way in which the value of an exogenous variable affects the value of an endogenous one. A fictional example of a model based on elasticities (see section 15.3, and Chapters 3 and 6) is of a 1 per cent increase in the income level leading to a 0.4 per cent increase in car ownership. The value of 0.4 is the coefficient here. Determining the coefficients in a model is known as calibration. The aim is to make the model's results connect as well as possible to the collected empirical data, often from a specific base year. If methods from mathematical statistics are used to achieve this, it is called statistical estimation. Other methods are used as well, however, though regrettably these are sometimes still of a high trial-and-error calibre.

Descriptive versus exploratory models

Descriptive models represent the correlations between variables without explicitly considering causality, whereas explanatory models will map out causes as well as consequences.

In general, statistical data and their analyses provide only limited indications as to what the theory should be like. It is possible to develop a perfect statistical model about the relationship between the number of storks and the number of births (most storks are to be found in developing countries, where the number of births is high). But statistical correlation between variables does not provide an insight into the causality and its direction (e.g. does the presence of an airport in a region cause extra economic growth or are airports typically situated in regions with a strongly developing economy?). Therefore it is important to distinguish between descriptive models and explanatory models.

Although any model should ideally be based on an explicit theory, descriptive models can nonetheless often be useful. For example, by using the data mining technique a file containing a large number of variables can be searched systematically for the existence of correlations. A concrete example from traffic safety is the relating of accident frequencies to road characteristics, regional characteristics, time of day and period, and so on. Afterwards it can be studied to determine whether or not the relationships found can be explained in any useful way. In other words, are the relationships theoretically underpinned? In general, in any research project, it is advisable to start with descriptive analysis (e.g. by frequency and cross-tabulations) before developing models.

Spatial versus non-spatial models

Models can be divided into non-spatial and spatial models. In the former, physical space does not play a role; in the latter, the location of activities in space is an explicit component in the model.

An example of a model of the former kind is a model for car use in a country, estimated on the basis of a time series and with the population in the age group 20–65, the gross domestic product (GDP), fuel prices, road capacity and the price per kilometre for road transport as exogenous explanatory (or independent) variables (for example, van Mourik, 2008).

In the case of spatial models, the research area is often divided into zones (e.g. 1000 zones). Furthermore, there are networks for cars and other transport modes, often with (many thousands of) nodes and links (junctions and stretches of road). Each zone has its centroid, which is connected to the networks for cars and public transport through so-called feeder links. One of the objectives of such a spatial model is to establish the origin–destination matrices, which indicate for each transport mode the numbers of trips between every pair of zones in the study area. Because of the finite number of zones, such a model is called discrete. In principle, continuous models, in which the study area is not compartmented into zones, are possible as well, but for a number of reasons these are less common. An example of a continuous model is a model for the number of trips between A and B as a function of the distance between A and B.

Aggregated versus disaggregated models

In the traditional aggregated four-step model (see section 15.4), the zone is the unit of observation. In disaggregated models (see section 15.5), on the contrary, attention is focused on the individual traveller.

The aggregated model dates back to the 1950s. It was developed in a period when the 'predict-and-provide' approach was at the forefront. Owing to prosperity, car ownership and car use strongly increased, which required extra road infrastructure. For example, a forecast was produced for car use over, for example, the next 20 years, and based on this forecast the extra number of lanes to be built was determined. During the 1970s this line of thought shifted in favour of the idea that it would be impossible to meet the demand for car mobility infinitely (the 'predict-and-prevent' paradigm). Instead, it was decided to opt for a different approach: a combination of containing measures, regulation, promoting alternatives for the car, optimal use of existing infrastructure and limited building of new infrastructure. Such an integral policy requires a different kind of model, which resulted in the emergence of the disaggregated models.

Owing to the progress achieved in statistical estimation techniques from the 1970s onwards, disaggregated models can nowadays be estimated much more efficiently by using individual variation instead of zone averages. Note though that the models thus developed need to be aggregated to make them workable for forecasting. After all, it is not the mobility of one specific individual that transport policy-makers are interested in; their interest is in the aggregated traffic flows on the roads and in public transport.

Static versus dynamic models

In dynamic models, as opposed to static models, the time factor plays an explicit part. Changes in behaviour are assumed not to occur instantly, but over a certain time period.

One example is a model for the effect of the price of fuel on car use. If the fuel price per litre were to rise from €1.35 to €2.70 tomorrow, the short-term effect would be limited. Many commuters go to work by car and would have to do so tomorrow as well, for lack of alternatives. Any effects would only become visible in the long run. These effects may be: a different method of transport, different destinations closer by (e.g. for shopping), less frequent trips, finding a job closer to home, moving to a place closer to work, keeping the currently owned car for a longer time (to keep down depreciation costs) or rather switching to a more fuel-efficient car, cutbacks on expenditures other than those for the car, and so on. In this case, the effect of the variable (e.g. the fuel price) on other variables (e.g. the commuting distance) is 'lagged'.

In transport studies, static models are usually estimated on the basis of what are called cross-section data. These data constitute a snapshot taken at one moment in time. Dynamic models require time-series data. Preferably, these data concern one constant group of individuals over time – so-called panel data. An advantage of such data is that the development of the exogenous as well as the endogenous variables over time is known for each individual.

Models based on revealed preference or based on stated preference

When collecting data on individual choice behaviour, two options are open: that of revealed preference (RP) and that of stated preference (SP). In the case of RP, the actual choices made by the individual are observed in a real situation. In the case of SP, the researcher confronts the individual with hypothetical selection situations.

In the case of RP, people are asked, for example, to keep a travel diary indicating data like origin, destination, travel mode, time of day and travel purpose for every trip. In the case of SP, people for example are shown 5–10 cards each with two trip alternatives, from which they recurrently are asked to choose their preferred one. This kind of model has two major advantages. Firstly, it is often cheaper. Secondly, it can be used for making predictions for as yet non-existent situations, like a high-speed Maglev train to Berlin. Of course there are also drawbacks, the most important of which is that the models have to be adjusted for respondents' ignorance and exaggerations.

An example of a study using SP can be found in Mabit (2009). An example of a study using RP is in Whelan et al. (2010).

Models for travel versus models for activities

Traditional models consider travel behaviour in isolation. Activity-based models (see also Chapter 3) link travel to the total activity pattern.

Trips in essence constitute a demand derived from the full pattern of activities. A trip is one of the mutually coherent, subsequent activities in time and in space. In the models discussed earlier, these interactions between subsequent trips are largely lost. It is not only the income and time-budget restrictions that affect travel behaviour, for example. There are all kinds of constraints that are connected mainly to the life-cycle stage in which the consumer household finds itself. Synchronization of all kinds of activities (like meals) within the family put extra time restrictions on the choice of behaviour concerning activities. These constraints ultimately determine the collection of feasible selection alternatives (see Chapter 3). The advantages of these 'activity-based models' are (1) that they provide a better description of behaviour and, in connection with this, (2) the fact that, from a conceptual point of view, they are more satisfactory. After all, trips are a derivative of activities. The major drawbacks are that the models rapidly get very complex and that their development is expensive. An example can be found in section 15.7.

Methods that do versus models that do not take into account the effects of transport on land use and on the economy

Traditional models describe transportation as a derivative of spatial planning and economic development. In actual practice, however, a partial reverse effect may be possible: the presence or absence of transport infrastructure does affect the pattern of spatial and economic activities (see also Chapter 2, and Figure 2.1). There are two kinds of model that count, with interactions in two directions: the so-called land-use transportation interaction (LUTI) models and the models for determining the economic effects of transportation.

The spatial effects of the transport infrastructure can be considerable, especially in the long run (think of cities that came into being near natural harbours or river crossings). The building of a new road or railway, for example, may well lead to businesses settling in their vicinity, or business may leave an area with heavy congestion.

Investments in infrastructure lead to direct effects in the form of travel time savings and travel cost savings. Often, however, there are indirect effects as well. These indirect effects manifest themselves when direct benefits of investments in infrastructure are transferred, for example through price reductions, as a result of which companies and families that do not use the improved infrastructure nevertheless profit from the benefits of these investments. Sometimes there are additional indirect effects on the standard of living as well: positive or negative benefits on top of the direct and indirect effects on welfare. These effects may come about if market imperfections exist and/or if effects are exchanged with foreign countries.

A drawback of the use of these spatial interaction models is, however, that their effects only become apparent after a long period of time and need not be due solely to infrastructure. Therefore the possibility for validating these models remains rather restricted where the long-term dynamics of the economy are concerned.

An example of a LUTI model is SASI, developed for a European Union project (Wegener, 2008). More information about LUTI models can be found in Geurs and Ritsema van Eck (2004).

Models for passenger transport versus freight transport

The majority of the transport studies models deal with passenger transport. On most roads, the share of freight transport is limited (traditionally about 10 per cent). Yet models for freight transport are gaining in importance. Freight transport is expected to grow much more than passenger transport. Trucks are an increasingly important part of road transport because, through technical measures, the CO_2 emission by passenger cars is decreasing more rapidly than that of trucks. Besides, heavy trucks damage the road surface, and many slow, heavy trucks lead more quickly to traffic jams. One of the specific problems in developing freight transport models is the availability of data. Aggregated economic statistics often use a too coarse classification of economic sectors. Besides, the availability of international data for the EU has decreased through the disappearance of interior borders in the EU. Disaggregated data are hard to collect, because their collection is dependent on the voluntary cooperation of businesses and because one company encompasses many different actors (management, administration, production departments, logistics department, transportation department). See section 15.7 for an example.

15.3 Elasticities

Many models contain what are called ‘elasticities’. This has already been briefly discussed in Chapters 3 and 6. In this section the notion is further elaborated on.

Elasticities enable us quickly to get an insight into the following question: to what extent does a change in one variable lead to a change in other variables?

An example is the elasticity of car use (in kilometres) as a function of the fuel price. If this elasticity is –0.4, car use will diminish by 4 per cent if the fuel price rises by 10 per cent. In general, elasticity ε of a variable *y* because of a change in a variable *x* is defined as:

$$\varepsilon = \frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}} \tag{1}$$

Two kinds of elasticities are distinguished: direct elasticities and cross elasticities. The current example is about direct elasticities: the effect of an attribute of a car on car use itself. If we look at the effect of fuel price changes on, for example, public transport, we are dealing with cross elasticity. If we focus on price elasticities, direct elasticities are negative whereas cross elasticities are positive. Cross elasticity of the use of public transport as a function of the fuel price might, for example, be 0.1: in that case, the use of public transport would increase by 1 per cent if the fuel price rose by 10 per cent. In this example, therefore, the cross elasticity is only small: this corresponds with what we have learned from research, namely, that substitution between the car and public transport is very limited (e.g. Bakker and Zwaneveld, 2009; Mackett, 2009). If totally independent goods or services are at stake, the cross elasticity is 0. This may apply to car fuel and dishwashing detergents, for example.

Elasticities are especially popular because they are so very simple and understandable, but the risk of misuse is considerable. Elasticities are usually not constant: a rise in fuel prices from €1.35 to €2.70 (+100 per cent) may have a different effect than a fuel price rise from €2 to €4 would (which is also +100 per cent). The effects of an increase (e.g. of +20 per cent) may be different from the effects of a decrease (of –20 per cent). The relative effects of an increase of 100 per cent may be different from the relative effects of an increase of 10 per cent. And for a given level of overlap between markets for car use and public transport the values of cross elasticities depend on the shares of the car and public transport. Instead of making simple approximations with elasticities, one would have to consider the total demand curve of *y* as a function of *x* and possibly other goods. To deduce that demand curve, one usually would have to estimate models. In addition, it is risky to 'transfer' elasticities over place and time: elasticities as found in one country do not necessarily apply to another country. And elasticities change over time. For example, the relationship between income per capita and car ownership levels is S-shaped, implying that elasticities vary across income levels, and therefore over time.

15.4 The traditional aggregated model

As was mentioned earlier, in the case of spatial models the study area is divided into zones (e.g. 1000 zones). Furthermore, there are networks for the car, for public transport and for slow modes of transport (although the latter two are not always involved), with (many thousands of) nodes and links. Each zone has its own centroid, which is connected to the networks through feeder links.

In the aggregated model, usually four steps (sometimes five steps) can be distinguished:

1. Generation (production) and attraction: how many trips depart from a zone and how many arrive there?
2. Distribution: what is the direction of the trips thus generated?
3. Selection of the transport mode (modal split): which transport mode is used?
4. Assignment: which route is taken to get from one zone to the next? And sometimes:
5. Time-of-day selection: at which time of the day (peak or off-peak)? Many of the older models confined themselves to modelling the peak period only, since that was the busiest period and thus determined the maximum required number of extra lanes.

In the origin–destination (O–D) matrix, generation models compute the total number of trips O_i from zone i during a certain time-span. That number is based on the characteristics of the zone concerned, like population size, retail area and employment. These characteristics are known as 'land-use variables'. Trip-attraction models compute the total number of trips D_j in the O–D matrix to a certain zone j.

The oldest distribution models were derived from the gravity theory (Newton's laws about attraction between celestial bodies). At present, though, a distribution model usually looks like this:

$$T_{ij} = A_i O_i B_j D_j f(c_{ij}) \tag{2}$$

where T_{ij} is the number of trips between zones i and j, and where O_i and D_j are the computed generation from zone i and the attraction from zone j, respectively, as derived from the generation and attraction model.

The balancing factors A_i and B_j for the zone of origin i and the zone of destination j accomplish that the trip-distribution model's results are consistent with the results from the trip-generation and trip-attraction models. This results in what is called a doubly constrained model. Single-constraint models are used as well but they are not discussed here.

In the impedance function $f(c_{ij})$, c_{ij} represents the generalized travel costs between zone i and zone j. For example, c_{ij} = a×travel distance + b×travel time + c×toll charge (where a, b and c are weighting factors). In the generalized travel costs, times are translated into costs or, precisely the other way around, costs are translated into times. The translation takes place through the so-called value of time, which is the amount of what one hour of travel time is worth (see Chapter 14).

Modal split models distribute the number of trips T_{ij} for each i and j over the available transport modes. Mostly, the aggregated version of the logit model is used for this (for a general introduction to the logit model, see section 15.5 on disaggregated models).

Usually, the last step in the aggregated model is the allocation of trips to the network. Various methodologies are available for this. The first and most simple methodology is that of all-or-nothing (AON). In this case, it is assumed that everyone takes the same, shortest route between i and j. Shortest is defined here in terms of travel time or generalized costs. The second methodology is that of stochastic allocation, also called multiple routeing. In this case, differences in the perception and taste of individuals are included in the model, as a result of which other routes, besides the shortest one, will also be considered. The third methodology takes congestion into account. Congestion causes changes in the travel times on a stretch of road. The relationship between the traffic intensity on a stretch of road, on the one hand, and its capacity, on the other hand, is described by a speed-flow curve (see Chapter 7). Because of delays due to congestion, the shortest route between i and j will run differently. The ultimate shortest route can be found through an iterative process. In the balance, the so-called Wardrop principle applies: all routes used between i and j have identical travel times and no one can improve on his travel time by opting for a different route. If, in a situation of congestion, the methodology of multiple routeing is taken into account as well, the fourth methodology, stochastic user equilibrium assignment, is applied. For further details, see Ortúzar and Willumsen (2011).

Models that take the time of day into account are relatively recent ones and are mostly of the disaggregated kind, which will be discussed in section 15.5.

Even today, the aggregated model is still widely used. See, for example, USDOT (1999), which is a fairly old but still a good reference. Disaggregated models do make more efficient use of individual data than aggregated models. Nevertheless, in many cases those individual data have to be specifically collected for the study at hand. In addition a disaggregated model has to be aggregated before it is ready for forecasting (see section 15.5). Data for aggregated models at a zone level are often more readily available from routinely collected statistical data.

15.5 Disaggregated models

The disaggregated approach (Ben-Akiva and Lerman, 1985; Ortúzar and Willumsen, 2011) uses the individual or the household as the basic unit for the analysis. When participating in traffic, an individual needs to choose between alternatives. His or her travel behaviour can be unravelled into a number of decisions:

1. To make a certain trip, or not to make it (generation).
2. The choice of the destination (distribution).
3. The choice of the means of transport (travel mode selection).
4. The choice of the route to be taken (route selection).
5. The choice of the time of day for the trip (time-of-day selection).

The model needs to indicate the alternative that the individual will select from the set of alternatives offered. The rational individual will prefer the alternative that will yield the highest utility. However, this utility cannot be perceived exactly. It is therefore assumed that total utility consists of the sum total of a deterministic component and a stochastic component (representing the dispersion between individuals). The deterministic component is a function of those characteristics of alternative i that are relevant for the individual (like speed, costs and comfort) and of the individual's socio-economic characteristics (like gender, age, education and income). The stochastic component, which describes those characteristics of the individual and of the alternatives that cannot be observed by the researchers, means that we are unable to explain the individual's choice exactly but can only assign it a probability.

The most used disaggregated model is the multinomial logit model:

$$\mathrm{P}_i = \frac{\exp(V_i)}{\sum_j \exp(V_j)} \tag{3}$$

where P_i is the likelihood that the individual will select alternative i, and V_i is the deterministic part of the utility of alternative i.

This approach can be applied to the selection of the mode of transport, for example. In that case, for each mode of transport a number of explanatory variables are specified that will affect the individual's decision-making, like travel times, travel costs, waiting times and transfer times. Furthermore, various socio-economic characteristics of the individual, like income and age, will also affect the decision-making. The logit model determining the likelihood that travel mode i is selected will then look like this, for instance:

$$V_{\text{public transport}} = c_1 \times \text{travel time} + c_2 \times \text{costs} + c_3 \times \text{waiting time} + c_4 \times \text{transfer time} + c_5 \times \text{dummy} \qquad (4)$$

with dummy=0 if the individual does not own a car, and dummy=1 if he does.

And:

$$V_{\text{car}} = c_6 \times \text{travel time} + c_7 \times \text{fuel costs} + c_8 \times \text{parking costs} \qquad (5)$$

The coefficients (in the above example c_1 to c_5, and c_6 to c_8) in the utility function V_i indicate how the effects of the various variables and characteristics affect the likelihood that a specific travel mode will be selected. The larger the coefficient, the more it affects the utility an individual derives from the travel mode concerned. Negative coefficients indicate a variable's disutility: the variable affects the utility perceived by the individual in a negative way (an example of such a variable is fuel costs).

The values of the coefficients are determined by using statistical estimation techniques (on the basis of what is called 'maximum likelihood'). The basis for this is a random sample of trips actually made, which is usually derived from a survey. In these surveys, respondents have filled out trip diaries. To each trip, travel times, travel costs and other explanatory variables are added by using the available networks in the study area.

A distribution model can also be a logit model. Here also, characteristics can be identified that affect the likelihood of the individual selecting destination d from a collection of possible destinations D. On the one hand, the effort required from the individual to reach a certain destination plays a part, and on the other hand there is the attractiveness of a destination. The effort required to get to d is usually expressed as a kind of weighted sum total of the effort for each travel mode to get to d.

This so-called logsum variable is determined from the underlying modal split model. In this structure the choice process is sequential. An individual first chooses a destination, and next a travel mode. This is often the case for commuting journeys, for example. It is also possible that the choices of destination and travel mode are simultaneous, for example, for shopping, either by car to a suburban shopping mall or by tram to the city centre.

When modelling an individual's trip frequencies in a disaggregated trip-generation model, various explanatory variables can be distinguished. Socio-economic characteristics like income, profession and age prove to be important in this. Furthermore, characteristics of the household that the individual is a part of, like the number of members of the household and whether or not there are children, are important. Besides, factors like the accessibility of destinations from the individual's residence and the attractiveness of making a trip from the residence (sometimes) play a part.

A disaggregated model for time-of-day selection usually uses the current time of departure and a number of periods (e.g. per quarter of an hour) before and after as alternatives. The utility of each alternative period is a function of the travel time and the travel costs in that period. If there is a lot of congestion, the travel time will be high, and travel costs can get high if a peak charge is introduced, for example. This diminishes the utility of such a period, and with it the likelihood that that particular period will be selected.

Various other disaggregated models exist besides the logit model. In the logit model it is assumed that the stochastic component of the utility follows a Weibull distribution (a family of continuous probability distributions). One alternative is the 'probit' model, in which the stochastic component follows a normal distribution. Furthermore, over the last few years the mixed logit model has become very popular, especially among academic researchers. In this model, the coefficients (c_1 to c_8 in the example presented earlier) do not have a fixed value; instead, they are individual specific, with a certain statistical distribution over the population (Ortúzar and Willumsen, 2011). Because of its complexity, 'mixed logit' is better suited for describing behaviour than for predicting future behaviour.

As already indicated in section 15.4, the results of disaggregated models need to be aggregated before they can be used for forecasting. A description of an aggregation procedure is beyond the scope of this chapter; see Ortúzar and Willumsen (2011). After aggregated O–D matrices have been obtained through the aggregation process, these can be used for making the allocations, in exactly the same way as described above for aggregated models.

15.6 Validation of models

Validation is defined as the assessment of whether or not the model describes reality correctly. The question is phrased simply, but the answer is usually complicated (Barton Aschman Associates and Cambridge Systematics, 1997). For validation models designers use independent data not used during the construction of the model. They want to know the reliability of the model for a wide range of situations.

Criteria during validation of a transport model may be:

1. **The level of validation.** For example, if the model is designed for motorways, the model may not need to correspond exactly on secondary roads.
2. **Application area.** For example, when using a peak model designed for working days, model users cannot expect the model to perform well for peak travel during holidays.
3. **Qualitative criteria.** For example, for which policy measures should the model be suitable in general?
4. **Quantitative criteria.** For example, on which roads should the model be able to calculate travel time losses due to congestion? All roads or just a few important roads under scrutiny? The model designers must also define their levels of confidence about the forecasts that they judge to be acceptable.

The analysis of the non-conformities of the model will ideally give clues about required improvements:

1. At the lowest level, additional or other model parameters could be added.
2. At the level of the computer model, the accuracy could be enlarged, for example by calculating with more digits or taking smaller time steps (in dynamic models).
3. In iterative model-solving algorithms, the step size could be reduced or the number of iterations could be increased.
4. The mathematical model might not sufficiently adhere to the conceptual model and hence might need to be enhanced.
5. Finally, at the highest level, the conceptual model might need some modifications because the dependencies are not correctly modelled or because the theory needs amendment.

15.7 Some examples of models

In this section, three models are described that are frequently used at the national level.

The Dutch National Model System

The Dutch National Model System (in Dutch: Landelijk Model Systeem – LMS) is an internationally renowned, unique measuring instrument for designing transportation policies. Rijkswaterstaat, the public works department of the Dutch Ministry of Infrastructure and the Environment, has been using it since 1986. The LMS is owned by the Centre for Transport and Navigation (DVS), which is a part of Rijkswaterstaat. Besides the LMS there are four regional clones, together covering the whole of the Netherlands and fully consistent with LMS.

The LMS is a disaggregated model system that can estimate future traffic flows, both on the trunk-road network and in public transport. The LMS is a spatial model, which means that the Netherlands and small parts of bordering countries have been compartmented into about 1500 zones, each with its own characteristics (e.g. employment, number of students, income, size of the working population in the base year). These characteristics were mainly derived from statistics prepared by the CBS (the Dutch Central Statistical Office). For the forecasting year, these data were derived from other models.

When using the model, the following steps can be distinguished:

- **Step 1 – Selection behaviour for each type of household**
  The model starts with the selection behaviour of individuals or households. The individual makes a comparative assessment of costs versus time. For example, using the car in the morning peak hours will take more time than making the same trip at a later point in time. These choices are based on behaviour actually observed (RP; see section 15.2), as found either in the trip diaries from the annual OVIN, the Dutch national mobility study, or derived from the responses from a study in which individuals were asked to state what their behaviour might be in reaction to a change (SP; see section 15.2). The latter applies to the planned introduction of the kilometre tax (road pricing), for example.
- **Step 2 – Kinds of travellers in the base year**
  Using the choices made by individuals and households, the LMS establishes what are called traveller types. What are the characteristics of the people making the choices? The following characteristics may be distinguished, among others: age (grouped into 18 classes), gender (2 classes), driving licence holder or not (2 classes), participation in society (working, retired, etc., 6 classes), income (10 classes), education level (6 classes) and student or not (2 classes). For households there is a similar classification. This is done for every zone, in both the

base year and the forecasting year, and is necessary in being able to estimate future traffic flows. For example, if the number of people over 50 years of age in 2030 is twice as large, the model takes into account an increase in the number of trips made especially by people in that age group.

- **Step 3 – Selection models**
  In step 2 we determined traveller types and their selection behaviour. In this step the selection models are computed. The following choices are taken into account:

1. the choice in favour of a driving licence and car ownership;
2. the decision as to whether or not to make the trip;
3. the selection of the travel destination, the travel mode and the time of day (in the latest version of the LMS this is modelled simultaneously; see section 15.5);
4. the selection of the route to take.

The LMS thus mainly consists of four selection models which are directly related to the choices listed above. To determine the transport demand, the first three selection models in particular are important. The LMS deduces the transport supply from the network data and thus from data on accessibility. Using the driving licence and car ownership model, the LMS determines car ownership per household and zone but uses the DYNAMO forecast (see 'DYNAMO' below) for the whole of the Netherlands as a constraint. The car ownership model, therefore, works as a distribution model for a given external national total. The LMS determines the choice of destination, travel mode and time of day on the basis of the tour generation and accessibility. Except for the choice of the route, all LMS selection models are disaggregated logit models.

Figure 15.1 visualizes the coherence between various selection modules from the LMS.

For the assignment, QBLOK is used. The essential difference between QBLOK and other equilibrium assignments (see section 15.4) is the calculation of the link travel times. During the calculation of the travel time on a link, QBLOK takes into account the inflow from preceding links. The inflow on the link is constrained when there is congestion on the preceding links. It is also investigated whether blockades occur. These limit the maximum outflow from a link, owing to the distribution of congestion over the network. A blockade occurs when traffic on a

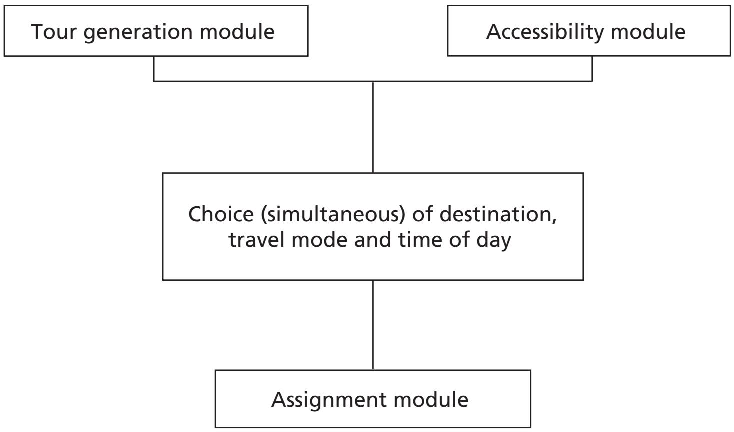


Figure 15.1 Coherence between selection modules from the LMS

link is halted by a bottleneck elsewhere, but will not pass that particular bottleneck itself. Another term frequently used for that phenomenon is 'blocking back'.

- **Step 4 – Types of travellers in the forecasting year**
  After the selection models that are to be used have been established, the model can compute the situation in a future year. The model now determines the size and composition of the future traveller population for each location in the Netherlands, thereby using all input on demographic and socio-economic data as well as spatial developments. The model thus computes the types of travellers in 2030, for example.
- **Step 5 – Changed selection circumstances**
  In this step, the planned policy options that may affect the selection behaviour are taken into account. An example of such a policy option is improving the railway connections between Amsterdam and the IJsselmeer polders. This may mean that people would sooner opt for public transport. A second example is making parking in Rotterdam more expensive, in which case people would use the tram more often and leave their car at home. A further example is widening or improving the utility of a stretch of road of the A4 motorway, in which case the trip to work would take less time, so car owners would sooner opt for using the car or travelling by this route.
- **Step 6 – Forecast: new travelling behaviour**
  The model computes both the short-term and the long-term changes in the selection behaviour. A heavier load on a stretch of road means, in the short run, that people opt for a different route, whereas in the long run they will opt for a different time of departure, while in the even longer run they will opt for a different travel mode and ultimately maybe even

for a different destination. The model forecasts the number of travellers for each transport mode as well as the number of kilometres they travel. It also computes the transport flows within and between zones. Finally, the model allocates these trips to the trunk-road network and public transport.

In order to be able to run the LMS, input is required. As could be deduced from the outline presented above, this input consists of:

1. road networks, including, for example, toll charges;
2. public transport systems;
3. parking costs;
4. socio-economic and employment data for each zone, both in the base year and in the forecast year;
5. driving licence and car ownership data;
6. a description of passenger mobility and freight transport in the base year.

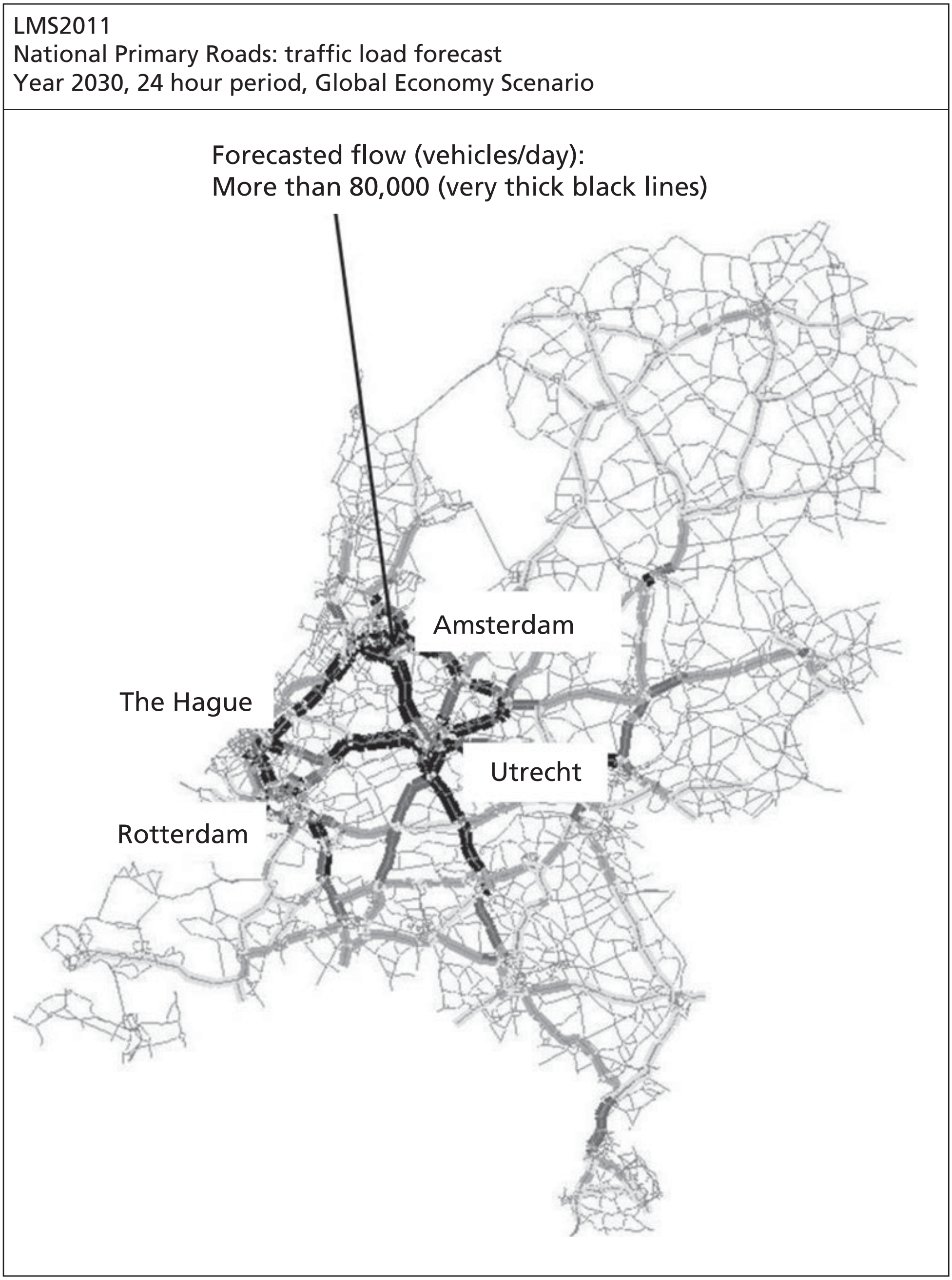
The LMS distinguishes various transport modes: car driver, car passenger, train, bus/tram/metro, bicycle, walk, and finally bus/tram/metro as access transport for the train.

There are 11 travel purposes distinguished in the LMS, covering commuting, business, school and private travel. Some travel purposes are non-home-based.

There are nine times of day distinguished, covering peak, off-peak, evening and night. The number of day periods is so large because the peaks widen more and more. To be ahead of traffic jams motorists leave earlier or later. For the benefit of a correct prediction of the travel time losses due to traffic jams, it is essential to predict the distribution across times of day as precisely as possible.

The output of the LMS consists of:

1. Forecasts about passenger mobility in the Netherlands in the forecast year. For the transport modes and travel purposes listed, as well as for the time of day, the LMS distinguishes: morning peak, evening peak and remainder of the day.
2. Forecasts of the load on the trunk-road network in the forecast year. A simple example is presented in Figure 15.2.



Source: Rijkswaterstaat.

Figure 15.2 Illustration of the amounts of traffic as computed by the LMS

Quality of the model

Ten years after the completion of the LMS, in 1996, the quality of the model was tested by way of an audit (Bates et al., 1996). The researchers concluded that the LMS had been prepared using the latest academic insights. On various occasions, further quality checks were carried out by comparing the forecasts to the actual developments (Gunn and van der Hoorn, 1998; de Jong et al., 2008). In this way, the 1986 forecasts for 1996 were compared with the 1996 reality. Except for the fact that the model underestimated the growth of the socio-recreational traffic, the quality of the forecasts proved to be reasonable. Since then, the development of household incomes has been incorporated in the model. Recently a comparison was made for 2010. Total mobility growth was predicted well. Car driver kilometres were overestimated; those of car passengers and walking/cycling were underestimated. A large part of the 'wrong' forecasts was caused by unexpected developments in society outside transportation. In particular, both the population and the work force (particularly women) grew more than expected. Incomes per household increased less than expected. The anticipated pricing measures (road pricing, kilometre charge) did not materialize. Public transport increased strongly through the introduction of the SOV free public transport pass for students. The biggest matter for concern was the underestimation of the growth of road congestion. This is why the periods of the day have been refined in the latest LMS. Equally the assignment method QBLOK has been improved. See also Chapter 13 for uncertainties about future developments.

DYNAMO

DYNAMO (Meurs and Haaijer, 2006) is a model to forecast the size, composition and use of the Dutch passenger car fleet for the period 2003–40, together with the resulting emissions. It is a dynamic model, which means that the effects in one year affect results in the next. The heart of the model is an equilibrium module where the prices for second-hand cars emerge in such a way that demand and supply are in equilibrium, both for the fleet as a whole and for individual household types. Size, use and composition of the car fleet are functions of the household and car characteristics (like the fixed and variable car expenses, among other things).

The model needs to be able to predict:

1. the number of cars per household;
2. the kind of fuel;

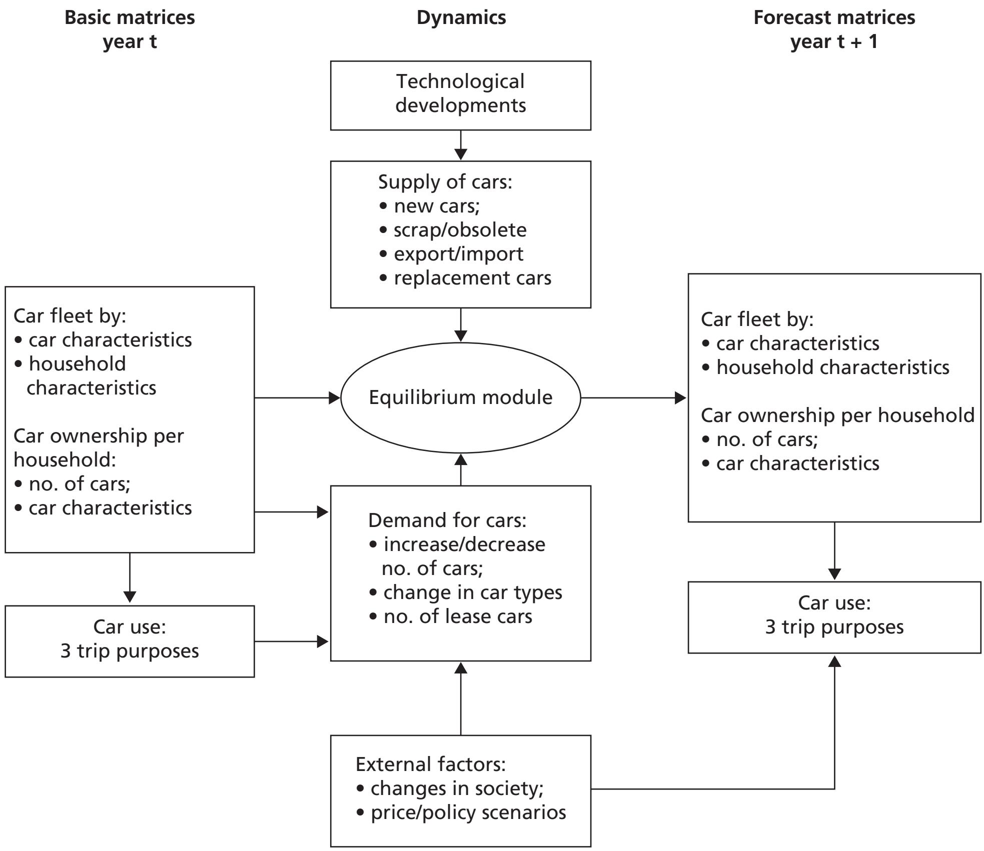


Figure 15.3 Structure of the Dynamo model

3. the weight category;
4. car age;
5. fuel consumption;
6. type of owner (private versus lease or company cars);
7. car use (the number of kilometres driven);
8. types of households (relative to the types of cars), which are defined by four characteristics: size of the household, number of working people, age of the oldest individual in the household and real disposable household income;
9. income effects for groups of households;
10. effects on government income.

Figure 15.3 represents the DYNAMO model's structure.

Activity-based model

As mentioned before, traditional models consider travel behaviour in isolation. Activity-based models link travel to the total activity pattern (see also Chapter 3).

A state-of-the-art example of an activity-based model is presented in Cirillo and Axhausen (2010). It is a model for the choice of activity type and timing, incorporating the dynamics of scheduling, estimated on a travel diary filled out by interviewees over six weeks.

The main focus of the study is the inclusion of past history of activity involvement and its influence on current activity choice. The model developed to generate the activity programmes is disaggregated and based on individual choices. The alternatives are discrete and are exogenously defined by the analyst. In this study, only out-of-home activities are considered, and the decision process is modelled at the tour level; however, multiple tours on the same day are possible. Tours with multiple activities are included in the model; in this case only the choice of the principal activity is modelled.

For each tour generated by an out-of-home activity episode, individuals are supposed to make a two-step choice: first they choose the type of activity to be pursued and then they schedule the chosen activity within a specific time frame. Four activity types are possible: shopping, leisure, personal business and pick-up/drop-off of passengers. The scheduling is intended to be the decision regarding the specific time frame when the activity is executed; in this model time frames are defined with respect to the main activity of the day.

The main activity of the day is by definition the work activity for working days, and it is the activity with the longest duration for non-working days. For working days the possible time frames are: morning tour, morning commute, evening commute and evening tour; work-based tour to home during lunchtime is also considered. The choice set for non-working days is similar; the main difference consists of the possibility of executing secondary activities on the outbound and return leg of the principal tour. The combinations of activity types and of time frames define the alternatives in the choice set. A total number of 38 alternatives are modelled. The availability of each alternative for a given individual is then determined on the basis of the activity patterns reported during the entire survey period. In particular, if an activity type was never reported by a respondent during the six-week period, this activity type is considered not available for the individual. Shopping alternatives are available from 8 a.m. to 6 p.m., Sundays excepted.

The variables used in model estimation are (1) socio-economic variables, (2) level-of-service variables and (3) individual activity involvement variables. Socio-economic variables included are age (by categories), number of young children in the household, number of working hours per week, employment status and sex. Level-of-service variables are included through the feedback (logsum – see also Chapter 9) from a mode choice model.

Each individual is supposed to know the average durations of the possible activities to be performed and the travel times and travel costs to reach the relative locations. Pattern related variables are defined on three different time horizons: the day, the week and the entire survey period. Durations are estimated across activities, and different coefficients are estimated for different time frames; the remaining pattern related variables are estimated as activity specific.

In the discrete–continuous choice model of activity type and activity duration, time is allocated to an activity only if the activity is chosen by the individual. The explanatory variables are assumed to be only the result of past choices. The model is sensitive to both work duration and time availability before and after working activities. Day-to-day dynamics are explained by the variable that counts the number of days passed since the same activity was performed last. Individual past history is also represented by the number of days spent at home since the beginning of the survey period.

Freight module in TRANS-TOOLS

TRANS-TOOLS was developed for the European Commission and covers both passenger transport and freight (Burgess et al., 2008). General features of the freight model are:

1. It makes maximum use of the trade information observed for the base year 2000: region-to-region flows per type of commodity and transport mode.
2. It models the link between economy (production/consumption sectors) and flows of goods).
3. It models the links between the transport demand and infrastructure networks in both generation and modal split phases of modelling with a high level of detail.

4. It includes a logistic module to simulate the transport chains on both land–land and land–sea–land relations.
5. It makes a good assessment of the various policy options possible, at different levels, including also the infrastructure development policy, pricing and so forth.

The trade model considers the economic relations between regions of different countries by production and consumption (by sector), where the generalized cost of transport (or, simplified, distance) between region 'weights' the related trade flows. The trade model provides (explains) the volume of flows of goods between each O–D pair by type of commodity for all transport modes, for each segment of the market. Main input information for the calibration of the trade model consists of:

1. aggregated region-to-region (country for outside Europe) trade flows (all modes) by type of commodity;
2. production and consumption factors by sector, for each region (country for outside Europe);
3. average generalized cost of transport for each O–D relation (or, simplified, distance, if this is the case).

In the modal split model the modes considered are: road, rail, inland waterways and sea. Different O–D relations are considered in different segments of the market (defined on the basis of total volume and distance for each type of commodity), and cross elasticities for time and generalized costs of different modes have been estimated.

The logistics module is motivated by the observation that freight transport measured in ton kilometres or vehicle kilometres usually grows faster than when measured in tonnes (though this is not necessarily true for each shorter period – see Chapter 4). An explanation for this phenomenon is increasing logistics services. The module can be used to determine unimodal freight transport matrices for a base year and for a future year. This transport matrix is input to the assignment algorithm. The freight model uses socio-economic data as well as time/costs as input. The socio-economic data are output from the economic model, and the time/costs result from the assignment.

15.8 What can and what can't a model be used for?

The main reason for using models is that they provide an insight into the magnitude of effects of autonomous developments or policy measures on transportation and on indicators that convey something about the quality characteristics or the effects of transportation. Examples of quality characteristics and effects are the emissions of hazardous substances, the level of congestion or the question as to which roads have a lesser capacity than the expected demand. This may involve the effects of demographic, economic, spatial or infrastructural developments.

The question whether or not it is advisable to answer a research question or solve a policy issue by means of model simulations cannot be answered in a general way. In some situations, it may be sufficient to use the results of earlier empirical research or model simulations carried out earlier. In other cases, new model simulations may be required. Much depends on the amount of time available for answering the question, and on whether any (and if so how much) money is available for new studies and what the outcomes will be used for. The greater the effect of the answer, the higher the expenditures (in time and money) that will be made available, and the sooner model simulations will come into the picture. Put generally, it is important to weigh possible extra costs against the enhanced quality of the answer. For some questions, a rough estimation of the magnitude of effects may suffice, whereas for other questions a higher quality is required.

Furthermore, it is important to set up a structure of a model corresponding to the real world. The structure of a model should be compared to a schematic conceptual representation of that particular part of reality that is relevant for answering the questions. An example elucidates this. Suppose policy-makers want to know the effect of the acquisition of more fuel-efficient types of cars on the emissions of noxious substances. The effect of more fuel-efficient cars is, on the one hand, a decrease per kilometre in the average fuel use and in some emissions and, on the other hand, an increase in the number of kilometres, as the fuel costs of more fuel-efficient cars are, on average, lower. A model which first models car ownership and car use and then translates the fuel use and emissions back to car ownership and car use without considering the changes in the fuel use disregards the second phenomenon. Therefore it will overestimate the effect of a more fuel-efficient car fleet on the reduction in fuel use and emissions (see also Chapters 2 and 6 for such relationships). Needless to say, a model that does take into account the effect of decreasing fuel costs on car use due to the shift toward more fuel-efficient cars does not have this drawback. In another example, suppose that policy-makers want to know the effect of ABS on traffic safety. Presumably, ABS will lead to different driving behaviour. A model that does not take this into account overestimates the positive effect of ABS on traffic safety. Put differently, researchers who aim to support policy-making first need to get an idea of what will change in actual practice and how these changes will affect human behaviour, and next need to establish whether these changes can adequately be modelled by the model they would like to use. This is an important advantage of the use of models, or rather of considering the use of models: it stimulates thinking about the way in which the problem is put together.

Besides, it should be realized that most models are better able to study the effect of relatively small changes than to study the effect of major changes. The effect of an increase in fuel prices of half a euro can reasonably well be estimated by using the current models, whereas that would not hold for an increase by five euros.

An important advantage of using models is that they enable us to make the research or policy questions more explicit. The researcher who poses the question needs to make a general question concrete. A question like 'What happens if fuel prices increase?' is very vague. A model forces the researcher to make the question explicit; otherwise simulations are impossible. Therefore the researcher needs to answer questions like: Which kinds of fuel are becoming more expensive? And to what extent? When will the price increase become effective? For which year, or for which years, does one need to get answers? Which effects does one want to know? Only the effect on car use? Or also the effect on car ownership, the use of other means of transport, moving behaviour, emissions of hazardous substances, or noise pollution? Of course, a research question always has to be concretized, but the use of model simulations can be very helpful in doing so.

A further advantage of using models is that, usually, a model can be used rather quickly to compute the effects on variants for socio-economic developments and policy variants. Because of the fact that expressly only a few variables are changed and the other factors are kept constant, comparability of those variants is warranted and the right conclusions can be drawn as to the effects of the supposed changes.

One more advantage, to conclude with, is that the comparability of the results from individual studies can be enhanced by using models. Various studies on the same issue often (seemingly) arrive at different results. For example, many studies have been carried out on the effects of increases in fuel prices. The effects are rather diverse: when expressed in the fuel price elasticity of car use, the range goes from less than −0.1 to more than −1.0, or a difference of more than a factor 10. Models that can be used to compute the effects of increasing fuel prices teach us that the effects will be larger for people with lower incomes when compared to people with higher incomes.

A study carried out in Portugal in 1980 would therefore, in principle, show a higher price sensitivity than would a study carried out in the US in 2000. Furthermore, models show that the effect of an increase in (solely) the price of petrol will be less if other, cheaper kinds of fuel are available. So the effect of an increase in the price of petrol will be larger if no LPG is available and if the price difference when compared to diesel oil is limited. If we look at the effect of fuel price increases on the demand for fuel, model simulations teach us that, in the short run, the only effect is a decrease in car use. In the long run, however, a further effect will be that people acquire more fuel-efficient cars if fuel prices increase. In some cases, this phenomenon also explains part of the seeming differences in the study results. In the even longer run, a further effect may be that people will select a different place to work or to live, which will increase the effect on car use. On the other hand, in most countries incomes are gradually increasing, which partly counterbalances the effects of price increases. So model simulations can mean that researchers get a better insight into the causes of (seeming) differences between empirical studies.

In our opinion, a drawback of model use is that policy-makers and lawyers overestimate the importance of model simulations. A model is nothing more than a tool, and one should refrain from attaching too much absolute value to its results. Especially if the result of simulations is that certain government objectives are, or are not, achieved by a small margin only, these simulations have a great impact on policy-makers' way of thinking. It is hardly relevant whether a certain package of policy measures, according to the model and in a certain context, will only just, or will only just not, lead to achieving a goal in 2020, for example the objective concerning NO_x emissions. The message then is first and foremost that, in that context, the package will lead to NO_x emissions in the magnitude of the set aim. And, in the case of air pollution, lawyers often assume the results to have a great degree of certainty, even though these results are relatively uncertain. Researchers also sometimes attach too much weight to a model's results. Sometimes they hide behind the model and then make statements like 'Well, that is what the model has come up with.' Once again, a model is no more than a tool. The researcher answers a research question and in doing so is responsible for whether or not to use a model and, if so, in which way, and how to use and interpret the results.

15.9 Conclusions

The major conclusions to be drawn from the above are:

1. A model is a simplified representation of a part of reality.
2. Transportation models are first and foremost a means to gain insights

into the effects of various developments (like policy developments) on transportation.

3. There are many kinds of models. They can be classified, among other things, for the transport mode they focus on (e.g. passenger cars, aeroplanes), or the period of time they focus on, or the technical characteristics (e.g. dynamic versus static models), or the question as to whether or not they contain spatially varying data.
4. Whether or not it is advisable to answer a certain research or policy question by using a model cannot easily be established. It requires adequately weighing the available alternatives as well as the pros and cons of model simulations and the relevant alternatives.

REFERENCES

Bakker, P. and P. Zwaneveld (2009), *Het belang van openbaar vervoer: de maatschappelijke effecten op een rij* [The importance of public transport: a list of the social effects], The Hague: Kennisinstituut voor Mobiliteitsbeleid (Ministerie van Verkeer en Waterstaat), Centraal Planbureau, accessed 4 February 2012 at http://www.rijksoverheid.nl/ministeries/ienm/kennisinstituut-voor-mobiliteitsbeleid.

Barton-Aschman Associates and Cambridge Systematics (1997), *Model Validation and Reasonableness Checking Manual*, prepared for Travel Model Improvement Program, Federal Highway Administration, USA, accessed 4 February 2012 at http://www.ctre.iastate.edu/educweb/ce451/LECTURES/Validation/finalval.pdf.

Bates, J., M. Dasgupta, N. Daha, G.R.M. Jansen and M.J.M. van der Vlist (1996), *Audit of the Dutch National Model: Final Report*, Report INRO-VVG 1996-29, Delft: TNO INRO.

Ben-Akiva, M. and S. Lerman (1985), *Discrete Choice Analysis*, Cambridge, MA: MIT Press.

Burgess, A., T.M. Chen, M. Snelder, N. Schneekloth, A. Korzhenevych, E. Szimba, M. Kraft, M. Krail, O. Nielsen, C. Hansen, A. Martino, D. Fiorello and P. Christidis (2008), *Final Report TRANS-TOOLS (TOOLS for Transport Forecasting and Scenario Testing)*, Deliverable 6, Funded by Sixth Framework RTD Programme, Delft: TNO INRO.

Cirillo, C. and K.W. Axhausen (2010), 'Dynamic model of activity type choice and scheduling', *Transportation*, 37 (1), 15–38.

De Jong, G., G.J. Tuinenga and M. Kouwenhoven (2008), *Prognoses van het Landelijk Model Systeem: komen ze uit?* [Prognoses of the LMS: do they come true?], Contribution to the Colloquium Vervoersplanologisch Speurwerk, Santpoort, November, accessed at http://www.cvs-congres.nl/.

Geurs, K. and J. Ritsema van Eck (eds) (2004), 'Special issue: land-use and transport interaction models', *European Journal of Transport Infrastructure Research*, 4 (3), 247–249.

Gunn, H. and A. van der Hoorn (1998), 'The predictive power of operational demand models', Paper D14, prepared for the European Transport Conference (formerly PTRC), Loughborough University, 14–18 September, *Proceedings of the Seminar Transportation Planning Methods*, London: Association for European Transport, pp. 249–264, available at http://www.etcproceedings.org/.

Mabit, S. (2009), 'Danish preferences concerning alternative-fuelled vehicles', paper presented to the European Transport Conference, accessed 4 February 2012 at http://www.etcproceedings.org/paper/danish-preferences-concerning-alternative-fuelled-vehicles.

Mackett, R.L. (2009), 'Why is it so difficult to reduce car use?', paper presented to the European

Transport Conference, Noordwijkerhout, accessed 4 February 2012 at http://www.etcproceedings.org/paper/why-is-it-so-difficult-to-reduce-car-use.

Meurs, H. and R. Haaijer (2006), *DYNAMO: een nieuw dynamisch automarkt model voor Nederland* [DYNAMO: a new dynamic model for the Dutch passenger car market], Contribution to the Colloquium Vervoersplanologisch Speurwerk, Amsterdam, November, available at http://www.cvs-congres.nl/.

Ortúzar, J. and L. Willumsen (2011), *Modelling Transport*, 4th edition, Chichester: John Wiley.

USDOT, Federal Highway Administration (1999), *Guidebook on Statewide Travel Forecasting*, accessed 4 February 2012 at http://www.fhwa.dot.gov/planning/statewide/swtravel.pdf.

Van Mourik, H. (2008), *Verkenning Autoverkeer 2012* [An investigation of car traffic in 2012], The Hague: Kennisinstituut voor Mobiliteitsbeleid, Ministerie van Verkeer en Waterstaat, accessed at www.kimnet.nl.

Wegener, M. (2008), 'SASI model description', Working Paper 08/01, Spiekermann and Wegener Urban and Regional Research, Dortmund, accessed 4 February 2012 at http://www.spiekermann-wegener.de/mod/pdf/AP_0801.pdf.

Whelan, G., J. Crockett and S. Vitouladati (2010), 'A new model of car ownership in London: geo-spatial analysis of policy interventions', paper for the European Transport Conference, Glasgow, accessed 4 February 2012 at http://www.etcproceedings.org/paper/a-new-model-of-car-ownership-in-london-geo-spatial-analysis-of-policy-interven.

Index

Aarts, H. 34
Aarts, L. 256, 274
abilities (NOA model) 24, 25, 26
acceleration 150, 235–6
access 86–7, 209, 276
accessibility 5, 10, 207–24
  benefits 95–6
  components 209–10
  definitions 13, 207, 208–9
  impact on the economy 13–14
  impact of technology on 11, 12
  measures 207–8
    choosing and using 215–17
    communicability 217
    examples 218–22
    impact of land use on travel behaviour 91–2
    interpretability 217
    operationalization 211–15, 217
  perspectives on 208–9
  transport policies 300, 301
accidents 9, 117, 191, 286, 300–302
acidification/acid rain 12, 168, 175, 234
action point models 151
action space 42–3
active safety 11, 12, 192, 194
active travel 247
activity duration, and travel time 44
activity programmes 45, 46
activity-based approach 41
activity-based models 223–4, 357–8
Adams, J. 117, 287
adaptive approaches, futures research 321–5
adaptive behaviour 116
adaptive cruise control (ACC) 198–9
adoption dynamics, technology 173–4, 199–200
advanced driver assistance systems (ADAS) 192
advisory (ISA) systems 192
affect 31–2
age
  of cars 9
  traffic safety 257, 260, 261, 262, 272
  travel behaviour 7, 21, 22, 23, 32
aggregated models 356, 361–3
aggression, and road safety 257–8
Agusdinata, D.B. 318
Ahmed, K. 149
air pollution 5, 168, 174, 175
  environmental impacts 234–6
  estimates of external costs 286
  technological innovations 175–9
air pressure pulsations 187
air transport
  costs 9
  environmental impacts 228, 231, 237, 239
  estimates of external costs 286
  freight transport demand 53, 57
  income elasticity 39
  likelihood of death from 9
  technologies 186, 189–91
  travel time 9
airbags 256
Albrantes, P.A.L. 105
alcohol, crash risk 257
alcohol ignition interlocks 194
all-or-nothing (AON) methodology 362
alternatives, assessment of 333, 345
Annema, J.A. 348
anthropogenic warming 230–31
anthropological explanation, travel time 106
appraisal, transport policy 329–50
  multi-criteria analysis 344–8
  social cost benefit analysis 330–37
  utility-based accessibility measures 222
  valuation of environmental effects 337–40
  valuation of time travel 340–44
Arunotayanun, K. 120
assignment, aggregated models 361, 363
attitudes 28–9, 30, 33, 84, 87–8
attractive environments 82, 243
Australia 147, 261, 267, 292, 300
authority constraints 42, 44

automatic number plate recognition (ANPR) 198
automatic vehicle identification (AVI) 133
avoidance costs 287

backcasting 312, 319–21
Badoe, D.A. 86
Bagley, M.N. 87
balancing factors, accessibility measures 219
Balcombe, R. 115
Bamberg, S. 31
Banister, D. 24, 283, 319
base locations, time geography 42, 43
basic flow theories 145
basic policy, DAP process 323–4
Bates, J. 115
beacon transponder systems 198
behavioural choices, travel *see* travel behaviour
behavioural geography 40
behavioural theory, traffic flow models 148
Belgium 108, 242, 260, 261
Ben-Akiva, M. 220
benefits *see* co-benefits; external benefits; marginal costs/benefits
best estimate models, policy analysis 313
Best Practices in Road Safety 274
biases, data collection 264
bicycles 89–90, 111, 260
  see also cycling
bio fuels 182–3, 184, 185, 233, 246, 247, 248
biodiversity 227, 248
"black swans" metaphor 322–3, 324
blood alcohol content (BAC) 257
Blue Map/Shifts 317
Bohte, W. 87
Börjeson, L. 312
Botma, H. 137
bottom-up (BUM) methods 239
bounded rationality 167
Bovy, P.H.L. 154
Brackstone, M. 151
BRIC countries 53, 67, 245, 247, 292
Brons, M. 86
Brundtland Report (1987) 228
Buckley, D. 135, 136
Buehler, R. 23–4
built environment, impacts of transport 227
B–C criterion, in SCBA 333–4

Cadot, O. 348
California 116, 175–6, 185, 223–4, 228
Cao, X. 84, 87
capability approach 296
capability constraints 41, 44
capacity (-drop) 8, 12, 135, 136, 137, 139–40, 142, 152, 157
car ownership 21, 22, 86, 87, 106, 113, 114, 242–3
car use
  changes in 19
  cost-benefit analysis 294, 295
  effort resistance factors 114, 115–16, 117
  elasticities 110, 112–13, 242, 360
  environmental impacts 12
  German/US comparison 24
  money cost components 111
  motivational factors 30, 32, 34–5
car-following models 149–51
carbon dioxide (CO_2) 176, 180
  emissions 12, 14
    data and trends in 230–34, 245
    mode of transport 239–40
    price elasticities 243
    reduction 182, 183, 184, 186, 244–5, 245–7
    rights 71
    standards/policies 247, 248–9
carbon monoxide (CO) 11, 14, 175, 176, 229, 230, 241
cars
  average age of 9
  fuel-efficient 11, 112, 113, 181–2, 183
  levies and taxes 242
  noise emissions 12
  road fatalities 260
  safety technology 191–2, 193, 256
  status and choice of 7
  see also diesel cars; electric vehicles; petrol cars
catalyst technologies 176, 177, 178, 179
cellular positioning 198
censored observations 137–8
centroids 355
Cervero, R. 86, 87
Chandler, R.E. 150
chapters 189
choice, and travel *see* travel behaviour
choice probability model 220–21
Chorus, C.G. 116, 221
circular action spaces 42–3
climate change 12, 174, 179–86, 286
  see also global warming

climate neutral energy system 248–9
Club of Rome 166, 228
co-benefits 247
co-evolution 170
coefficients 354, 364
cognitive disorders 257
cognitive dissonance reduction 25–6
cognitive ergonomics 276
collision warning systems 192
comfort 5, 9, 115, 343
compensation criterion 289
competition effects 219
complementarity 38
composite headway models 135–6
computational experiments 318
conceptual model, travel behaviour 24–8, 79
congestion 5, 8, 12, 157, 174
  charging 19, 292
  costs 105
  delay 286
  technology 195–9
  time 103
conjoint analysis 339
consistency, scenario assessment 313
consolidation, supply chain partners 73
constant money cost budgets 111–12
constant time budgets (TTB) 95–6, 105–8
constraints 7, 33, 37, 41–2, 44
consumer surplus 92, 220–21, 331–2
consumptive expenditures, optimal allocation 36–7
containers 163
context, valuation of environmental effects 339
contextual factors, travel behaviour 33
contingent valuation 339, 340
continuity equation 131–3, 140
continuous models 355–6
continuum traffic flow models 154
cooperation, supply chain partners 73
cooperative adaptive cruise control 199
Corporate Average Fuel Economy (CAFE) standards 180–81
correlation 354, 355
cost control 61–2
cost overruns 297
cost sensitivity function 218
cost-benefit analyses (CBA) 94, 289
  car use 294, 295
  criticisms of 295–6
  road infrastructure travel time savings 105
  transport innovation 174
  transport policy 292–3, 316
  travel behaviour 28–30
  see also social cost benefit analysis
cost-effectiveness 298
costs
  globalization and changes in 69–71
  improved road networks and reduced 11
  road crashes 265–6
  and transport resistance 8–9
  see also external costs; generalized logistics costs; generalized transport costs; inventory costs; marginal costs/benefits
country differences, and study results 89–90
coupling constraints 41–2, 44
Cowan, R.J. 135
crash cushions 193
crash helmets 191, 193, 256, 260
crash tests, standardized 191
crash-prone theory 270
crash-protective roadsides 193
crash-protective vehicle design 193
credibility, scenario assessment 313
critical gaps 138
cross elasticity(ies) 38, 39, 110–11, 360, 376
cross-section data 357
Cullen, I.G. 43
cultural factors 10, 24, 79, 84, 89, 90, 113, 171, 271
cycle-by-pass measures 179
cycling 33, 39, 79, 103, 114, 115, 116, 117, 247

Daganzo, C.F. 147, 155, 156
DALYS (disability adjusted life years) 236–7, 265
damage-only crashes 264
danger, measuring 263–6
Dargay, J. 85, 86, 114, 242
data collection 224, 264
data communication systems 198
data mining 318, 355
daytime running lights 193
De Brucker, K. 344, 347
De Jong, G. 108, 220
Decade of Action for Road Safety 266
decarbonization 245
decentralized distribution structures 71
decision-making, habit in 110
decoupling 68, 245
dedicated short range communication (DSRC) 198
deep uncertainty 321

Delucchi, M.A. 287
demand function 37–40, 330, 331
demand and supply, transport volumes 102
demand uncertainty, modal choice 65
demography 10, 20–24, 26, 44, 83–4, 87
density(ies)
  traffic flow *see* traffic flow theory
  and travel behaviour 45, 46, 80–81, 84, 85–6, 88–9
deregulation 70–71
derived demand 11, 24, 59, 80, 358
descriptive models 354, 355
design quality, and travel behaviour 82, 86
desired speed distributions 137–8
deterministic models 151–3, 363
diesel cars 176–8, 182, 230, 239, 241, 242
diffusion of innovations 200
Dijst, M. 44
Dingus, T.A. 259
direct effects, infrastructure investment 358
direct elasticities 360
direct factors, mode environmental performance 239
direct travel time losses, congestion costs 105
disaggregated accessibility measures 223–4
disaggregated models 356, 362–5, 366–76
discomfort 115
discount rate (SCBA) 333
discrete models 355
distance estimation 138
distance headways 129, 130
distances
  air pollution source and receptor 234, 235
  freight transport 58, 63
  travel behaviour 22, 23, 39–40, 82, 86–7
distraction 187, 258
distribution models 361–2, 364
distribution systems/structures 71
disutility 81, 95, 96, 210, 215, 220
Dong, X. 221
double counting 336, 346
doubly constrained models 362
driver sub-tasks 148, 149
driving behaviour 5, 12, 148, 179, 241, 258–9, 269
driving experience 257
driving licenses 20–21, 22
drugs, crash risk 257
Duncan, M. 87
Dutch National Model System 367–72
dynamic adaptive policy 323–5
dynamic models 356–7
dynamic traffic management 195–6, 244
DYNAMO 368, 372–5

early adopters 200
economic cycles, product sensitivity 58
economic growth 19, 26, 66–9
economic integration, global freight flows 68
economic perspective, travel behaviour 35–40
economics *see* evolutionary economics; neo-classical economics
economies of scale 57, 60, 61, 62, 64, 69, 73, 170, 174
economists 9–10, 167
economy
  impact of accessibility on 13–14
  transportation models and 358–9
Edie, L. 133
effectiveness, of policies 298
efficiency, of policies 295, 298
effort resistance factors 114–18
El-Geneidy, A. 86
elasticities 102
  resistance factors 102
  transportation models 359–60
  see also income elasticity; price elasticity; travel time elasticity
elderly 22, 262, 272, 294
electric vehicles 169–70, 183, 184–5, 249
elliptic action spaces 43
Elvik, R. 347
emissions 55
  standards/policy 176, 177, 178, 187, 241, 242, 247, 300–302
  see also individual pollutants
emotions 25, 257–8
end-of-pipe technologies 180
endogenous variables 354
energy
  climate neutral systems 248–9
  limitations 94
  security 286, 287
  sustainably produced 94
  use 55, 58, 85–6, 180, 229, 239, 243
environment(s)
  adjustment to, and road safety 276–7
  attractive 82, 243
  discomfort factor 115
  impact of technology on 11
  infrastructure and quality of 287–8
  transport policies 300–302

environmental impact assessments 296, 329
environmental impacts 14, 227–50
  air pollution 234–6
  CO_2 emissions and oil use 230–34
  infrastructure 288
  land use and travel behaviour 93
  long-term challenges 244–9
  moral concerns and travel behaviour 31
  noise 236–8
  overview of problems 228–30
  performance of modalities 238–40
  policy instruments 240–44
  valuation of 337–40
  vehicle use 12
environmental values 33
Environmentally Sustainable Transport project 319
equilibrium speed 143
equity
  in SCBA 334–5
  and transport policy 284, 293–6, 298, 301
EURO NCAP 191, 192
European Automobile Manufacturers' Association 174
European Commission 110, 182, 183, 236, 254, 274, 375
European Conference of Ministers of Transport (ECMT) 178, 179, 195
European Enhanced Vehicle-safety Committee (EEVC) 192
European Union (EU) 13, 337
  ambition to reduce road deaths 254
  emission standards 176, 177, 187
  emissions 182, 229–30, 244–5
  energy use 229
  freight transport 52–3, 66, 67, 120
  public transport subsidies 294
  safety initiatives 191–2
Eurostar 305, 306
evaluation methods 94
evaluation of policy 13, 247–8
Evans, G.W. 117
Evans, J.S. 235
evidence-based policies 271
evolutionary economics 167–74
exercise 93
exhaust gas recirculation (EGR) 178
exogenous variables 354
explanatory models 354, 355
explanatory variables 365
explorative scenarios 312
exploratory modeling and analysis (EMA) approach 317–18
external benefits, transport policy 288
external collaboration, economies of scale 62
external costs 71, 285–8, 291
external effects 167, 168, 285–8
external forces, policy analysis 307, 313, 314

fatigue 258
feasibility of behaviour (NOA model) 24, 25
Feitelson, E. 173, 174, 185, 198, 200
Feo, M. 118, 120
financial aspects, land use and transport 93–4
financial crisis, transport distances 58
first-generation biofuels 182–3, 247
"fit" technologies 169
fixation in time and space 43–4
fixed costs 9
flexibility
  futures research 321–5
  hybrid networks 73, 74
  transport policies 299
flexible activities, time geography 42, 43
floating car data 127, 133, 195–6
flow *see* traffic flow theory
Flyvbjerg, B. 297, 312
focal goal-frame 32
Forbes, T. 150
forecasting 66, 353, 356, 365
forgivingness, road safety 277
forward-looking scenarios 312–17
France 106, 165, 181, 242, 244, 260, 261, 267, 293, 348
Frank, L. 86
free flow phase 145, 147
free riders 339
free speed 137–8
free-flow time 103
freight transport
  environmental challenges 245
  external benefits 288
  models 359
  technologies to reduce emissions 186
  see also mode of transport
freight transport demand 11, 51–75
  determinants 51, 59–66
  drivers of change 51, 66–74
  indicators of 51, 52–9
  resistance factors 118–21
Fridstrom, L. 347

fuel cell vehicles 184–5
fuel costs 9
fuel efficiency 9, 179, 180–82, 190, 247
fuel efficient cars 11, 112, 113, 181–2, 183
fuel price elasticities 38, 112–14, 242
fuel standards 183, 300, 301
Fujii, S. 34
full-speed-range ACC systems 199
functionality 34, 277
fundamental diagrams 140–41, 143, 147, 155–6
fundamental risks, road traffic 255–6
futures research 305–26
  backcasting 319–21
  exploratory modelling and analysis 317–18
  flexible and adaptive approaches 321–5
  forward-looking scenarios 312–17
  transport policy analysis 307–12

gain goal-frame 32
gains (travel time), valuation 341–2
gap acceptance 138
Gateley, D. 114
GDP elasticity 105
Geilenkirchen, G.P. 113
gender
  and crash risk 257
  travel behaviour 20–21, 22, 23, 32
general packet radio system (GPRS) 198
generalized logistics costs (GLC) 60, 63, 64, 65
generalized traffic flow variables 133–4
generalized transport costs (GTC) 8, 38, 80, 95, 96, 102
  see also monetary costs
generation models 361, 365
generic problems, road safety analysis 272
genetics, and travel time 106–8
geographers 10
geographic information systems (GIS) 41, 223
geographical perspective, travel behaviour 40–45
geographical scale, differences in study results 88–9
Germany 24, 106, 181, 243, 261, 267, 293
Geroliminis, N. 155, 156
Geurs, K.T. 216, 217, 221, 222, 288, 332
Gilbert, R. 233
Giuliano, G. 86
Givoni, M. 115
glare sensitivity 257
Glassbrenner, D. 256
Glista, D.J. 287
global navigation satellite system (GNSS) 198
global positioning systems (GPS) 45, 174
global targets, CO_2 emissions 246–7
global warming 168, 230–31
globalization 69–71
goal-framing theory 32–3
Godson, V. 43
Goodwin, P. 41
Gordon, D. 166, 185
Gordon, I. 85
Gould, P. 207
gravity-based accessibility 212–13, 218–19
greenhouse gases 182–3, 230, 231, 244–5
  see also individual gases
gross densities 80
gross distance headway 129
gross time headway 128
Gruebler, A. 165
Gunn, H. 108

habit(s) 30, 33–5, 110
Haddon matrix 270–71
Hägerstrand, T. 41
Handy, S.L. 207
Hanly, M. 85
Hansen, W.G. 218
headway distributions 135–7
health 93, 234, 236, 247, 257
heavy duty vehicles 178, 179, 187, 188, 192, 194, 230, 247
hedonic goal-frame 32
hedonic price approach 338, 340
Heinen, E. 115, 117
heterogeneity 53–5, 116, 167
Hicks, J.R. 289
high-income countries 266, 267
high-income groups 7, 23, 39, 40
high-risk locations 273
high-speed trains 106, 237, 239, 343–4
higher-speed technologies 165
homogeneity, road safety 277
Hoogendoorn, S.P. 135, 136, 138, 154
horizontal collaboration 62
household characteristics, disaggregated models 365
household variables, travel behaviour 19, 21, 22–3, 26, 27, 40, 84
hybrid networks 73, 74
hybrid vehicles 183–4
hydrogen (H_2) 184, 185

Hymel, K.M. 109
hypotheses, EMA approach 318
hypothetical compensation, in SCBA 335

illness, and crash risk 257
"images of the future", in backcasting 319–20
impact assessments 193–4, 296, 313, 316, 329
impedance functions 218, 362
imperfect transport systems 167
implementation, of transport policies 298–9
implementation research 321
in-vehicle time 103, 104–5, 215, 342
in-vehicle-only systems, OBUs 198
inattention, and near crashes 259
income
  and safety 191, 266, 267, 269
  travel behaviour 7, 21, 22–3, 32, 36, 38, 40, 84
  travel time valuation 341
income elasticity 39–40
indirect effects, infrastructure investment 358–9
indirect factors, mode environmental performance 239
indirect travel time losses, congestion costs 105
individual, in disaggregated models 363
individual component, of accessibility 210, 214
individual variables, travel behaviour 19–47, 84
induced phase transitions 145
induced traffic hypothesis 108–9
inductive loops 127–8, 130, 132
inferior goods 39, 40
information 9, 25, 116, 191, 195, 244, 343
information technology (IT/ICT) 62, 69–70, 108, 192, 224, 344
infrastructural adaptation 268
infrastructure
  external costs due to 287–8
  and GTCs 8
  improvements, and spill-overs 336–7
  investment 300, 358–9
  modal split and traffic volume 244
  projects, and cost overruns 297
  spatial effects 358
  underestimation of economic benefits 336
infrastructure-based accessibility measures 211, 212, 214, 223
Ingram, D.R. 218
injury 255–6, 263, 264, 287
innovation 167–70
Insight, B. 224
instantaneous variables 129, 130, 133
instrumental motives 25, 32
intelligent speed adaptation (ISA) 12, 192–5
intelligent transport systems (ITS) 192, 198–9
intentional errors 275
intentional violations 258–60
intentions 28–30
Intergovernmental Panel on Climate Change (IPCC) 230
internal combustion engines (ICEs) 170, 183–4, 231, 234, 235
internalization, external costs 71, 291–2
International Civil Aviation Organization (ICAO) 189, 190–91
International Energy Agency (IEA) 233
international freight flows 68
international projects 337
International Transport Forum 274
intrinsic value, of travel 24
intuitive hypothesis 150
inventories 61–2
inventory costs 65, 66
investing society 268
Iseki, H. 103

Japan 176, 261, 267, 293
Jepsen, M. 137
jet airliners 163, 190
Johansson, P.O. 338
Jordans, M. 63
"just-in-time" concept 11, 61, 62

Kaldor, N. 289
Kaldor–Hicks efficiency criterion 289, 335
Kelly, C. 292
Kenworthy, J.R. 83, 85, 89
Kerner, B.S. 144, 145, 147
kilometres per day 22, 23, 39–40
kinetic energy 255
Kitamura, R. 87
knowledge 25, 167, 169
Kockelman, K. 86
Koornstra, M.J. 274
Kwakkel, J.H. 318
Kwan, M.-P. 41
Kyoto Protocol Agreement 245

Lal, D. 106
Lam, T.C. 116
land-based freight transport 66, 67

land-use
accessibility in relation to 208–9
impact on travel behaviour 78–97
differences in study results 88–90
empirical examples 84–8
environment versus accessibility 95–6
indicators and evaluation methods 91–4
interplay between variables 83–4, 85
theoretical foundation 79–84
and transport policies 13
land-use component, of accessibility 209–10, 214, 215
land-use patterns 8, 13, 78, 90
land-use planning 243
land-use policy, use of logsum measure 222
land-use transportation interaction (LUTI) models 358, 359
land-use variables 80–84, 361
landscape developments 172
lane changing modelling 149
lane departure warning system 12
latent errors 275, 276
lateral driver tasks 148, 149
lead emissions 241
leaded petrol 176
learning society 268
Lerman, S.R. 220
Level 1 uncertainties 310, 311, 312, 321
Level 2 uncertainties 310, 311, 312, 317
Level 3 uncertainties 310, 311, 312–18, 324
Level 4 uncertainties 310, 311, 312, 318, 321–5, 329
Levinson, M. 163, 164
Li, Z. 119
life cycle assessment (LCA) 240
lifestyles 27, 35, 84, 87–8, 245
The Limits to Growth 166–7, 228
linear action spaces 43
Linneker, B.J. 207
Litman, T.A. 84
liveability 14, 196, 227, 301
loading factors, traffic performance 58
local variables 129, 130, 133
location-based accessibility measures 211, 212–13, 214, 223
locations of activity 8, 42, 43, 79, 80
locations of production 61
lock-in 170
log-normal distributions 135
logarithm unit decibel (dB) 187
logistics 11, 59–66, 71, 72, 73, 239
logsum accessibility measure 220–22
logsum variable 364–5
London 19, 207, 243
long-distance travel 106–8, 319, 320
long-term choices, travel behaviour 27
long-term fuel price elasticity 242
longitudinal driver tasks 148, 149
losses (travel time), valuation 341–2
low-carbon fuels 182
low-carbon transport 246, 248
low-income countries 191, 266, 269
low-income groups 32, 39
luxury goods 39, 40

Mabit, S. 357
McCubbin, D.R. 287
McDonald, M. 151
Mackie, P.J. 292
macro-level (NOA model) 26
macroscopic feedback control strategies 156
macroscopic fundamental diagram 155–6
macroscopic traffic flow models 151–4
macroscopic traffic flow variables 129–34
Maglev technology 165
Manaugh, K. 86
mandatory (ISA) systems 195
Marchetti, C. 106
marginal costs/benefits 196, 289, 290, 291, 294
market failure 284, 285–8
market imperfections 336
marketing literature 244
Marshall, J.D. 235
mass differences, road safety risk 256, 263
mass individualization 71–3
MASTIC model 45
mathematical formulas, in models 354
Maut system 243
maximum flows 139
Meadows, D.H. 166, 228
Mellios, G. 184
mental processes 34
mental strain 117
Metz, D. 22
Michaels, R. 151
micro-level (NOA model) 26–7
microscopic traffic flow models 149–51
microscopic traffic flow variables 127–9
mid-term choices, travel behaviour 27
middle-income countries 191, 266, 269

Miller, E.J. 86
Miller, H.J. 41
Minderhoud, M. 140
minimum headways 150
misperceptions, and habit 34
mixed land use, and travel behaviour 81–2, 86
mixed logit model 365
mobile phones 195–6, 258
mobility choices, travel behaviour 27
modal split models 362, 376
mode of transport
  choice 30, 63–6, 82–3, 241, 242, 243
  environmental performance 238–40
  monetary costs and switches in 119–20
  payment of total social costs 294
  and road safety 260, 269
  transport performance by 57
  see also individual modes of transport
model validation 315, 366
models 353–80
  aggregated 361–3
  disaggregated 363–5
  elasticities 359–60
  examples of 366–76
  kinds of 354–9
  traffic flow 147–54
  uses and non-uses 376–9
  validation 366
Mokhtarian, P.L. 84, 87
monetary constraints, and choice options 7
monetary costs 9, 38, 105
  freight transport and resistance factors 119–20
  road crashes and injuries 266
  role in passenger transport 37, 111–14
monitoring system, DAP process 324–5
mono-causal model, road crashes 270
moral concerns, and travel behaviour 30–31
morbidity rate, road crashes 264
mortality rates, road crashes 9, 264, 266, 267, 268
motivation, travel behaviour 24, 25, 28–33
Motor Vehicle Pollution Control Board 176
motorcycles 192, 260
motorways 9, 12, 263, 268
multi-criteria analysis (MCA) 94, 296, 316, 335, 344–8
multi-lane traffic flow facilities 147
multi-level perspective, system innovation 171, 172
multinomial logit (MNL) model 220–21, 363–5
multiple routeing 362
Myers, J.A. 143
Nader, R. 166
Næss, P. 87
Nakicenovic, N. 165
National Travel Surveys (NTS) 20
natural environment, impacts of transport 227
nature areas, valuation of 338–9
needs and desires 7–8
  freight transport 51–75
  passenger transport 19–47
needs (NOA model) 24, 25, 26
negative coefficients 364
negative cross-price elasticity 38
negative exponential function 218
negative impacts 4, 13, 91
  government policy 293–4, 305
  technology to reduce 163–202
  see also congestion; environmental impacts; safety
negative scenarios 324
neighbourhood design 82, 86, 243
neo-classical economics 300
net densities 80
net distance headway 129
net present values 333
net time headway 128–9
Netherlands
  car travel time elasticities 108
  CBA, transport policy 293
  congestion costs 105
  fuel efficiency 181–2
  GDP and freight transport 69
  importance of bicycles 89–90
  improvement in road safety 267–9
  modal split 242
  road fatalities 260, 261, 268
  SCBA in policy analysis 348
  societal costs, road crashes 266
  Sustainable Safety 273–7
  transportation models 367–76
  travel behaviour 20–24, 39–40, 44
network degeneration 156–7
network dynamics 155–7
network fundamental diagram (NFD) 155–6
Newman, P.W, G. 83, 85, 89
Niemeier, D.A. 207
Nijkamp, P. 344
nitrogen (N_2) 176, 178
nitrogen oxides (NO_x) 11, 12, 14, 175, 176, 178, 179, 229, 230, 234, 241, 338
"no-regret" measures 246, 322

NOA model 24–8, 35–6
noise
emissions standards/regulations 187, 189, 241–2
environmental impacts 12, 169, 236–8, 249, 337
estimates of external costs 286
spatial and temporal distribution of traffic 5
technology to reduce negative impacts 187–91
valuation of environmental effects 338
noise barriers 187, 238
"non-emissions-related liveability" impacts 229
non-linear valuation, environmental quality 339
non-oil-based fossil fuels 233
non-renewable raw materials, use of 249
non-spatial models 355–6
non-use phase of vehicles 287
norm activation model (NAM) 31, 32
normative concerns, travel behaviour 30–31
normative goal-frame 32
normative scenarios 312
Norway 196, 243, 261, 293, 347
novice drivers 257
Nussbaum, M.C. 296

OECD 195, 274, 319
OECD countries 67, 112, 176, 241, 245, 270
oil/dependency 166–7, 168, 174, 179–86, 229, 231, 232–4, 245–7
on-board diagnostic (OBD) systems 178, 179
on-board electronic stability programmes 194
on-board reformation process 185
on-board units (OBUs) 198
one pipe regime 147
operational factors, mode environmental performance 239
opportunities (NOA model) 24, 25, 26
optimal allocation, consumptive expenditures 36–7
optimization, SPITS MODEL 59–60
optimizing technologies 199, 201
option values 92, 332
origin-destination (O–D) matrices 355, 361, 365
Ortúzar, J. 353, 365
outcomes, of SCBA 333
outcomes of interest, policy analysis 308, 313, 314, 315
overtaking lanes 194
"own" price elasticity 38
oxidation catalyst technology 177

PACTS report 274
Pan, H. 86
panel data 357
parametric estimation 137–8, 150–51
Pareto efficiency 289, 335
parking charges 243
Parry, I. 294
partial differential equations 154
particulate filters 177
particulate matter 11, 14, 175, 176, 177, 178, 229–30, 234
Pas, E.I. 24
passenger transport
environmental challenges 245
models 359
role of travel monetary costs 111–14
role of travel time in 102–11
see also mode of transport
passive safety 11, 12, 192, 193–4, 268, 269
path dependency 170
paths (time geography) 41
Paulley, N. 110, 115
peak oil 233–4
"peaks in distribution" 273
Peden, M. 266
perceived behaviour control (PBC) 29
perception capabilities 138
"perception of problems", technology adoption 173–4
Perl, A. 233
person-based accessibility measures 211, 213, 214, 223
PESASP 44–5
petrol cars 169, 176, 239, 241
physical direction separation 277
physical effort 114, 115
pipeline inventory costs 65
Pipes, L. 149, 150
Pirie, G.H. 217
Pivo, G. 86
plausibility, futures research 305, 313
plug-in hybrids 184
Polak, J.W. 120
policy(ies)
analysis 307–12
and developments in transport 4
impact on land use patterns 90
see also evidence-based policies; pricing policies; time policies; transport policies
political economy model 173–4
political explanations, infrastructure project problems 297

political feasibility, road charging 198
population size 19, 26
porous asphalt 189
positive crankcase ventilation 176
positive cross-price elasticity 38
positive impacts, transport policy 13
POSSUM 319
potential accessibility measures 212–13, 218–19
potential action spaces 42–3
potential policies 320
pre-queue maximum flow 139
"predict-and-provide" approach 356
predictability, road safety 277
predictions, traffic conditions 153
predictive scenarios 312
preferences
  impact on transport 7
  modal choice 63
  residential choice 93
  in SCBA 330
  and travel behaviour 24, 35, 41, 84, 87–8
price elasticity 38, 112–14, 119, 242–3
price increases, consumer surplus 332
pricing policies 196–8, 242–3, 289, 292
prism concept 42–3, 44
private companies 300
private transport costs, and GTCs 8
pro memory (PM) indicator 335–6
probability density function 135–6
probit model 365
product characteristics, mode choice 63
product sensitivity, economic cycles 58
production/systems, freight transport demand 60–61, 70, 71
products, mass individualization 71–3
prohibition of access, to traffic 276
proton exchange membrane (PEM) fuel cell 185
proximity to facilities, and travel behaviour 86
psycho-active substances, crash risk 257
psycho-spacing models 151
psychological explanations, infrastructure project problems 297
psychological perspective, travel behaviour 28–35
public authorities, system innovations 172–3
public choice theory 296–7, 300
public health indicators 264, 265
public transport
  attitudes towards 30
  costs 8
  country differences, and differences in study results 90
  distance to connections and travel behaviour 82, 86–7
  effort resistance factors 114, 115, 116–17
  elasticities 38, 113, 360
  improvement, responsiveness to 110
  income and use of 40
  information provision 244
  investments 19
  money cost components 111
  policies 244
  quality, and travel behaviour 33, 244
  subsidies 242, 294
  travel time components 103
public-private partnerships (PPAs) 300

QALY (quality adjusted life years) 265
quality, of transport/service 33, 61–2, 118, 120–21, 244, 269
quantitative approaches, policy analysis 310
queue discharge flow 139
queues 125, 156–7
queuing theory 151–3
"quotient of opportunities", accessibility measures 219

rabbits 147
radio frequency identification (RFID) 198
rail transport
  environmental impacts 228, 237, 238, 239, 240
  estimates of external costs 286
  freight transport demand 53, 54
  infrastructure 295
  passenger estimation 310, 311
  projects, problems 297
  see also train use
ramp management 195
RAND Europe 313
raw materials, substitution 249
rebound effects, emission standards 247
recreational travel 11
redesign 199
reference alternative, in SCBA 333
reference scenarios 217, 315, 316–17, 319, 321
regulations 8, 241–2
relevance, scenario assessment 313, 314
reliability 9, 115–16, 343
renewables 233
representation, traffic flow models 148

Research in Transportation Economics 300
residential choice preferences 93
residential self-selection 84, 87
resilience, futures research 322
resistance, futures research 322
resistance factors 12, 101–21
  effort 114–18
  impact on transport 8–10
  role of travel monetary costs 111–14, 242
  role of travel time in passenger transport 102–11
response-frequency measure, habit strength 34
responsiveness
  logistics structures 72, 73
  to monetary changes
    freight transport 119–20
    passenger transport 112–14
  to public transport improvement 110
return on investment, in SCBA 334
revealed preference approach 341, 357
revenue generation, transport policy 284, 302
risk factors
  road safety policies 272–3
  road traffic 254, 255–8
  and transport resistance 117
risk-seeking behaviour 258
Ritsema van Eck, J.R. 217
road crashes 254
  costs 166, 265
  defined 263
  differences in tracing causes 270–71
  outcomes 264
  registration 264, 270
  risk factors 255, 256, 257, 258, 259, 263
  statistics 264
road design, and traffic safety 263
road fatalities 9, 168, 254, 260, 261, 262, 263, 264, 266, 267, 269
road pricing 174, 196–8, 242, 243, 298
  see also congestion charging
road safety paradigms, shifts in 269–73
road transport
  environmental impacts 168, 187, 228, 229, 231, 234, 236, 237, 239, 240
  estimates of external costs 286
  freight transport demand 53, 54, 55, 57, 59, 71
  infrastructure 82, 105
  networks 11, 12, 244, 269
  resistance factors 118, 119–20
  socio-technical system for 171
  spatial and temporal distribution 5, 12–13
  technologies to reduce negative impacts 186, 187–9
  travel time 9
road type 189, 262–3
robustness
  land use and transport 94
  of transport policies 299, 322
Rogers, E.M. 200
Rölle, D. 31
rubber road surfaces 189
rule of half 331
Rumar, K. 271

SACTRA 109
Safe System Approach 271–2
safe-distance models 149–50
safety 168, 254–78
  developments in 266–9
  and driving behaviour 12
  feelings of, as a resistance factor 116–17
  impact of transport on 14
  infrastructure and vehicle characteristics 8
  land use and transport 92
  legislation 191, 268
  measuring 263–6
  mode, age and road type 260–63
  risk factors 255–8
  shifts in road safety paradigms 269–73
  Sustainable Safety 273–7
  and technology 11–12, 191–5
  transport policies 300, 301
  and travel resistance 9
  “unintentional errors” or “intentional violations” 258–60
  see also active safety; passive safety
safety culture 271, 276
safety management decisions 272
Safety Science 267–8
safety stocks 65, 66
Salomon, I. 173, 174, 185, 198, 200
Santos, G. 299
scale of research, differences in study results 89
SCBA (social cost benefit analysis) 330–37, 344–8
scenarios 312–17
Schafer, A. 105, 111
scheduling costs 343
Schippl, J. 321
Schwanen, T. 44, 87
Schwarz, P. 313

scientific quality, accessibility measures 216
sea transport
  biofuel substitutes 247
  environmental impacts 240
  freight transport demand 53, 57
  future technologies 186
  resistance factors 120, 121
search theory 116
seatbelts 191, 193, 256, 269
second-generation biofuels 183, 185, 246
second-generation turbofans 190
selection 167–70
selective attention 34
selective catalytic reduction (SCR) catalysts 178, 179
self-enhancement values 30
self-interest 284, 296–7
self-organization 144–7
self-selection 84, 87, 338
self-transcendence values 30
Sen, A. 296
sensitivity analysis 310, 333, 337
serendipity 167
serial and combined hybrid vehicles 183–4
settlement size, travel behaviour 86
shadow price method 340
shipment size, modal choice 63, 64
shockwave theory 153–4
"short term pain" approach 323
short-term choices, travel behaviour 27
Simmons, M. 233
simulation models 44–5, 149
single constraint models 362
situational conditions, and VOT 37
sleep disorders 187, 258
slugs 147
Small, K. 116, 294
smart technology 178, 193
smartphones 224
Smeed Report (1964) 196
Smith, K.R. 235
social acceptance, equity notion and 335
social cost benefit analysis (SCBA) 330–37, 344–8
social factors, travel behaviour 25
social function, transport policy 294, 296
social impact assessment 296
social optimum 291
social scientists 10
societal costs, road crashes 266
societal feasibility, road charging 198
socio-demographics, travel behaviour 20–24, 26, 44
socio-economic variables, and travel behaviour 83–4, 86, 87
socio-technical systems 170–73
Sovacool, B.K. 167, 169
space
  fixations in 43–4
  gap acceptance process 138
  impact of 7
space-mean speeds 130, 131, 132, 133
spatial configurations, activity programmes 45, 46
spatial distribution, of traffic 5, 12–13
spatial models 355–6
spatial policies, activity programmes 45
spatial scale, densities 81
spatio-temporal qualities, of queues 153–4
specific constants 117–18
specific problems, road safety analysis 272
speed
  emission levels 188, 235–6
  intellgent adaptation 12, 192–5
  modal choice 64
  road safety risk 255–6, 263
  traffic flow *see* traffic flow theory
  travel behaviour 37, 39
  travel times and transport resistance 12
speed cameras 194
speed limiters 194
speed-flow curve 362
Spence, N.A. 207
Sperling, D. 166, 185
spill-back 156–7
spill-overs, transport investment 336–7
SPITS MODEL 59–66
spontaneous phase transitions 145–6
Starnes, M. 256
state awareness, road safety 277
stated preference approach 340, 341, 343, 357
static models 356–7
statistics 310, 354, 355, 356
Statline 20
status, vehicle choice 7
Stead, D. 86
steam (H_20) 176, 178
Steg, L. 32
Stewart, J. 180
stimulus response models 150–51
stochastic allocation 362
stochastic component, disaggregated models 363, 365

stochastic fluctuations 143
stochastic queuing theory 151–3
Strategic Petroleum Reserve 287
strategic planning 271
stress 117, 187
structural equation models (SEM) 84, 85, 87
subjective norms 29–30
subjective valuations 339
subsidies, public transport 242, 294
suburban density, activity programmes 45, 46
sulphur content, petrol and diesel 176
sulphur dioxide (SO_2) 175, 230, 234, 337
The summit conquered 269
SUNflower studies 267
supply chains 61–2, 63, 71–3
sustainable development 19, 86, 319
Sustainable Safety 273–7
sustainably produced energy 94
Sweden 196, 260, 261, 267, 268, 271, 293
sweet spot 156
SWOV 269
Sydney Harbour Tunnel 300
symbolic motives 32
synchronized flow 145, 146
synthetic data 223–4
system approach, road crashes 270
system diagrams, policy analysis 313
system innovations 170–73
Szalai, A. 105

tailpipe standards 176
Taleb, N.N. 322
targets, in backcasting 319
Tavasszy, L.A. 66
technical explanations, infrastructure project problems 297
technical factors, mode environmental performance 239
technological niches 172
technology *see* information technology; transport technology
temporal component, of accessibility 210, 214, 223
temporal distribution, of traffic 5, 13
test cycles 178, 179, 183, 184
theory, transportation models 354
theory of planned behaviour (TPB) 28–30, 31, 32, 33
third parties, external costs 285
third-generation biofuels 183, 185, 246
Thissen, W.A.H. 313
Thomson, J.M. 155
three-phase traffic flows 144–5
three-way catalytic converters 176
Tight, M. 115
time, fixations in 43–4
time factor
  dynamic models 356–7
  travel behaviour 7, 37
time geography 40–45
time headways 128–9
time horizon 89
time intervals 42, 43, 44, 45
time policies 45
Time to Decide 110
time windows 42, 45
time-mean speeds 130, 131, 132
time-of-day selection models 365
time-related costs 38
Timmermans, H.J.P. 221
ton kilometres, freight demand 55, 56, 57, 58, 59
tons lifted, freight demand 55, 56, 58, 59
top-down (TDM) methods 239
total social costs 294
Towards Zero: Ambitious Road Safety Targets and the Safe System Approach 274
"track and trace" systems 70
trade liberalization 70
trade-offs, inventory and transport costs 66
trading networks, complex 70
traffic flow theory 125–57
  flow dynamics and self-organization 142–7
  macroscopic flow variables 129–34
  microscopic and macroscopic flow characteristics 134–41
  models 147–54
  multi-lane traffic flow facilities 147
  network dynamics 155–7
  vehicle trajectories and microscopic flow variables 127–9
traffic hysteresis 143
traffic jams 105, 116, 125, 145, 146–7, 168, 195, 291, 359
traffic management 195–6, 244
traffic performance 58
traffic stability analysis 150
tragic choices, in CBA 296
train use 40, 86–7, 343–4
TRANS-TOOLS 375–6
transient states 143

transit time 118
transitions
  system innovations 170–73
  traffic phases 145–6
transport
  factors impacting on 5, 6, 7–10
  interventions 19
  models *see* models
  performance, and freight demand 55–7
  problems 19, 41, 306–7
  relevance of developments in 4–5
  and service quality 33, 61–2, 118, 120–21, 244, 269
transport component, of accessibility 210, 214
transport policies 24, 283–302
  activity programmes 45
  current 299–302
  environmental 240–44
  evaluation of options 13, 247–8
    see also appraisal
  general criteria for healthy 297–9
  reasons for government intervention 283–4
    equity 284, 293–6
    market failure 284, 285–8
    maximizing welfare 288–93
    public choice theory 296–7
    revenue generation 284
  travel time elasticities 109
  see also futures research; models
transport research
  method and differences in results 88–90
  resistance factors 9–10
  traditional 24
  see also futures research
transport system
  deregulation and cost decreases 70–71
  impact of changes in 4
  policy analysis model 307–8
  role in fulfilling people's needs 7
transport technology 163–202
  air pollution 175–9
  climate change and oil dependency 179–86
  congestion 195–9
  evolution of 163, 164–6
  impacts of 5, 11–12
  nature and dynamics of 199–201
  need for rare materials 249
  noise 187–91
  policy perspective 166–74
  potential to reduce emissions 246
  safety 191–5
  to reduce oil dependency and CO_2 emissions 245–7
travel behaviour 19–47
  changes in 111
  conceptual model 24–8, 79
  economic perspective 35–40
  geographical perspective 40–45
  and land use *see* land use
  models 357–8
  need and desires 7–8
  psychological perspective 28–35
  socio-demographics 20–24
  see also passenger transport
travel cost method 338–9, 340
travel time(s)
  factors impacting on 5, 9, 12
  role in passenger transport 102–11
  valuation of 340–44
travel time elasticity 39, 108–11
travel time ratio 44
Treiterer, J. 143
triggers, DAP process 325
trip purposes 37, 39, 91, 139, 341
trip-attraction models 361–3
trip-based approach, transport problems 41
trips per day 22, 23
two pipe regime 147
two-wheelers (motorized) 256, 260, 262, 269
tyre-road surface contact noise 188, 189, 237

uncertainty
  emission reduction potential 245, 246
  in policy analysis 309–12, 329
    see also Level 1 uncertainties; Level 2 uncertainties; Level 3 uncertainties; Level 4 uncertainties
unintentional error 258–60, 275
unit transport costs 71
United Kingdom
  accessibility 207
  car fuel economy 181
  CBA, transport policy 293
  congestion costs 105
  modal split 242
  passenger estimation, rail transport 310, 311
  road fatalities 261, 267, 268
  travel behaviour 20–24
  travel time elasticities 109, 110
  walk and wait times 103

United States
car use 24
CBA, transport policy 293
congestion costs 105
development in passenger travel 164, 165
energy use 229
expenditure on noise barriers 238
external costs 287
freight transport mode 53
fuel efficiency 180–81
fuel standards 183
induced demand effect 109
road fatalities 260, 261, 267
transport-related emissions 14, 229–30
see also California
universal mobile telecommunications system (UMTS) 198
unpredictability 116, 117
urban density, travel behaviour 85, 86
urban roads, road safety risk 263
urbanization, travel behaviour 21–2
urea function, SCR catalysts 178
USDOT 362
useful field of view (UFOV) 257
utilitarian travel demand 80
utility functions 35, 36, 37, 364
utility theory(ies) 40, 101
utility-based accessibility measures 211, 213, 214, 220–22, 223

valuation
by people, of land-use scenarios 93
of environmental effects 337–40
of travel time 340–44
valuation of outcomes, policy analysis 308
value density, modal choice 64–5
value of reliability (VOR) 116
value of statistical life (VOSL) 265–6
value of time VOT/value of time travel savings (VTTS) 37–8, 104–5, 118
value-added logistics 71
values, travel behaviour 26, 30–31, 33
Van Acker, V. 87
Van der Heijden, K. 313
Van der Pas, J.W.G.M. 318
van Pelt, M.J.F. 347
Van Schagen, I. 256
Van Wee, B. 39, 95, 216, 239, 296, 297, 298, 299, 322
variable costs 9
variables, transportation models 354
vehicle kilometres, freight demand 58, 59
vehicle trajectory, traffic flow theory 127–9
Verhoef, E.T. 285, 287
vertebrate mortality 287
VIBAT 319
Vickerman, R.W. 218
Victor, D.G. 105
Vidakovic, V. 44
video cameras 198
volatile organic compounds (VOCs) 14, 175, 176, 230, 241
volume of traffic 8, 11, 12, 19, 102, 125, 187, 235, 243
voluntary (ISA) systems 192
vulnerabilities
identifying, DAP process 324
road safety risk 256, 262, 263

waiting time, public transport 103
walkability 86
Walker, W.E. 307, 309
walking 39, 103, 104, 111, 114, 116, 117, 247
wants, economic models 36
Wardman, M.R. 105
Wardrop, J.G. 155
Wardrop principle 362
Wasielewski, P. 136
water pollution 286
waterborne transport
environmental impacts 228, 231, 240
estimates of external costs 286
freight transport demand 54
Wegman, F.C.M. 274
weight lifted, freight transport demand 52–5
weighting factors, measurement of effort resistance 117–18
Weilbull distribution 365
Weiss, M. 164
welfare theory 288–93
Wener, R.E. 117
Whelan, G. 357
wide moving jams 145, 146–7
wildlife habitat 288
Williams, J.R. 164
willingness to accept (WTA) 285, 339, 342
willingness to pay (WTP) 105, 116, 121, 285, 330, 332, 339, 342
Willumsen, L. 353, 365
Wilson, A.G. 219

World Bank 66–8, 274
World Commission on Environment and Development 228
World Health Organization (WHO) 191, 234, 236, 274
Wyman, O. 164
young people 7, 22, 32, 260, 262, 272

Zahavi, Y. 111, 155
zero emission vehicles 184, 185–6
zones, spatial models 355