ELSEVIER
JAI

RESEARCH IN TRANSPORTATION ECONOMICS
VOLUME 19

INVESTMENT AND THE USE OF TAX AND TOLL REVENUES IN THE TRANSPORT SECTOR

ANDRÉ DE PALMA
ROBIN LINDSEY
STEF PROOST
Editors

INVESTMENT AND THE USE OF TAX AND TOLL REVENUES IN THE TRANSPORT SECTOR

RESEARCH IN TRANSPORTATION ECONOMICS

Series Editor: Martin Dresner

RESEARCH IN TRANSPORTATION ECONOMICS
VOLUME 19

INVESTMENT AND THE USE OF TAX AND TOLL REVENUES IN THE TRANSPORT SECTOR

EDITED BY

ANDRÉ DE PALMA

THEMA, Université de Cergy-Pontoise and Ecole Nationale des Ponts et Chaussées, France

ROBIN LINDSEY

Department of Economics, University of Alberta, Edmonton AB, Canada

STEF PROOST

Center for Economic Studies, Faculty of Economics and Applied Economics, Katholieke Universiteit Leuven, Leuven, Belgium

ELSEVIER
JAI

Amsterdam – Boston – Heidelberg – London – New York – Oxford
Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo

JAI Press is an imprint of Elsevier

JAI Press is an imprint of Elsevier
Linacre House, Jordan Hill, Oxford OX2 8DP, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

First edition 2007

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

ISBN: 978-0-7623-1246-7
ISSN: 0739-8859 (Series)

Printed and bound in the United Kingdom

07 08 09 10 11 10 9 8 7 6 5 4 3 2 1

Dedicated To

Céline and Matthieu
Laura and Amanda
Marie-Anne, Bregtje, Antje, Thomas

This page intentionally left blank

CONTENTS

PART II: REAL WORLD EXPERIENCES

LIST OF CONTRIBUTORS

Jon-Terje Bekken	Molde University College, Molde, Norway
Laurent Cretegny	Ecoplan, Economic Research and Policy Consultancy, Berne, Switzerland
André de Palma	THEMA, Université de Cergy-Pontoise and Ecole Nationale des Ponts et Chaussées, France
Bruno De Borger	Department of Economics, University of Antwerp, Belgium
Claus Doll	Fraunhofer-Institute for Systems and Innovation Research, Karlsruhe, Germany
Pia Koskenoja	EVTEK, University of Applied Sciences, School of Business Administration, Vantaa, Finland
James Laird	Institute for Transport Studies, University of Leeds, Leeds, UK
Robin Lindsey	Department of Economics, University of Alberta, Edmonton, Alberta, Canada
Heike Link	German Institute for Economic Research (DIW Berlin), Department of Energy, Transport, Environment, Berlin, Germany
Aurélie Mercier	Laboratoire d'Economie des Transports, (CNRS, Université Lyon 2, ENTPE), Lyon, France
David Meunier	Ecole Nationale des Ponts et Chaussées, Paris, France
Chris Nash	Institute for Transport Studies, University of Leeds, Leeds, UK

Bård Norheim	Urbanet Research, Oslo, Norway
Stef Proost	Centre For Economic Studies, Faculty of Economics and Applied Economics, Katholieke Universiteit Leuven, Leuven, Belgium
Emile Quinet	Ecole Nationale des Ponts et Chaussées, Paris, France
Charles Raux	Laboratoire d'Economie des Transports (CNRS, Université Lyon 2, ENTPE), Lyon, France
Simon Shepherd	Institute for Transport Studies, University of Leeds, Leeds, UK
Kenneth Small	University of California at Irvine, CA, USA
Stéphanie Souche	Laboratoire d'Economie des Transports (CNRS, Université Lyon 2, ENTPE), Lyon, France
Urs Springer	Ecoplan, Economic Research and Policy Consultancy, Berne, Switzerland
Stefan Suter	Ecoplan, Economic Research and Policy Consultancy, Berne, Switzerland
Saskia Van der Loo	Centre For Economic Studies, Katholieke Universiteit Leuven, Leuven, Belgium

FOREWORD

There is an irony in modern understandings of transport pricing and finance. Transport is characteristically starved for funds to keep up with ever-growing infrastructure needs. Conventional funding sources are limited by the efficiency losses of raising taxes, especially labor taxes where they are already high. Thus it would seem that a source of funds that enhances rather than diminishes efficiency, and that can be raised in abundance, would be extremely desirable.

Pricing reform is precisely such a source. Many studies have shown that charging for road use at times of peak congestion would greatly enhance efficiency by reducing travel times and making them more predictable. Consumers, workers, and businesses would all benefit. Furthermore, with road use charged for in this way, public transport prices could also be raised, even while expanding service to handle diversion from users of private modes. Revenue estimates from such pricing reforms are astoundingly large.

The irony is that the very size of these revenues is a political liability. Their magnitude is almost an embarrassment, making it harder to win acceptance of pricing reform. The reason is not hard to grasp. While pricing reform creates large net benefits to the economy, they arise from two even larger but mostly offsetting quantities: revenues to the government on the one hand, and losses to travelers on the other – the losses caused by the fact that the payments these travelers must make are greater than the value to them of the travel-time savings. If travelers do not trust that the revenues will be spent in a way that fully benefits them, they may well conclude that the net benefits to themselves are negative.

Indeed, in some analyses of political feasibility of pricing reform, the ratio of revenues raised to net benefits created has been used as a rough indicator of expected political resistance to the package. As a result, some effort has been expended in finding compromise policies that are almost as efficient as the "best" policy, but that raise *less* revenue. This response turns the initial dilemma of inadequate infrastructure finance on its head!

The program undertaken by this book promises to expose these factors to fuller understanding, which may in turn lead to solutions to the dilemma. The authors develop a rigorous, analytical approach: one that explicitly accounts for the economic inefficiencies involved in raising funds for government projects, the effects of constraints such as "earmarking" of revenues, the institutional arrangements for public–private sector cooperation, and the political economy of pricing reform. Through a number of earlier research programs, European researchers have already greatly enhanced our knowledge about the effects of transport pricing and investment and how to optimize them. The researchers represented here push the analysis further by explicitly incorporating all-important institutional factors. These are the factors that underlie the irony just discussed and that determine the feasibility of proposed measures.

The analytical ideas developed here are illustrated by a series of case studies that carry out many of the suggested calculations for actual or proposed implementations of transport pricing reforms in Europe. To those of us who have been studying and advocating pricing reform of various kinds, such case studies are of great interest in themselves by providing an empirical background for such concepts as optimal prices, welfare effects, tax distortions, and distributional impacts. Within the context of this book, they are even more valuable because they begin to show how to bridge the gap between the idealized models and actual decisions being made regarding implementation (or non-implementation) of pricing reform. The process is not always tidy – factors governing actual policy development are of course more complex than those included in analytical models. Undoubtedly, various mismatches between theory and case study will provide fuel for further research to better understand the underlying processes. Meanwhile, both researchers and policy makers will find much to enlighten them as their perspectives move a little closer thanks to the efforts of these scholars.

Kenneth A. Small
University of California at Irvine

CHAPTER 1

INVESTMENT AND THE USE OF TAX AND TOLL REVENUES IN THE TRANSPORT SECTOR: THE RESEARCH AGENDA

André de Palma, Robin Lindsey and Stef Proost

ABSTRACT

This chapter introduces the research agenda. The problems related to the use of revenues from tolling and charging in the transport sector are organised into nine research questions. These range from the optimal level of user charges to the optimal allocation of the revenues and the appropriate choice of institutions to accomplish this. The theoretical contributions and the case studies of the book are briefly outlined.

1. PROBLEM, OBJECTIVES AND METHOD

This book analyses the use of revenues derived from Transport pricing.[1] Transport policy faces a classical contradiction: to serve a global policy determined by the will of voters, while being open to the legitimate interests of stakeholders including users, transport service operators, transport infrastructure owners and the infrastructure construction industry. We are all

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 1–26

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19001-6

aware of the huge problems of investment financing in transport. Notable European examples are the Channel Tunnel under the Strait of Dover, and the 30 priority projects of the Trans European Networks, which will require hundreds of billions of Euros of investment subsidies.

Pricing and financing transport infrastructure is a complex problem since there is a potential tension between marginal social cost pricing (MSCP) to support efficient use of the infrastructure, and pricing to assure adequate funds for transport investment and maintenance as well as other public revenue requirements. The field requires a large scope as it needs to embrace political decision-making processes, fiscal policy in relation to taxation and redistribution effects, Public–Private Partnership in financing and managing large infrastructures, acceptability criteria such as efficiency, equity and so on.

The goal of the research programme[2] was to tackle this ambitious research agenda using a combination of theoretical contributions and case studies whose function is presented in Section 4 of this chapter. We begin with a short presentation of the so-called earmarking question.

1.1. The Question of Revenue Earmarking in Transport Policy

1.1.1. Formulation of the Problem

Earmarking or hypothecation entails dedication of money raised by user charges, taxes or other state revenues to a specific use (be it the health system, religious schools, pensions or roads), rather than adding this money to the national Treasury. Earmarking is resisted by political scientists attached to the transcendence of the State,[3] not to mention numerous civil servants of the Ministries of Finance – and many Ministers of Finance. Indeed, as a traditional rule, in modern states, public revenues are not to be hypothecated to any particular purposes;[4] they remain open for the public decision-making process, and they are to forget their respective origins: *pecunia non olet*.

Notwithstanding, there are some good reasons to consider allocation of transport revenues to specific transportation expenditures. What is more, there is little doubt that in transportation policy, there are many deviations from the general fiscal principle of non-allocation of revenue.

1.1.2. An Example: The British Road Fund of 1909

A standard case illustrating how these deviations were possible is given by the history of the British Development and Road Improvement Fund, established in 1909. As a rule, the Treasury would not have access to the revenues produced by a new tax on motor vehicles: all would be spent on roads. The Road Board did not actually undertake itself the construction or the maintenance of roads; it only allowed grants to local authorities. It was not long before events intervened to upset plans for the Fund. The Finance Act of 1915 allowed the Chancellor of the Exchequer to retain the whole tax revenue of the Road Fund for the war budget. Later, Winston Churchill was to raid the Fund twice, in 1926 and 1927. The Finance Act of 1936 wound up the Fund. From then onwards, the Ministry of Transport was to bid for money from the Treasury, as any other government department. During the heated debate of 1936, Winston Churchill pleaded for the abolition of the Fund, arguing that its continuation would lead car users to think they had some moral ownership of the roads on the grounds that they had paid for them.[5]

1.1.3. Recent EU Policy

The debate over hypothecation of transport revenue to specific expenditures is thus a classical one, and it has a history in the field of transport. But curiously, it has not, until very recently, received explicit treatment in EU-funded studies on transportation policy. Current EU policy tends towards rather loose earmarking of revenue from infrastructure charging for the transport budget. It allows revenue raised from one mode to be spent on infrastructure associated with another mode, and also for revenue collected in one region to be spent in another region. However, this hypothecation policy is only one option regarding the use of revenue; other, more inclusive options are, for example, using this revenue for general taxation purposes such as reducing labour taxes. One could also consider more restrictive options such as directing revenue towards transport projects within the specific region or the specific mode from which it was raised.

As a matter of fact, most recent EU projects on transport pricing have focused on the optimal design of charging schemes.[6] More generally, existing cost benefit policy guidelines for transport charging emphasise the importance of deriving accurate estimates of costs and benefits of transport investments. In contrast, our research project focuses on the role of interactions between the transport sector and the rest of the economy, and the equity and acceptability aspects of user charges and revenue use.

The research project originated from a growing realisation that the long-term effects of any pricing reform would depend not only on price levels, but also on how the revenues are used:

> … it has clearly emerged that the impact of pricing policies will heavily depend (in terms of effectiveness, efficiency, equity, acceptability) on the use that will be made of the revenues generated by transport pricing schemes. The REVENUE project has been

designed to address this specific issue, thus providing further input to the formulation and development of EU policies in the area of infrastructure charging. (Laird et al., 2004, p. 2)

We see how, of necessity, any discussion of transport revenue earmarking leads to considerations of general policy questions. Some of these questions or problems are now identified.

1.1.4. Specific Problems Addressed

The problems presented in this section are discussed from a theoretical perspective in Part I, and examined in the case studies of Part II. The main problems from a decision-making perspective go beyond consideration of efficiency since their solutions often call for some form of institutional creativity. They are identified on the left-hand side of Fig. 1.1.

1. What types of transport charges should be levied on users? What are the appropriate charge levels?

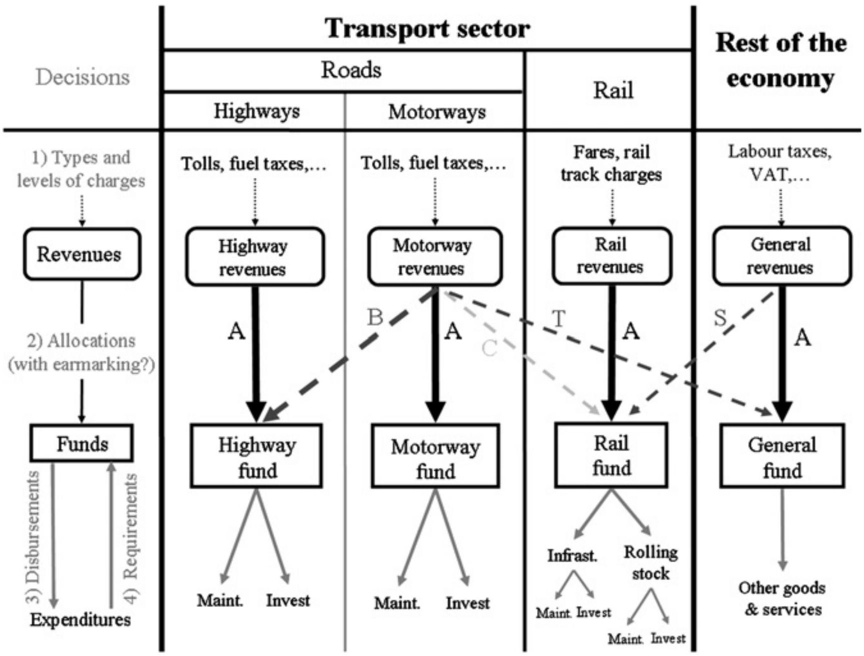

Fig. 1.1. Scheme for Transport user Charge and Revenue use Decisions.

2. How should the revenues from user charges be allocated between economic sectors and/or transport modes? Should some, or all, of the revenues be earmarked to transport funds for particular modes?
3. How should the allocated funds be distributed or disbursed among investments, maintenance and other expenditures?
4. How should any predetermined investment or other expenditures be paid for?

Fig. 1.1 depicts these four problems for a simplified interurban transport setting.

The transport sector consists of roads and rail transport, and roads are divided into highways and motorways.[7] All other sectors are lumped together on the right-hand column of Fig. 1.1 as "Rest of the economy". User charges from each mode and sector generate revenues as represented by the rounded boxes. Revenues in turn are allocated to revenue pools or funds dedicated to particular sectors or modes.[8] Finally, the funds are disbursed to finance expenditures of various sorts.[9] The arrows with letter labels (A, B, C, S, T) in Fig. 1.1 will be explained in connection with Problem 2 of the list above.

1.1.5. Advantages and Drawbacks of Earmarking

Various rules can be used to allocate transport revenues to funds. One is to earmark revenues for a particular purpose. Despite the fact that parliamentary decisions on taxes have traditionally been kept independent of decisions on expenditures, a substantial portion of transport tax revenues in Europe is earmarked to infrastructure investments and maintenance for particular transport modes. Earmarking revenues can have significant efficiency and welfare-distributional effects. Indeed, since the revenues from user charges may exceed the (monetised Pigouvian) efficiency gains by a large multiple, allocation decisions can have larger effects on the economy than pricing decisions (Parry & Bento, 2001; Mayeres & Proost, 1997, 2001).

Economists vary in their stance towards earmarking generally, and earmarking of road-usage charge revenues specifically.[10] According to normative public finance theory, tax revenues should not be locked into any particular expenditure pattern because spending priorities change over time and in unforeseen ways. Moreover, most EU research projects on transport pricing have concluded that governments should be free to use transport revenues in whatever way provides the greatest benefit (Laird et al., 2004, p. 1). Earmarking creates inflexibility in the allocation of funds, hampers effective budget control and can result in shortages of revenues for some modes and excesses for others.

In opposition to this view, a number of arguments have been advanced in favour of earmarking. One is that those who incur charges should get for what they pay. A second is that earmarking facilitates long-term planning and can reduce project costs by lowering interest rates. A third is that earmarking helps to prevent political abuse of funds (see Chapter 3).

Yet another argument that has gained widespread currency is that earmarking revenues for specific uses could make policy reform more acceptable to voters, and consequently improves the chances that reforms will successfully be implemented. This perception underlies the decision in the UK to require hypothecation of local road user charge revenues for at least 10 years following inception of a scheme. Ison refers to this hypothecation scheme as "all-important" and draws the conclusion that "… the largest proportion of the revenue generated from road user charging should be utilised to improve public transport, particularly in the area in which the charge is introduced" (Ison, 2004, pp. 174–175). This book sets out to define general guidelines for best use of the revenues, and to examine via case studies selected instances of current practice. Naturally, no attempt is made here to treat exhaustively such a complex subject – encompassing as it does not only various aspects of transport economics, but also questions of public finance and public choice economics in multilevel political systems.

2. CONCEPTS AND DEFINITIONS

We briefly introduce here the general terminology used in the book. Other terms will be introduced in later chapters. Some of these concepts have evolved over time with developments in the literature on transport pricing. Their precise meaning will be elucidated in later chapters.[11]

General terms of public finance, such as *earmarking* or *hypothecation*, were already presented above; let us consider first the terminology related to pricing.

Transport pricing refers not only to access charges, tolls and fares levied by the transportation service provider, but also to taxes and additional levies or subsidies introduced by some public authority. Taxes may be imposed to generate revenues for maintaining and expanding the infrastructure. They may also be used to cover or "internalise" social costs that are disregarded by the user and service provider such as pollution and congestion. The sum of internal and external costs is the *Marginal Social Cost*, which is often higher than the *Private Marginal Cost* that the user is prepared to take into account when assessing the advantages and the drawbacks of a trip. This taking into account of social marginal cost introduces the Pigouvian idea of *externalities*: economic costs not taken into account in markets and in the decisions made by market players, i.e. costs borne by others than those who decide. By *internalisation*, we intend the incorporation of an externality into the market decision-making process through pricing, taxation or regulatory intervention. For example, internalisation of pollution costs is implemented by charging the polluters with the damage costs of the pollution generated by them, according to the polluter pay principle.

The 2001 EU White Paper on Transport Policy is based on the principle of MSCP:[12] each user should pay the full marginal social cost related to that use. The traditional justification of MSCP is that it is ''allocatively efficient'', i.e. it optimises the allocation of resources and thus maximises the welfare of society as a whole. The argument, founded on well-established principles of economic science, is summarised in this White Paper (CEC, 2001):

> Where charges are too low, excessive demand is likely, generating higher costs than benefits, and individual operators have less incentive to reduce the costs that they impose on society. Where charges are too high, some users who would be able to pay the costs they impose would be discouraged from using the infrastructure, thereby reducing its social benefit.

The validity of this argument depends on the so-called ''first best'' conditions: markets are competitive; there are no external effects, etc. In the real world, first-best conditions are not satisfied, even approximately, and the more complex problem of ''second-best'' pricing must therefore be tackled (see Laffont, 1984).[13] We present three reasons for these distortions from first-best conditions.

The first reason is inefficient pricing of substitute or complementary transport modes. Automobile transport is typically underpriced in urban areas; and this has been a longstanding argument for subsidising public transit. Similarly, the damages to the environment caused by freight transport are higher for truck than for rail; and accordingly the EU has been advocating pricing in order to increase the rail share. Modal shifts can be encouraged by raising taxes or charges on modes with high social costs, and by lowering taxes or subsidising investments for competing modes. To fund such investments, the 2001 White Paper (CEC, 2001) suggested that supplementary infrastructure charges could be levied on high-social-cost modes such as interurban motorways.

The second reason why first-best conditions fail is heavy reliance throughout the economy on labour taxes and Value Added Taxes (VAT). These taxes often impose an excess burden on the economy because they distort economic incentives, and they are costly to collect and administer. Because the revenues from (efficient) transport user charges can be used in lieu of revenues from more distortionary sources, a case can be made for deviating from MSCP in the transport sector to boost revenues.

The third source of distortions is bias in the ways transport policies are formulated, and transport infrastructure and services are provided. Politicians and other decision makers may pursue their own interests (see Chapter 3). And unregulated private-sector agents with market power create distortions by setting prices above competitive levels (see Chapter 4).

By *revenue use*, we evoke the whole discussion that arose out of the emergence of pricing budgets and the question of their allocation: investment devoted to the expansion of the transport capacity of the infrastructure, local or global deficit compensation, or even more general policy concerns such as regulatory policy, budget transfer, welfare policy, etc. Bös (2000) defines earmarking as dedication of the revenues from a particular tax to help provide a single public good with a total expenditure that exceeds the dedicated revenues. According to this definition, dedicated revenues provide only one part of the total revenues required to fund a public good: complementary sources must be tapped, and net funding will drop if they are withdrawn.

Any revenue has its own cost, and this is true for taxes and tolls. There exist many definitions and estimates of the *Marginal Cost of Public Funds* (MCPF). The most common definition is the efficiency cost of raising one unit of tax revenue, given that the tax revenue is spent on a public good that does not affect the consumption of taxed commodities (see Chapter 3).

A transport service or mode of transport is *self-financing* when user charges cover all the costs – including infrastructure costs and external costs – related to the activity. Self-financing is often supported on the basis of the user pay principle. Where users fail to cover their full costs, the deficit must be paid for by taxpayers or victims of external costs.

A *tax* is a levy that must be paid with either no discernible service required from the government or a service that is not in proportion to the payments. Taxes include income taxes, property taxes, corporate taxes, etc. A *toll* is a special charge levied at a particular point where vehicles pass (e.g. tunnel, motorway, etc.).

Regulation is often not well understood by newcomers in the field of transportation pricing. Generally, governments are responsible for setting the policy and regulatory framework within which infrastructure managers set prices, while service providers and ultimately users are affected by pricing decisions. However, there are interactions and conflicts between each set of actors as each may pursue a different set of objectives. In this sense, regulation was an important concern for us in this research endeavour: the theoretical guidelines developed were translated into specific regulation schemes to be tested using the MOLINO model.

In a more general sense, regulation relates to all means available in order to ensure that some policy goal may be actually attained. Let us mention here a remarkable reasoning formulated in the middle of the 19th century by the French engineer Jules Dupuit: the mere augmentation of the traffic subsequent to the suppression of a bridge pricing provides a sufficient reason to open the discussion on reintroducing this scheme:[14] charging can be a tool for regulation purposes, not only a source of money. His reasoning was not dependent on cost accounting, focused as it is on congestion regulation (of course, other, non-pricing measures are available for this goal). We see here in a nutshell why real-world politics is complex, as various means correspond in various ways to various goals.

We also had to consider a specific scale of regulation, where some political instance tries to regulate the management of some transport service operated by a concessionaire. This introduces the idea of the quality of regulation: the degree to which the convergence of private objectives of the concessionaire towards the public objectives of the regulator can be achieved.

Because of the large costs of *transport infrastructures*, and the large risks entailed, it has been widely argued that only national governments have pockets deep enough to finance them, and this is why most infrastructure is state-owned. Nevertheless, in some countries, private entities own and manage road networks, railways and airports. In many situations, *transport services* are operated and maintained by a private contractor as part of a concession. Transport operators may also contribute to the *financing* of infrastructures, etc. These various configurations traditionally call for complex and innovative institutional arrangements typically called *Public–Private Partnerships* (PPPs or P3s). For example, in a "concession type" PPP, the private sector takes on all the investment. Public and private parties divide the risks ex ante by contractual agreement (for example, the government bears the political risks, while the private party takes on the construction and commercial risks).[15]

Public Transport (PT) refers to all services for passenger and freight transport that are supplied according to a pre-defined timetable and are open to use by any individual or organisation. Examples of PT are rail, bus, air and ferry services. As is well known, PT is a complex case for pricing theory.[16] The transport of an additional person or unit of goods does not, in the short run, causes additional vehicle kilometres, as scheduled vehicles are used, which are running anyway. In the long run, due to increased capacity use, additional or larger vehicles could be scheduled; also, PT infrastructure could call for investments.[17]

Economic efficiency corresponds to the idea that aggregate welfare is maximised while no attention is paid to welfare distribution. Economic theory distinguishes between *productive efficiency* (producing a given output at minimum cost), *allocative efficiency* (dividing up a given output amongst people so that there can be no mutual gains from exchange) and *output efficiency* (producing the right mix of outputs of different goods). Welfare maximisation requires that all three types of efficiency conditions be satisfied. The measure of efficiency to be used is thus context dependent. *Welfare* is the well-being of the population; in economics, it is usually measured in terms of their own preferences or utility.

Though the terms *equity* and *acceptability* are sometimes used interchangeably, the two concepts are quite distinct. Equity relates to moral judgements on the distribution of welfare. Acceptability concerns approval or disapproval of a particular policy by voters or decision makers.[18] A head tax on identical households may be considered inequitable but acceptable by politicians; conversely, a policy that confers small benefits on poor people while concentrating the costs on a richer majority may well be considered unacceptable by the majority, but would clearly be judged equitable.

Concerning management funds, a distinction is commonly made in recent use between *first-generation* or conventional funds, and *second-generation*, more commercial-type funds, in which spending and management decisions are made by an independent transport infrastructure agency (TIA), rather than through the political process.

The term *actors* refers to social groups that (1) create costs by using specific means of transport, and (2) suffer costs by either paying for vehicle operation, providing infrastructure or bearing the external effects caused by transport. Individuals, PT providers, interest groups and every level of government are all actors. We could also speak of players, as some of our models are inspired by game theory.

A state of *asymmetric information* exists when one player in a game has more information than the other. The management of PPP raises the question of the economics of incentives, which "can be described as the design of rules and institutions for inducing economic agents to exert high level of effort (in a broad sense), and to reveal truthfully all socially relevant information they might have" (Laffont, 1996, p. 49; see also Laffont & Tirole, 1993). Each actor has two problems to solve: face its own problems of risk management, and make information available to other stakeholders; see the discussion in Leruth (2006), Lindsey, de Palma, & Proost (2006), and Chapter 4, Section 2.1, regarding the responsibilities that could be assigned to an independent TIA. In particular, a TIA may help to reduce information asymmetries among actors by generating useful new information and by diffusing it.

The *principle of subsidiarity* requires political tasks to be allocated to the lowest level of government that can perform them efficiently. This principle is not sufficient, of course, when cross-border traffic or externalities are to be taken into account, or when more than one local government wants to control transport infrastructure and pricing.

3. PROBLEMS AND TRADE-OFFs

We will propose here some remarks on much discussed topics. As the reader will see, most often they do not accept clear-cut answers; moreover, they do call for a constructive approach. This is why we call them *problems*.

3.1. Problems Concerning the Efficiency of Charging Policy and Revenue Use

3.1.1. Problem 1: What Charges should Users Pay?

For two reasons, transport charging is of central importance in the research presented in this book. First, the types and levels of user charges affect the volumes of passenger and freight transport flows, and consequently the allocative efficiency of transportation activity. Second, the system of charges determines the amount of revenue generated that can be allocated for various uses.

If first-best conditions held throughout the economy, the answer to Problem 1 would be straightforward: charges should be set accordingly to MSCP principles. The nature of the infrastructure, the service and the externalities (e.g. congestion, road damage, pollution, etc.), and the spatial and temporal variation of the costs, would dictate what types of charges are required such as highway tolls differentiated by time of day, vehicle size, axle weight and emissions control.

In theory, all the different types of second-best distortions identified above can be treated endogenously with Computable General Equilibrium models. In practice, distortions are usually accounted for in rough-and-ready fashion by factoring up the social cost of raising revenues by some estimate of the MCPF. This point is thoroughly assessed in Chapter 3, and discussed in the case studies.[19] The studies vary widely in the base-case values they adopt for the MPCF. The disparity in these values may reflect real differences between the countries in the social costs of raising public money (see Kleven & Kreiner, 2006). However, the "true" value of the MCPF in a given jurisdiction is difficult to determine precisely, and for sensitivity analysis two studies (for Germany (Chapter 10) and France (Chapter 11)) entertained alternative values.

3.1.2. Problem 2: How should Revenues be Allocated to Funds?

The second problem identified in Fig. 1.1 is how to use revenues generated from transport charging. One possibility is to improve or expand the facilities on which the charges are imposed. This choice is depicted in Fig. 1.1 by the vertical arrows labelled "A". Another option is to spend the revenues on other facilities within the same mode of transport. This form of cross-subsidisation is illustrated for motorways and highways by the diagonal "B" arrow. A third alternative is to cross-subsidise another mode as illustrated by the diagonal "C" arrow extending from the motorway revenues box to the rail fund. Cross-subsidisation of types B and C is common for facilities with low traffic volumes that generate too little revenue from user charges to finance themselves.

Unless the overall transport sector is self-financing, revenues will flow to or from the rest of the economy. In most of the EU, rail transport is subsidised from general revenues as indicated by the "S" arrow in Fig. 1.1. And in countries such as Britain road transport more than pays its way, and the general fund receives a net transfer as is represented by the "T" arrow. Such transfers can be used in a variety of ways. One option that has received wide support in recent years is to reduce existing labour taxes (Mayeres & Proost, 2001).

The merits of earmarking were a central concern of the REVENUE research project, and all the case studies in Part II include scenarios in which revenues are earmarked. Nevertheless, the analyses of earmarking in these studies are incomplete in three respects.

First, and contrary to Bös's definition of earmarking (Bös, 2000), some of the case studies did not restrict revenue allocation to a single public good, but rather considered allocations between two or more purposes in varying proportions.

Second, and more importantly, the studies did not systematically address the fact that earmarking entails deliberate constraints on how revenues can be used. The studies largely ignore the institutional and politico-economic factors that determine whether constraints are desirable despite the inflexibility inherent in earmarking.

Third, the studies largely ignore the danger that earmarking will be undermined by a claw-back of general funding. The Edinburgh study describes how the transport investment to be funded by the proposed congestion charge was presented as a separate package from the investments to be paid from existing revenue sources. The French study discusses the case of the recent French multi-modal national interurban investment transport fund Fonds d'Investissement pour les Transports Terrestres et les Voies Navigables (FITTVN), created in 1995 and abolished in 2001. Among other shortcomings, this fund was weakened because it only replaced money from the national general budget, rather than adding a net increase. But none of the studies entertain compensating behaviour in a formal scenario.

For all these reasons, it would be more accurate to say that the case-study scenarios as a whole examine alternative revenue allocation decision processes rather than earmarking decisions in the full sense of the term.

3.1.3. Problem 3: How should Funds be Spent?

Revenues that have been allocated to specific modes must then be disbursed for particular purposes. One choice is between maintenance and investment. The two decisions are intertemporally linked because investment leads over time to additional maintenance obligations as the new infrastructure ages. In the case of rail or public transit, there are also choices between infrastructure and rolling stock (see Fig. 1.1), as well as between different transit systems (e.g. buses, trams, light rail transit, etc).[20]

3.1.4. Problem 4: How should Predetermined Expenditures be Paid for?

Some of the case studies did not seek to identify optimal investment or expenditure plans, but rather asked how predetermined outlays should be paid for at least social cost. This is the case in the Swiss study for the New Alpine Rail Tunnel investments (Chapter 9), and in the French study (Chapter 11) for some motorway projects and the Lyon–Turin rail link.[21] Problem 4 is represented in Fig. 1.1 by the arrow labelled "Requirements" pointing upwards from Expenditures to Funds.

An important consideration when formulating or analysing transport policies is that optimal pricing, revenue allocation and expenditure decisions are interdependent. This is obvious when reading Fig. 1.1 from top down: the structure and levels of user charges determine how much revenue is generated, and by altering usage charges affect the benefits from investments. In addition, allocation and disbursement decisions determine how much of the revenue is spent for given purposes. The interdependence between decisions also runs from bottom up in Fig. 1.1. Pricing and allocation decisions are affected by funding requirements. Investments in capacity expansion will typically alleviate congestion, and hence reduce optimal congestion charges. But emissions, noise and road damage may rise because of greater traffic volumes, and any corresponding user charges will increase in tandem.

Optimal pricing, allocation and disbursement decisions are sensitive to such case-study-specific considerations as the capacity and condition of transport infrastructure, the availability of funds from various sources, etc. Local circumstances will also determine which of Problems 1–4 are the most crucial to address.

3.2. Trade-offs between Efficiency, Equity and Acceptability

Political processes have their own calendar, and this fact should not be neglected when discussing the interaction of rational analysis of efficiency, moral consideration on equity and political problems of acceptability. In a decision-making process, a poorly scheduled calendar can block a solution for decades if it happens that a good solution was not presented to the voter in due time.

We have to mention here the role of the lobbies and the opportunities open to different stakeholders, as well as constraints specific to the political agenda. In this discussion, the difference between users, citizens, voters, residents, etc, is important, as these distinctions are related to potential difficulties in the implementation process of some charging scheme (for example, the planning of the referendum in September, 2006, on Stockholm's congestion charging experiment was disputable, since users residing outside the city were not invited to vote). Policy evaluations are frequently made on the basis of changes in aggregate welfare. This approach may be considered adequate in some contexts, but it is deficient for analysing transport pricing and revenue allocation policies that affect diverse groups. One reason is that groups generally differ in economic status, so that equity concerns arise. Indeed, a candidate policy may be considered inequitable even if agents are identical ex ante if their ex post realisations are different. A second and related reason is that the benefits and costs of policies tend to fall unequally in the population, and those who perceive themselves to be losers may object. It is clear that an unpopular proposal is unlikely to be implemented.

Thus, more so than just aggregate welfare changes, the distributions of perceived benefits and costs matter in determining equity and acceptability, and there are many factors to take into account in determining them. For example, the welfare impacts of tolls depend not only on changes in monetary costs and travel time, but also on changes in the geographical distribution of emissions, accidents and other external transport costs. The welfare impacts also depend on how revenue is allocated between roads, PT service improvements and expenditures on other goods. In the absence of complete accounting data, assumptions have to be made about the incidence of benefits. The Swiss study, for instance, assumes that welfare gains are divided equally per capita to low- and high-income passengers (with none of the gains accruing to freight transporters).

Acceptability can be gauged through polls and surveys. Surveys can elicit what policies stakeholders favour, as well perhaps as what set of policies they would vote for. Surveys were conducted as part of the Oslo, Edinburgh and German studies that convey much useful information of this sort.[22]

Despite the distinction between equity and acceptability noted above, it is likely (although certainly not inevitable) that the two tend to go hand-in-hand. In principle, efficiency can also be positively correlated with equity and acceptability.[23] However, a majority of road pricing and other transport pricing studies have concluded that enhancing the equity and/or acceptability of a pricing *cum* revenue use package will require some sacrifice of efficiency. If so, two further problems can be raised.

3.2.1. Problem 5: How to Set a Trade-Off between Efficiency and Equity?

The political confrontation of actors leads to attempts to find an acceptable trade-off between efficiency and equity. The importance of this task cannot be overlooked if we are to ensure acceptability. For example, how should we choose between two criteria: the increase in social surplus derived from a policy and the percentage of voters who will support it? We would thus have to examine the trade-off between efficiency and equity effects of paying investments by user charges or by general tax revenue (see Chapter 3).

All the case studies assess efficiency. Some also compute changes in consumers' surplus disaggregated by population groups, and review evidence on acceptability. None of the studies attempt to identify the trade-off in quantitative terms[24] although some perception of the trade-offs can be gleaned from the Oslo, Edinburgh and German surveys.

3.3. Designing a Policy for the Management of Transport Funds

Several aspects of transport pricing and revenue use were tackled in the REVENUE project, but played a smaller role in the case studies.

3.3.1. Problem 6: Are Costs Recovered under Marginal Social Cost Pricing? Cost recovery of infrastructure expenditures is an important question. Chapter 2 of this volume reviews the theoretical literature on cost recovery and some empirical evidence by transport mode. The key theoretical result is the Cost Recovery Theorem: revenues from efficient user charges just suffice to pay for the long-run costs of building, operating and maintaining infrastructure. If the conditions of the theorem hold, each transport mode is self-financing. Neither surpluses nor deficits arise and there is no prima facie case for cross-subsidisation within modes or between modes, or for net transfers between the transport sector and the rest of the economy.

While the Cost Recovery Theorem is a useful benchmark result, the conditions that underlie the theorem are strong.

1. First-best conditions apply and user charges are set according to MSCP principles.
2. Capacity is perfectly divisible and can be expanded at constant marginal cost.
3. User costs depend only on the ratio of usage to capacity.[25]
4. Capacity is at its long-run optimal level.

For reasons discussed in connection with Problem 1, Condition 1 is most unlikely to hold. Condition 2 is also doubtful because of lumpiness of capacity as well as space and environmental constraints on expanding capacity. Empirical evidence on Condition 3 varies, and it is usually violated for PT because of economies of traffic density. Finally, due to the rigidity of infrastructure and the long lead times often required to add capacity, Condition 4 is unlikely to be fulfilled except at infrequent points in time.

The stringency of these conditions perhaps explains why none of the case studies investigated whether they are satisfied for the settings they examined – or indeed whether cost recovery would obtain under MSCP. The French study did examine whether the revenues from user charging schemes on existing infrastructure generate sufficient surplus to pay for particular new infrastructure investments. However, this question differs from the traditional cost recovery question in two respects. First, the user charges are not restricted to MSCP, and second the infrastructure costs considered include not only existing infrastructure but also lumpy new investments.[26]

3.3.2. Problem 7: How should Transport Funds be Managed?

The representation of transport funds as boxes in Fig. 1.1 might suggest that funds are no more than temporary receptacles for money. But real-world transport funds embody numerous institutional features that must be accounted for through a careful discussion of the political decision-making process (discussed in Chapter 4). For example, the Swiss multi-modal fund FINÖV and the French funding agency AFITF resemble commercial-like, second-generation funds. Experience has shown that the operational performance of funds varies widely. Chapter 4 describes some of these design issues, but the case studies do not address them directly.

3.3.3. Problem 8: How to Assign Responsibilities between Governments?

Political responsibility for transport policy formulation and implementation is currently vested with different levels of government. Often, responsibility for transport infrastructure is also shared: in Europe, local, regional, national and supranational (i.e. the EU) governments all play a role. Furthermore, the prevalence of through traffic between regions, and trans-boundary externalities such as pollution, create overlapping interests for governments in neighbouring jurisdictions.[27] Where such vertical or horizontal relationships between governments exist, the assignment of responsibilities becomes an important policy issue. Three problems related to charging and revenue have received attention: Who should set the charges? Who should collect the revenues? Who should decide how the revenues are spent? These problems were addressed in the Edinburgh study with respect to the diverging interests of residents in the city centre and residents of the region.

3.3.4. Problem 9: How to Engage the Private Sector?

The private sector can be harnessed to provide transport infrastructure in many ways ranging from a simple maintenance contract to a comprehensive Design-Finance-Build-Own-Operate concession. Private-sector involvement has several potential advantages vis à vis wholly public schemes. Private-sector financing helps to circumvent public-sector borrowing constraints by tapping an independent source of funds. Experience suggests that private firms are often better at identifying attractive investment projects and able to build infrastructure more quickly and cheaply. And the private sector may sometimes have a greater incentive than do public agencies to achieve productive efficiency, and to seek innovative ways to cut costs and/or improve service quality.

Private involvement also has potential drawbacks. Contracts must allow private operators to earn an adequate rate of return, and risks related to demand uncertainty, cost overruns and other contingencies must be dealt with. And private operators have an incentive to exercise market power while disregarding externalities such as emissions and noise that do not adversely affect customer demand.[28]

The preferred mix between public and private organisations depends on many factors: the scope for competition, uncertainties, asymmetries of information about demand and costs, the adequacy of regulations and so on (see Chapter 4). There is no definitive or exhaustive set of rules for determining the optimal mix, and each case must therefore be examined on its own merits. The literature on PPPs and optimal contracting has made strides in recent years, and there are many case studies of both successes and failures. However, none of the case studies included here explore the role of the private sector.

4. PRESENTATION OF THE BOOK

4.1. Outline of the Chapters

We present here a brief outline of the book. Chapter 2 discusses the Cost Recovery Theorem and its implications. The Theorem is especially noteworthy for establishing that there may be no conflict between short-run MSCP and cost recovery. Chapter 3 discusses the public finance and political economy aspects of transport charging. It proposes cost benefit rules for transport investments that depend on the way the investment is financed. The source of finance comes in via the MCPF. The resulting expressions are used in the MOLINO model, which was developed in this research project and used in three of the case studies. There is no single omniscient government. Some attention is therefore devoted to decision problems related to multi-level government. The chapter also takes a political-economy perspective on pricing and investment rules, and shows how the behaviour of politicians is influenced by the economic context and the political landscape. Chapter 4 discusses the institutional forms that infrastructure management can take, describes the possible role of a TIA, and considers the pro and cons of various forms of PPP. Chapter 5 integrates the theoretical prescriptions on optimal pricing, investment and revenue use developed in the preceding chapters into an assessment model for case studies (MOLINO) in order to test alternative regulation schemes.

Chapter 6 discusses the role of the case studies and the main policy questions they address. Chapters 7–11 present the case studies that are included in this volume: the Oslo and Edinburgh urban tolling schemes, the Swiss railway investment fund, the German Heavy Goods Vehicle motorway toll and the use of revenues from existing motorways in France to cross-finance either new motorways or rail projects. Finally, Chapter 12 provides a synthesis of the conclusions and considers the prospects of implementing pricing and revenue-use packages that are not only efficient but also politically acceptable. This is of interest for the design of policy guidelines, and suggests some interesting questions for future research.

4.2. Remarks about the Status of the Case Studies

4.2.1. Case Studies are not Descriptions: The Role of Modelling

Several simulation models were used (mainly the Molino model) in the case studies, which are not mere descriptions designed for comparison's or generalisation's sake. Modelling allows us to analyse, using predictions and confirmation or refutation, various efficiency and equity aspects of alternative pricing, investment and regulatory regimes, with emphasis on the allocation of revenues.

We first present here some features of the MOLINO model. This model is more a policy assessment model than a forecasting model. It is a partial equilibrium model of the transport market: income levels of the private transport users, and production levels of the firms using freight services as input, are taken as given. Primary outputs are equilibrium prices, transport volumes, travel times, cost efficiency of operations, toll revenues and financial balances, travellers' surplus and social welfare. The model includes a local and a central government, which can pursue different objectives and control different tax and subsidy instruments including fuel taxes, public transport subsidies and profit taxes.

The time horizon, which can be chosen by the modeller, typically covers 10–50 years. The model includes separate modules for demand, supply, equilibrium and the regulatory framework. In its present form the model contains two transport alternatives (two parallel roads, road and parallel railway, railway and competing air link, etc.). Transport users pay a generalised cost composed of a resource cost (e.g. fuel), taxes levied by central and local governments (e.g. fuel taxes, car taxes), a user fee (toll or rail fare) and a time cost. For a given infrastructure, travel time is assumed to be a linear function of traffic flows.

For each transport alternative a distinction can be made between an operator (who takes care of maintenance and can set tolls or user charges), and an infrastructure supplier (who decides on capacity extensions and on infrastructure charges). The costs of the operator have a linear structure: a fixed cost, constant variable maintenance and operation costs that depend on the type of vehicle or load, and finally a payment for infrastructure use that can be specified in different ways. The infrastructure provider also has a linear cost structure where the main costs are the investment and associated financial costs for the infrastructure. Operator and infrastructure suppliers can be private or public agents, and the cost level can depend on the contractual form.

Given the demand and cost functions, and the regulatory framework (see below) that specifies the behaviour of the governments, operators and infrastructure suppliers, the equilibrium module computes a fixed-point solution in terms of prices and levels of congestion for the two transport alternatives and this for the horizon that is selected.

4.2.2. Case Studies allow us to Assess Policies

Part II of this volume presents five case studies: two concerning urban transport (Oslo and Edinburgh) and three concerning interurban surface transport (Switzerland, Germany and France). The studies assess a range of scenarios encompassing existing policies, official proposals, policies currently under discussion and policies or scenarios developed by the case-study authors that may be welfare-superior and/or more acceptable to policy makers, the public and other stakeholders than existing or proposed policies. Each scenario is defined by a pricing regime, rules for allocating revenues between sectors and/or modes, and expenditure plans that may include infrastructure investments.

The Oslo and Edinburgh studies assess urban cordon tolls and supplementary measures that have been implemented (Oslo Packages 1 and 2), proposed (Oslo Package 3) or proposed and turned down in a referendum (Edinburgh). Meanwhile, the Swiss and French studies assess particular rail investments (Alpine rail tunnels in Switzerland, the Lyon–Turin rail link in France), whereas the German study considers revenue allocations to rail and road investments in varying proportions without identifying specific projects.

As far as pricing regimes, the studies are similar in that all feature pricing in the status quo, a MSCP regime and at least one intermediate regime. The

Swiss and French studies consider supplementary motorway user charges to fund new rail infrastructure. The three interurban studies include road and/or rail funds, whereas funds are not treated explicitly in the urban studies. All studies feature alternative revenue earmarking schemes. Earmarking is either a legislative requirement in the jurisdiction under consideration, or perceived to be an acceptability constraint on revenue use.

The institutional settings are similar in that management, operations, setting of charges, revenue collection and revenue allocation are explicitly or implicitly assumed to be the responsibility of public rather than private agencies.[29] The Edinburgh study is unique in examining a setting with overlapping governments that differ in their objectives.[30]

All case studies use partial-equilibrium models to assess the impacts of alternative pricing *cum* revenue use policy packages. The German and French studies use the MOLINO model that is described in Chapter 5.[31] The Swiss study uses a variant of the MOLINO model. The Oslo and Edinburgh urban case studies used other models with more detailed representations of the respective urban road networks than the MOLINO model can accommodate.

4.3. Conclusions and Perspectives

The message of this book is simple. Over the coming years, transport infrastructure developments will depend increasingly on the level of user charges. It is expected that these user charges will progressively replace government subsidies for infrastructure expansion and maintenance. This trend has to be put in its proper context, including the fiscal, political and economical problems of labour taxes, head taxes and VAT, as well as the working of the whole of the economy and the dynamics of political decision-making processes in a real social world, including labour force and social security management. We hope the book sheds some light on a particularly complex web of cross relations.

NOTES

1. As Section 2 explains, the term "Transport pricing" encompasses not only access charges, tolls and fares levied by transport service providers, but also taxes and additional levies or subsidies introduced by some public authority.
2. The REVENUE project builds on results derived from a number of studies on transport policy that were financed over the last decade by the European

commission. The REVENUE is part of the EU Fifth Framework Research Program (DG TREN, 2003–2005).

3. Indeed, many political scientists would maintain that the State is not an agent among others, and that it embodies the substantial difference that prevails between the universe of social relations which is build on the base of private contracts and such specifically political events as war, the law or the election of a national government. In this last sense, the State has to negate particular interests, as it has to claim the monopoly on violence.

4. As Dunn aptly remarks: "Earmarking is more than just a finance mechanism. It is a political issue, and for some individuals and groups it becomes a political credo. It has its roots in a Millsian belief in individual liberty and in a Lockean sense of social contract between the citizen as taxpayer and his government" (Dunn, 1978, p. 33).

5. See Dunn (1978) and Walker (1956) for a historical and institutional analysis of political decision making with respect to earmarking in transport policy. The French case study (Chapter 11 in this volume) provides another story of a short-lived transport fund, the FITTVN (see Section 3.1.2 below). There is an unwritten rule in France that when a fund is liquidated, any extra taxes that are used to finance it are not terminated; rather, the money is directed to the general Budget of the State. This was the story of the Fonds spécial des grands travaux (FSGT), established in 1982 after the abolition of the Fonds Spécial d'Investissements Routiers (SFRI), established in 1951 and terminated in 1988. A similar fate befell the Fonds Spécial d'Investissements Routiers (FSIR) established in 1951. The so-called stable resources for the fund were gradually reduced to insignificant levels through the influence of the almighty Ministère des Finances, which was deeply hostile to anything resembling a violation of the principle of annual budgeting.

6. An exception was the TRENEN – II program (1998–2001), that pleaded explicitly to use the revenues of external cost pricing for a reduction of existing labour taxes (see De Borger & Proost, 2001).

7. Figure 1.1 can be modified to describe an urban transport setting by replacing rail with public transport, and by replacing highways and motorways either with urban expressways and city streets or with just a single category of roads. Some of the case studies consider additional modes of transport (e.g. inland waterways in the German study; see Chapter 10 of this volume). Others consider finer divisions of roads (e.g. existing motorways and new motorways in the French study; see Chapter 11, or roads in certain travel corridors vs roads elsewhere in the country as in the Swiss study; see Chapter 9) or finer divisions of PT (e.g. tram lines and buses in the Edinburgh study; see Chapter 8). Figure 1.1 can be extended to include these additional segments without affecting the substance of what follows.

8. The two urban case studies (Chapters 7 and 8) do not actually feature explicit transport funds. For the purposes of discussion, however, the allocation of revenues in these studies can be thought of as being channelled through funds as well.

9. Transfers may be made between funds; to avoid clutter, transfers are not shown in Fig. 1.1.

10. A diversity of attitudes is evident in the contributions to a recent special issue on road pricing in *Transport Policy* (Saleh, 2005); see also de Palma & Quinet (2005).

11. For a general discussion of the economic approach to transport pricing, see the textbook by Small and Verhoef (2007).

12. See the discussion by Quinet (2001).

13. The intricacy of second-best pricing has been intensively analysed in the economics of the public sector (e.g. Bös, 1985, Section 2) as well as in transportation economics (e.g. Nowlan, 1993; Verhoef 2002).

14. Cf. Dupuit (1849): "… depuis la suppression du péage, la circulation avait décuplé … cette donnée d'expérience me paraît suffire pour présenter la question du rétablissement du péage …".

15. See, inter alios, Quinet and Vickerman (2004), Leruth (2006) and Vickerman (2006). Bernstein (2005) presents the Erie Canal as a significant historical example of this kind of PPP for the management of risk.

16. The field of PT owes much to the path-breaking study on pricing New York's subways by Vickrey (1955); see the discussion paper by Arnott (1997).

17. As is well known, the distinction between short- and long-run pricing is relevant for both private and PT, but often with different cost management traditions. This remark opens the way to an interesting discussion on various management approaches for collective transport services.

18. Acceptability to individuals can be assessed by whether their utility rises or falls – although complications such as incorrect perceptions and envy make this an imperfect test (Mayeres & Proost, 2003); see also Jones (2003).

19. An exception is the French study (Chapter 11), which derives a value for MCPF using an estimate of the toll elasticity of demand for a particular highway.

20. Automobiles, lorries and other privately operated mobile plant are excluded from the schema in Fig. 1.1 because they involve neither public-sector nor infrastructure decisions.

21. The Edinburgh study (Chapter 8) also featured predetermined candidate investments (tram lines and bus services) but focused on whether they were economically justified rather than on how they should be paid for.

22. Survey results are not always reliable. Strategic response bias is one generic problem. Another is that public perceptions of road pricing may change as it progresses from an abstract concept to a fully developed proposition, and finally to an operational scheme (UK Department for Transport, 2004, Appendix D, p. 14).

23. For example, a comprehensive and time-differentiated road-pricing scheme with no discounts or exemptions for any groups may not only be efficient, but also spatially equitable and seen as fair because drivers pay in proportion to the costs they impose. Indeed, of the five schemes that were analysed for the London congestion charging research programme, the most comprehensive scheme was found to be the most equitable (Richards, 2005, pp. 54–55).

24. Some recent studies have attempted to quantify trade-offs; see for example, Mayeres and Proost (1997, 2001, 2003), Raux and Souche (2004), Safirova et al. (2004), Kalmanje and Kockelman (2004), Gulipalli and Kockelman (2005), and Armelius and Hultkrantz (2006).

25. This means that if usage and capacity are both doubled, private user costs are unchanged (costs are homogeneous of degree zero in capacity and usage).

26. According to the Cost Recovery Theorem, if capacity is below its long-run optimal level then usage revenues suffice to cover the full costs of existing infrastructure and to fund a *marginal* capacity expansion.

27. See, in particular the mention, in Chapter 3, of the "double marginalisation problem" that arises in industrial organisation theory (Tirole, 1988).

28. This is not to say that similar incentive problems do not arise in the public sector. Politicians and other officials have their own agendas that may be imperfectly aligned with social welfare. Government agencies, for example, may succumb to the temptation to boost revenues by raising tolls above optimal levels, restricting capacity or reducing service quality to cut costs.

29. Infrastructure management and service operations for both road and rail are private in the French study but this is not an explicit consideration in analysing pricing or investment decisions.

30. The City of Edinburgh Council is assumed to consider only the interests of city-centre residents, whereas the regional authority also regards the welfare of residents in the wider region.

31. In addition to MOLINO, the German study uses the system dynamics macroeconomic model ASTRA to assess the long-run macroeconomic impacts of pricing and investment schemes.

ACKNOWLEDGEMENTS

The REVENUE research project, which constitutes the basis of the results presented in this book, received financial support from the European Commission (EU Fifth Framework Research Program, DG TREN, Contract No. GMA2-2002-2011). We are grateful to Professor Émile Quinet (ENPC, Paris) and Serge Pahaut (ULB, Brussels) for helpful and in-depth comments and for suggesting useful material. We also benefited from various discussions with several scholars and practitioners.

REFERENCES

Armelius, H., & Hultkrantz, L. (2006). The politico-economic link between public transport and road pricing: An ex-ante study of the stockholm road-pricing trial. *Transport Policy*, *13*(2), 162–172.

Arnott, R. (1997). *William S. Vickrey: Contributions to public policy*, Boston College Working Paper no. 387. Newton, MA.

Bernstein, P. L. (2005). *Wedding of the waters: The Erie Canal and the making of a great nation*. New York: Norton.

Bös, D. (1985). Public sector pricing. In: A. J. Auerbach & M. Feldstein (Eds), *Handbook of public economics I* (pp. 129–211). North Holland: Elsevier Science.

Bös, D. (2000). Earmarked taxation: Welfare versus political support. *Journal of Public Economics*, *75*(3), 439–462.

CEC (2001). White Paper: European transport policy for 2010: Time to decide. Commission of the European Communities, COM (2001) 370, September 2001, Brussels.

De Borger, B., & Proost, S. (2001). *Reforming transport pricing in the European Union: A modelling approach*. Cheltenham: Edward Elgar Publishing Ltd.

de Palma, A., & Quinet, E. (Eds). (2005). *La Tarification des Transports: enjeux et défis*, Economica, Paris.

Dunn, J. A., Jr. (1978). The importance of being earmarked. *Comparative Studies in Society and History*, *20*(1), 29–53.

Dupuit, J. (1849). De l'influence des péages sur l'utilité des voies de communication. *Annales des Ponts et Chaussées*, 2 série. n 207, 170–248.

Gulipalli, P. K., & Kockelman, K. M. (2005). Credit-based congestion pricing: A Dallas-Fort Worth application. Paper presented at the 85th annual meeting of the transportation research board, Washington, DC.

Ison, S. (2004). *Road user charging: Issues and policies*. Hampshire, UK: Ashgate Publishing Ltd.

Jones, P. (2003). Acceptability of road user charging: Meeting the challenge. In: J. Schade & B. Schlag (Eds), *Acceptability of transport pricing strategies* (pp. 27–62). Amsterdam: Elsevier.

Kalmanje, S., & Kockelman, K. M. (2004). Toll roads in Texas: Traffic and welfare impacts. Paper presented at 84th annual meeting of the transportation research board, Washington, DC.

Kleven, H. J., & Kreiner, C. T. (2006). The marginal cost of public funds: Hours of work versus labor force participation. *Journal of Public Economics*, *90*, 1955–1973.

Laffont, J.-J. (1984). *Théorie micro-économique*. Paris: Economica.

Laffont, J.-J. (1996). William Vickrey: A pionneer in the economics of incentives. Nobel Lecture, December 27, 1996.

Laffont, J.-J., & Tirole, J. (1993). *A theory of incentives in regulation and procurement*. Cambridge: MIT Press.

Laird, J. J., Nash, C. A., Nellthorp, J., Macário, R., Van der Hoofd, M., Carmona, M., Proost, S., Suter, S., & Lieb, C., with contributions from Pires, E., & Ambrosio, A. (TIS). (2004). Deliverable 1: State of the Art and Conceptual Background. Project REVENUE (Revenue Use from Transport Pricing), Funded by the European Commission, Competitive and Sustainable Growth Programme, Contract: GMA2-2002-52011.

Leruth, L. (2006). Public/private cooperation in infrastructure development: A story of contingent liabilities, fiscal risks, and other (un)pleasant surprises, Keynote address to the First International Conference on Funding Transportation Infrastructure, Banff, Alberta, Canada, August 2 (Lindsey, R., de Palma, A., & Proost, S., 2006).

Lindsey, R., de Palma, A., & Proost, S. (Eds). (2006). International conference on funding transportation infrastructure, Banff, Alberta, Canada (2-3/08/2006).

Mayeres, I., & Proost, S. (1997). Optimal tax and public investment rules for congestion type of externalities. *Scandinavian Journal of Economics*, *99*(2), 261–279.

Mayeres, I., & Proost, S. (2001). Marginal tax reform, externalities and income distribution. *Journal of Public Economics*, *79*(2), 343–363.

Mayeres, I., & Proost, S. (2003). Acceptability of transport pricing strategies: An introduction. In: J. Schade & B. Schlag (Eds), *Acceptability of transport pricing strategies* (pp. 93–106). Amsterdam: Elsevier.

Nowlan, D. M. (1993). Optimal pricing of urban trips with budget restrictions and distributional concerns. *Journal of Transport Economics and Policy*, *27*(3), 253–276.

Parry, I. W. H., & Bento, A. (2001). Revenue recycling and the welfare effects of road pricing. *Scandinavian Journal of Economics*, *103*(4), 645–671.

Quinet, E. (2001). European pricing doctrines and the EU reform, IMPRINT Seminar, Brussels, November 21–22.

Quinet, E., & Vickerman, R. (2004). *Principles of transport economics*. Cheltenham and Northampton: Edward Elgar Publishing Ltd.

Raux, C., & Souche, S. (2004). The acceptability of urban road pricing: A theoretical analysis applied to experience in Lyon. *Journal of Transport Economics and Policy*, *38*(2), 191–216.

Richards, M. G. (2005). *Congestion charging in London: The policy and the politics*. Houndsmill, Basingstoke, Hampshire and New York: Palgrave Macmillan.

Safirova, E., Gillingham, K., Parry, I.W.H., Nelson, P., Harrington, W., & Mason, D. (2004). Welfare and distributional effects of HOT lanes and other road pricing policies in metropolitan Washington, DC. In: G. Santos (Ed.), *Road pricing: Theory and evidence. Research in transportation economics* (Vol. 9, pp. 179–206). Amsterdam: Elsevier Science.

Saleh, W. (2005). Road user charging: Theory and practice. *Transport Policy*, *12*, 373–376.

Small, K. A., & Verhoef, E. T. (2007). *Urban transportation economics*. Chur, CH: Harwood Academic Publishers.

Tirole, J. (1988). *The theory of industrial organization*. Cambridge: MIT Press.

UK Department for Transport (2004). Feasibility Study of Road Pricing in the UK – Report, http://www.dft.gov.uk/stellent/groups/dft_roads/documents/divisionhomepage/029709.hcsp, accessed July 28, 2004.

Verhoef, E. T. (2002). Second-best congestion pricing in general static transportation networks with elastic demands. *Regional Science and Urban Economics*, *32*, 281–310.

Vickerman, R. (2006). Risk, regulation and rates of return: Problems for public private partnerships, pricing, financing, and investment in transport, 3rd International Kuhmo Conference and Nectar Cluster 2, Tuusula, Finland, 11–14 July 2006.

Vickrey, W. S. (1955). A proposal for revising New York's subway fare structure. *Journal of the Operations Research Society of America*, *3*, 38–68.

Walker, G. (1956). Highway finance. *The Journal of Industrial Economics*, *4*(3), 161–178.

PART I:
THEORY

This page intentionally left blank

CHAPTER 2

TRANSPORT USER CHARGES AND COST RECOVERY

André de Palma and Robin Lindsey

ABSTRACT

According to the celebrated cost recovery theorem, the degree of cost recovery from optimal user charges depends on the extent of scale economies in user costs, infrastructure construction costs and operating costs. This chapter presents the theorem, and reviews various generalisations of it. It then summarises empirical evidence by transport mode on the degree of scale economies or diseconomies in usage and in infrastructure, and the predicted surpluses or deficits with efficient pricing and investment. It also discusses some of the practical challenges in translating the cost recovery theorem into policy.

1. INTRODUCTION

A long-standing question in transportation economics is to what degree the revenues from user charges cover the capital and operating costs of transportation infrastructure. Many factors come into play: the objectives of the facility operator; the scope and flexibility of the charges; the functional dependence of user costs on traffic volume and capacity; the mix of user types; the degree of cost economies in capacity investment; capacity

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 29–57

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19002-8

indivisibilities; competitiveness of factor markets and so on. The main goal of this chapter is to assess how the level of cost recovery depends on these considerations.

It is useful at the outset to identify the scope and nature of the analysis. First, it focuses on user charges based on marginal-cost pricing. There are several reasons for this emphasis. First, the pursuit of productive and allocative efficiency calls for marginal-cost pricing under first-best conditions, and marginal-cost pricing serves as a starting point in the design of second-best pricing schemes when first-best conditions do not hold. Second, it has been European Commission (EC) policy since the mid-1990s to promote marginal-cost pricing in transportation and the REVENUE project is part of the corresponding research stream.

Finally, the leading alternatives to marginal-cost pricing incorporate cost recovery (either partial or full) as a constraint.These alternatives include average-cost pricing, Ramsey–Boiteux pricing, non-linear pricing and various cost-allocation methods. Such schemes are designed to meet a given cost recovery target at minimum social deadweight loss (e.g. as with Ramsey–Boiteux pricing) and/or in adherence with some notion of fairness (e.g. as with cost allocation). This chapter concerns itself not with optimal departures from marginal-cost pricing, but rather with the degree of cost recovery as an *outcome* of marginal-cost pricing.[1]

A second feature of the analysis is that it is based wholly on partial-equilibrium models.[2] No account is taken of distortions outside the transportation sector such as the excess burden of income taxes or rigidities in labour markets. The use of partial-equilibrium models is effectively dictated by the literature which, since the seminal work of Mohring and Harwitz (1962), has relied almost exclusively on partial-equilibrium analysis. Furthermore, all the case studies in Part II use partial-equilibrium models.

A third aspect of the analysis is its piecemeal character. This, too, reflects the approach generally taken in the literature of relaxing just one or two assumptions of the Mohring and Harwitz (1962) model at a time. A piecemeal approach has the advantage of making the analysis more tractable and transparent. Simulation models, such as those adopted by the case studies in Part II, are generally necessary if a number of relaxations and extensions are undertaken simultaneously.

Section 2 provides a statement of the Mohring–Harwitz cost recovery theorem, and comments on the main lessons of the theorem. Section 3 examines whether the cost recovery theorem continues to hold when assumptions are relaxed. Section 4 summarises empirical evidence on the key assumptions and parameters underlying cost recovery. Section 5 reviews some of the practical challenges in translating the cost recovery theorem into policy. Section 6 concludes.

2. THE COST RECOVERY THEOREM

The cost recovery theorem is due to Mohring and Harwitz (1962). The theorem and various generalisations of it are reviewed in a number of surveys.[3] Coverage here will be brief and relatively informal. We begin with a general version of the theorem that will serve as a benchmark for the various extensions reviewed in Section 3 and the empirical/policy applications in Sections 4 and 5.

2.1. Statement of the Cost Recovery Theorem

Let N denote usage of a facility and K denote the facility's capacity. Assume that the user cost function $c(N,K)$ has partial derivatives $c_N > 0$, $c_K < 0$ and $c_{NK} < 0$. Where $c(N,K)$ is assumed to be homogeneous, the degree of homogeneity is denoted by h. Let $F(K)$ be the annualised capital (or capacity) and operating cost of a facility of capacity K,[4] and define $\varepsilon \equiv (F_K K)/F$ as the (local) cost elasticity of $F(K)$. Denote by $p(N)$ the inverse demand curve with derivative $p_N < 0$. Finally, let R denote the revenue from user charges and define the cost recovery ratio $\rho \equiv R/F(K)$. The following version of the cost recovery theorem (henceforth, the CRT) is derived in the appendix.

The Cost Recovery Theorem. Assume $c(N,K)$ is homogeneous of degree h, and capacity is perfectly divisible. Then, with marginal-cost pricing and optimal (social-surplus maximising) capacity, the cost recovery ratio is

$$\rho = \varepsilon + \frac{h \times c(N,K)N}{F(K)} \tag{1}$$

It follows as a corollary of the CRT that if the user cost function exhibits constant returns ($h = 0$) and capacity is supplied at constant marginal cost ($\varepsilon = 1$), then $\rho = 1$ and user charges just suffice to finance optimal capacity: the facility earns a zero surplus and is self-financing.[5]

Henceforth Eq. (1) will be called the *cost recovery* formula. Eq. (1) is problematic for determining the *extent* of cost recovery inasmuch as it depends (if $h \neq 0$) on N and K which are endogenous variables. A closed-form solution for ρ can be derived if $c(N,K)$ and $F(K)$ have constant elasticities with respect to N and K:

The constant elasticity functions

$$c(N,K) = d\left(\frac{N}{K}\right)^{s} N^{h}, \quad d>0 \text{ and } s>0 \tag{2}$$

$$F(K) = kK^{\varepsilon}, \quad k>0 \text{ and } \varepsilon>0 \tag{3}$$

Given (2) and (3), it is easy to show that the cost recovery ratio is

$$\rho = \varepsilon\left(1+\frac{h}{s}\right) \tag{4}$$

The cost recovery ratio deviates from ε further the larger the absolute value of h; i.e. the greater the degree of economies or diseconomies of scale in usage.[6]

2.2. Lessons from the Cost Recovery Theorem

The CRT is justifiably celebrated as a landmark theoretical result in transportation economics as well as in the economics of public utilities. The CRT prescribes the fraction of total costs that should be paid for by user fees under first-best conditions. If actual cost recovery falls short of this amount, the CRT gives policymakers a case for increasing user charges (Arnott & Kraus, 1998b) and consequently for reducing reliance on other, typically distortionary, taxes (Verhoef & Rouwendal, 2004).[7] Likewise, if cost recovery exceeds the stipulated fraction, users can argue that charges should be reduced. And if the cost recovery formula indicates that a facility should be exactly self-financing, the CRT is consistent with the *user-pays* principle that users should cover the full costs of the infrastructure used to serve them.

When the conditions for full cost recovery apply, the CRT offers another lesson in the form of a signal when capacity should be adjusted.

The Investment Rule (adjustment to optimal capacity). Assume the user cost function $c(N,K)$ is homogeneous of degree 0, capacity is perfectly divisible and the capital cost function is linear; i.e. $\varepsilon = 1$. Then if toll revenues exceed capital costs, capacity should be expanded. Contrarily, if revenues fall short of capital costs, capacity should be reduced or allowed to depreciate.[8]

The intuition for the *Investment Rule* is that under the stated conditions, the long-run average total cost curve is horizontal and with efficient (competitive) pricing and investment, profits are zero. The *Investment Rule* has appeal from a practical and ethical as well as a theoretical perspective. As Roy (2005, p. 10) remarks, in the case of a surplus "the transparency of the new revenues would make it all but impossible for government to deny justified claims for new investment."

Three caveats about the *Investment Rule* deserve highlighting. First, the rule does not apply if the conditions for self-financing do not hold because the sign of profits then does not provide a reliable investment signal. In particular, if there are scale diseconomies then profits will be earned at the optimum but capacity should not be expanded. It is therefore practically useful to know how closely the self-financing conditions are met in a given situation.

A second caveat is that all user charge revenues should not be invested in new capacity. Since $F(K)$ is the annualised cost of capacity, it includes the cost of debt incurred to finance the capacity as well as that part of the operating costs that do not depend on usage. At least some of the revenues must therefore be devoted to balance the books, and at a steady-state optimum nothing remains to expand capacity. Finally, the *Investment Rule* indicates only whether capacity should be increased or decreased, not by how much. It therefore does not prescribe whether a given investment is worth undertaking. This is a significant limitation if capacity is lumpy as discussed in the next section.

3. GENERALISATIONS OF THE COST RECOVERY THEOREM

The CRT has been extended in a number of directions. The extensions of most relevance to the REVENUE project are summarised here with a view to determining whether the cost recovery formula remains valid and, if it does not, whether cost recovery is likely to be larger or smaller than what the formula prescribes.

3.1. User Heterogeneity

Users in the Mohring–Harwitz model are identical except for differences in their willingness to pay for usage. In practice passenger and freight vehicles usually share the road. More generally, users differ in their costs of travel time, trip-timing preferences and other characteristics. And the vehicles they occupy differ in size, acceleration capability, emissions and so on.

To analyse the implications of heterogeneity, Arnott and Kraus (1995) assumed there are G groups of users. The user cost function for group g, $c_g(N_1, \ldots, N_G, K)$, is assumed to be a homogeneous-of-degree-zero function of the amount of usage by each group. Arnott and Kraus show that the cost recovery formula ($\rho = \varepsilon$) holds up if marginal-cost pricing can be implemented separately for each user group.

This result raises two questions: whether charges should be differentiated by group, and if so whether differentiation is feasible. The answers depend on how groups differ. If they differ only in their values of time then the optimal toll is independent of user type because values of time affect only private costs. But if users vary in the congestion costs they impose, differentiated tolls are called for. Whether this is practical depends on whether the individual or vehicle features that contribute to differences in external costs are observable. Vehicle attributes such as numbers of axles and configuration of trailers are readily observable. But other features, such as vehicle weight or driver attentiveness, are less easily measured and there may be legal barriers against discriminating on the basis of them.[9]

If vehicles differ in speed, scale economies in usage may exist because fast vehicles can pass slow vehicles more easily when there are two or more traffic lanes in each direction, multiple rail lines, etc. And if there are neutral- or positive-scale economies in infrastructure as well, a deficit will result.

What can be said about cost recovery if tolls cannot be fully differentiated? If no differentiation is possible then the optimal uniform charge is a weighted average of the first-best group-specific charges, with weights that are proportional to the toll elasticities of group demand as in the classical Ramsey-pricing formula. A deficit will result if groups that generate below-average marginal external costs have above-average absolute demand elasticities, and vice versa. The correlation between costs and elasticity is an empirical question that will depend on the mode and region, and no general conclusions are possible.

3.2. Indivisibilities in Capacity

One of the assumptions of the CRT is that capacity is perfectly divisible. Yet many components of transport infrastructure are lumpy such as traffic lanes, tunnels, rail lines and airport runways.

Where indivisibilities are present, it is generally not possible to choose a capacity where marginal benefits and costs are equal, and the cost recovery formula does not hold. Following Kraus (1981a), this can be illustrated by considering a road on which the number of lanes in both directions is restricted to be an integer multiple of 2. The cost of capacity is assumed to be $F(L) = F_0 + c_L L$, where L is the number of lanes and F_0 a fixed cost that accounts for the cost of shoulders and other highway elements that do not depend on the number of lanes. User costs are assumed to be homogeneous of degree zero. Without indivisibilities the Long Run Marginal Cost (LRMC) curve would be a horizontal line, and the long-run average cost (LRAC) curve would be downward sloping as shown in Fig. 2.1. At the optimum, revenues would fall short of costs by F_0, and the cost recovery ratio would be

$$\rho = \varepsilon = \frac{\mathrm{d}F(L)}{\mathrm{d}L}\frac{L}{F(L)} = \frac{c_L L}{F_0 + c_L L} < 1$$

With lane indivisibilities the situation is quite different. The short-run average cost (SRAC) curve corresponding to two traffic lanes is depicted in Fig. 2.1 by the curve SRAC(2) and the corresponding short-run marginal cost (SRMC) curve is SRMC(2). Similarly, the short-run average and marginal curves for four lanes are SRAC(4) and SRMC(4), respectively. If capacity

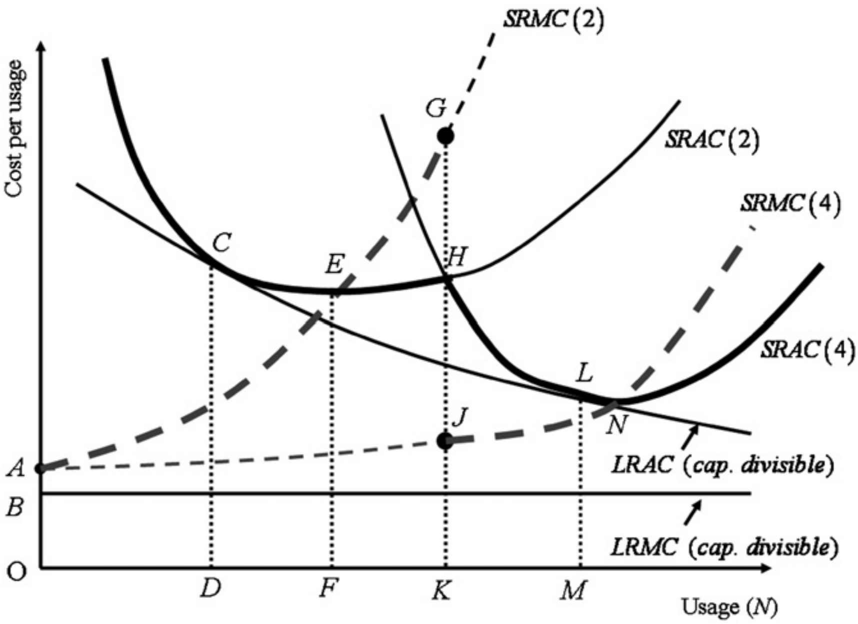

Fig. 2.1. Long-Run and Short-Run Cost Curves with Capacity Indivisibilities. *Source:* Adapted from Kraus (1981a, Fig. 1)

choice is limited to two lanes or four lanes, the LRAC curve is the scalloped lower envelope of SRAC(2) and SRAC(4) shown by the thick solid line passing through points C, E, H, L and N. And the LRMC curve is the thick broken line passing through points A, E, G, J and N, which has a discontinuity at the level of usage *OK*.

The irregular shapes of the LRAC and SRMC curves when there are indivisibilities have several implications for cost recovery. First, if the demand curve (not shown in Fig. 2.1) intersects the LRAC curve on an upward-sloping segment, i.e. between points E and H or to the right of point N, then marginal cost exceeds average cost and a surplus will result even though the LRAC curve with no indivisibilities slopes downward everywhere. Second, even if the demand curve intersects the LRAC curve on a downward-sloping segment, unless the intersection occurs at point C or point L, where LRAC is tangent to one of the SRAC curves, the deficit will not be equal to F_0. Finally, the demand curve could intersect both the SRMC(2) and SRMC(4) curves at points where social surplus is equal. If so, a choice will exist between multiple optima with potentially very different cost recovery ratios.

Do indivisibilities in capacity undermine self-financing in practice if the other assumptions of the CRT hold? As Verhoef (2006) notes, the answer appears to be yes if there is a minimum feasible capacity constraint that binds on a large fraction of the links on the network (which may be true for roads in rural areas). But if traffic levels vary strongly over the network, some links will produce surpluses and other links deficits, and the surpluses and deficits may approximately cancel out in aggregate. Furthermore, if demand grows over extended periods of time, an alternating series of deficits and surpluses may be generated on each link as its capacity is periodically increased. This possibility is examined under the topic of "Non-Stationary Environments" in Section 3.4.

3.3. Time-Varying Demand

The CRT has been extended to account for systematic fluctuations in demand. The most general treatment, due to Arnott and Kraus (1998a), allows for interdependence in demand between time periods, multiple elements of capacity and heterogeneity in user types. Arnott and Kraus show that, under homogeneity assumptions analogous to those of the Mohring–Harwitz model, the cost recovery formula holds if user charges can be varied freely over time. This requirement parallels the requirement in Arnott and Kraus (1995) that user charges can be differentiated across user types. Indeed, the two results are isomorphic if demand is independent between periods since individuals who use the facility in different periods can then be viewed as different types.

The requirement that user charges be freely variable over time is stringent. Although modern technology permits road tolls to be varied frequently over time, the infrastructure can be expensive. And various legal and acceptability barriers continue to impede the implementation of complex road-pricing schemes as well as time-of-day dependent charges for use of runways and other transportation facilities.

A general analysis of cost recovery with time-constrained charges has yet to be undertaken. Bichsel (2001) provides some insights using a simple peak/off-peak model. He shows that constraints on time variation can result in either a surplus or a deficit, although a deficit appears likely. The reason is that (similar to the case with heterogeneous users) the optimal uniform charge is an elasticity-weighted average of the optimal peak and off-peak charges. Since demand is likely to be more elastic in the off-peak, the optimal uniform charge is heavily weighted towards the lower off-peak charge, which drives revenues down. Constraints on toll variations therefore appear to militate against full cost recovery.

More flexible tariffs tend to be more costly to administer (see the discussion of administration costs in Section 5). However, optimal uniform tariffs require information about own- and cross-price demand elasticities that is also costly to collect. A trade-off thus exists between the complexity of administering time-varying charges and the amount of information required to set optimal uniform or time-constrained charges.

3.4. Non-Stationary Environments

The analysis of time-varying demand in the previous sub-section is limited to short time intervals during which capacity is fixed. Over longer time horizons it may be possible to alter capacity. And with depreciation of capacity, technological progress, changes in regulations and other developments, capacity may need to be both added and replaced.

Arnott and Kraus (1998b) investigate the implications of such non-stationary conditions for cost recovery under marginal-cost pricing. They use a series of independent model specifications that cover depreciation and maintenance, adjustment costs and irreversibility, intermittent capacity additions or replacements and fixed or lumpy capacity increments. They also allow for the discount rate to vary over time, but assume that users are identical except for their willingness to pay.

Arnott and Kraus show that for all their model specifications except capacity indivisibilities, the CRT holds in present-discounted-value (PDV) terms: PDV revenues cover the initial cost of capacity plus the PDV of investment, maintenance and adjustment costs. Technically, this is because the cost function exhibits constant long-run ray average costs which, under marginal-cost pricing, results in breakeven operation (Baumol, Panzar, & Willig, 1982, §13F). If capacity is indivisible, the CRT does not hold because long-run ray average costs are not constant (in one dimension the LRAC curve is scallop-shaped, as in Fig. 2.1). Thus, cost recovery does not require that capacity additions be made at an optimal time, but it does require that the size of capacity additions be optimal conditional on the timing.

3.5. Infrastructure Damage

To this point, the analysis has focused on the revenues derived from congestion charges. Infrastructure damage is another external cost that users bear in aggregate. Discussion here will be limited to road damage because of its large costs, and because Pigouvian taxation of road damage externalities has been analysed in the literature as well as implemented in various guises in a number of countries.

There are significant scale economies in road durability with respect to pavement thickness. In light of the CRT, this might suggest that road damage charges will cover only a small fraction of pavement costs and therefore result in a large ‘‘pavement deficit’’. However, this is not necessarily true if congestion pricing is implemented in tandem with road-damage charges. The reason for this is that strengthening roads increases capacity costs as well as durability costs, and widening roads increases durability costs as well as capacity costs. As Small, Winston, and Evans (1989) explains, there are diseconomies of scope in supplying capacity and durability jointly.

By using different models, Newbery (1988, 1989) and Small et al. (1989) show that an optimal combined system of road congestion and damage charges approximately pays for capital construction costs and maintenance costs. Costs are recovered even if the time interval between pavement overlays is not optimal but rather determined by a condition-response rule based on the roughness of the surface.[10]

3.6. Networks

The CRT was derived by Mohring and Harwitz for a single isolated facility. But as Yang and Meng (2002) show, if the assumptions of the CRT hold for each link on a network then the cost recovery formula holds for each link. Formally, the cost recovery ratio on link a equals the local elasticity of the capacity cost function on link a; $\rho_a = \varepsilon_a$. Urban road networks generally comprise various link types (expressways, arterials, local streets, connectors, etc.) with different cost elasticities, and optimal tolling may therefore yield surpluses on some links and deficits on others.

As a corollary it follows that if each link can be expanded at constant marginal cost ($\varepsilon_a = 1$ for all a) then each link will be self-financing. Thus, not only will user charges in aggregate pay for the whole network, but also each link will cover its own costs. Consequently, the principle of user-pays will be satisfied locally as well as globally.

As noted in Section 3.2, if capacity is lumpy, the cost recovery formula is inapplicable because capacity cannot be chosen optimally at the margin. However, if capacity on a particular link a is divisible, then the cost recovery formula holds for link a as long as all links on the network are priced optimally conditional on their capacities.[11] The capacities of other links do *not* have to be optimal (either because of indivisibilities or for other reasons).

Optimal pricing is necessary for the CRT to hold, and this generally requires that charges vary across links. The practicality of differentiation depends on the transport mode and on circumstances. As far as road pricing is concerned, it is not technologically difficult to set different tolls on different links, or at least to differentiate tolls by type of link as Singapore does with its electronic road-pricing system. However, setting different tolls on each link may be confusing for drivers on dense urban road networks. Where toll differentiation is not possible, the optimal constrained toll is (again) an elasticity-weighted average of first-best link tolls.

A more serious practical limitation on applicability of the CRT is that it holds only if all links are tolled. At least for roads this is unlikely to happen for some time because of tolling infrastructure and operating costs, political opposition and other implementation barriers. And if only some links can be tolled, the tolls should be set according to second best rather than first-best pricing principles. Second-best pricing on networks is a relatively new research area, and little attention has yet been given to cost recovery. A few insights have been derived by considering two-link networks with one link untolled.

In the case of two parallel, substitute, links the toll on the tolled link is set below the first-best level in order to alleviate congestion on the untolled link. If capacity of the tolled link can be optimised, the first-best capacity rule applies and the degree of cost recovery on the tolled link is unambiguously below the first-best standard (Verhoef, 2006). The larger the capacity of the tolled link and the shorter its free-flow travel time relative to the untolled link, the higher is the cost recovery ratio because the tolled link then carries a greater fraction of total traffic.

If the two links are in series, rather than parallel, they are complements rather than substitutes. Tolling one link reduces congestion on the other link, and consequently the toll is set above the first-best level. The first-best capacity rule again applies, and cost recovery on the tolled link exceeds the first-best standard (Verhoef, 2006).

On large networks a link can be a substitute for some links, a complement for others and independent of yet others, and the lessons from the two-link networks are blurred. The degree of cost recovery will depend inter alia on how many links are tolled and where, and on the topology of the network.

3.7. Non-Optimal Pricing and/or Capacity

Pricing and capacity decisions may deviate from first-best optimality for various reasons. For example, prices may not be optimal due to constraints on differentiating tolls by user category, time of day or link (see Sections 3.1, 3.3, 3.6). And non-optimal prices or investments can result from errors in calculation or forecasting demand or costs. Socially optimal decisions will also not result if the operator is a profit maximiser with market power.

3.7.1. Non-Optimal Pricing

If usage is not optimally priced the question arises whether capacity will be set at its first-best level, or whether it will be adjusted to the second-best level conditional on the flawed pricing. In practice, this will depend on the relative timing of pricing and investment decisions, on whether capacity can be adjusted after it has been installed, on institutional factors and so on. Attention is limited here to the case of second-best capacity since this is reasonable if non-optimal pricing is either the norm or anticipated when capacity is chosen.

If prices are not set optimally, the envelope theorem does not hold, and the first-order condition for optimal capacity includes a term that accounts for change in usage:

$$\frac{\partial \Omega}{\partial K} = \underbrace{-c_K N - F_K}_{[a]} + \underbrace{[\tau - \tau^*(N, K)] \frac{\mathrm{d}N}{\mathrm{d}K}}_{[b]} = 0 \tag{5}$$

where Ω denotes social surplus, τ the non-optimal charge or toll and $\tau^*(N, K)$ the first-best optimal toll. Term [a] in Eq. (5) is the first-order condition for the first-best optimal capacity. Term [*b*] incorporates the effects of induced demand. To understand the effect of this term, suppose usage is underpriced. Since $\mathrm{d}N/\mathrm{d}K>0$, term [*b*] is then negative. This suggests that second-best capacity is less than first-best capacity. However, with a sub-optimal toll both N and c_K in term [*a*] are greater in magnitude than in the first-best case.[12] The net effect of terms [*a*] and [*b*] is ambiguous a priori, and little can be said in general without specific assumptions about the user cost and demand functions.[13] Moreover, large deviations from optimal pricing may have qualitatively different effects from the marginal deviations described in Eq. (5). However, there is some presumption that cost recovery will be enhanced if prices are above optimal levels, and depressed if prices are set too low.

3.7.2. Non-Optimal Capacity

The effects of non-optimal capacity on cost recovery do not appear to have been systematically analysed in the literature despite many examples of egregious overinvestment in urban transit systems, airports and other transport infrastructure. As far as marginal deviations from optimal capacity are concerned, it is possible to show that

$$\varepsilon_{\rho K}<0 \quad \text{if} \quad \rho \geq 1$$

where $\varepsilon_{\rho K}$ is the elasticity of cost recovery with respect to capacity. Thus, if a facility breaks even or earns a surplus at the first-best optimum ($\rho \geq 1$) then overbuilding weakens cost recovery. This is the case because the user charge falls when capacity rises, and demand does not increase enough for revenues to keep up with capital costs.

3.7.3. Pricing and Capacity Choice for Profit Maximisation

An appreciable fraction of transport infrastructure in the European Union (EU) and elsewhere is operated by the private sector, which is typically interested in maximising profit rather than welfare. Although natural monopoly is rare in transport, and unregulated monopoly is virtually non-existent, it is nevertheless instructive to consider the pricing and capacity choices of an unregulated monopolist since the lessons are likely to apply (albeit in a diluted form) to oligopolistic markets.[14]

Monopolistic management of congestible facilities has been well studied in the literature (e.g. Small & Verhoef, 2007, Section 6.1). The monopolist sets a price above the first-best level by a markup that reflects market power. But the monopolist follows the first-best rule for choosing capacity, and therefore fully internalises the congestion cost borne by users.[15] The monopolist's capacity choice decision can therefore be treated as in Section 3.7.1 as a case of second-best optimal capacity given distorted pricing. There is a strong presumption that cost recovery for the monopolist will exceed the first-best level.

3.8. Uncertainty

Transport infrastructure is subject to uncertainty about construction costs, capacity availability and demand for usage. Major construction cost overruns are common (Flyvbjerg, Bruzelius, & Rothengatter, 2003). And once a facility is operational, available capacity can fluctuate unpredictably due to accidents, mechanical breakdowns, bad weather and so on. Demand, too, is subject to unanticipated peaks and troughs, and predictions of average demand levels can be well off the mark as evidenced by the poor track record of traffic volume forecasts for toll roads.

Despite its evident importance, the implications of uncertainty for self-financing have not been studied. One consideration is whether user charges can be adapted to changes in capacity and demand. On most existing tolled facilities charges are set according to a predetermined schedule, and cannot be adjusted – at least to short-run fluctuations. However, responsive pricing in near real-time has been successfully implemented on two toll roads in the U.S. and the practice could eventually become widespread.

3.9. Other Externalities

In addition to congestion and road damage, transportation generates pollution, noise, accidents and other external effects. Pigouvian taxes can be levied on these externalities that will generate revenues additional to those from congestion and infrastructure damage charges. However, except for accidents, the costs of these other externalities are not primarily borne by users[16] and there appears to be no prima facie case for using charge revenues to finance transport infrastructure; indeed, doing so would tend to exacerbate rather than alleviate the externalities. In any case, charges for greenhouse gas emissions and other external effects can be accommodated in the modeling framework by assuming that the inverse demand curve for travel is net of these charges.

4. EMPIRICAL EVIDENCE

This section briefly reviews empirical evidence on the primary determinants of cost recovery under first-best conditions: cost economies of scale in usage, cost economies of scale in infrastructure and capacity indivisibilities. Implications for cost recovery are then briefly reviewed for roads.

4.1. Cost Economies of Scale in Usage

As is noted in Section 2.1, capacity can be defined in such a way that the capacity cost function $F(K)$ has constant returns to scale. For empirical analysis, however, it is useful to work with natural units of capacity such as traffic lanes and rail lines, and this approach will be taken here. It is often assumed that doubling the number of traffic lanes allows twice as many vehicles to travel at the same cost. But adding lanes facilitates passing, and more generally allows users to travel at their preferred speeds. User scale economies of this sort can be particularly significant for railway infrastructure that serves both freight and passenger trains.

Urban bus and rail systems are characterised by economies of traffic density because average waiting times fall when demand rises and service frequency is increased (the Mohring effect). According to Mohring's (1972) square-root rule, usage costs are homogeneous of degree $h = -0.5$. But once account is taken of various complications – such as the spacing of bus stops, delays as passengers board and alight, vehicle capacity constraints and peak-period demands – usage economies are found to be rather smaller (Jansson, 1997; Tisato, 1991, 1998). To the extent that usage economies still exist, marginal costs fall short of average costs and marginal-cost pricing will result in a deficit.

4.2. Cost Economies of Scale in Infrastructure

Scale economies vary by mode. The empirical evidence is briefly reviewed here for roads, rail and public transport.

Roads. Keeler and Small (1977) found statistical evidence of constant or slightly increasing returns to scale for roads (their point estimate of scale economies is 1.03 with a standard error of 0.39). Using engineering data, Kraus (1981b) found moderate increasing returns (point estimate of scale economies of 1.19), which reflect the net effect of substantial scale economies for individual road segments and diseconomies for intersections. More recently, Levinson and Gillen (1998) identified mild scale diseconomies for passenger cars, but substantial economies for single trucks and even larger economies for combination trucks.[17]

As noted in Section 3.5, there are substantial scale economies in durability with respect to pavement thickness. But Small et al. (1989) conclude that because of diseconomies of scope with respect to road width and durability, charging for both congestion and damage charges would lead to recovery of at least 80% of long-term capital and maintenance costs. Summing up the evidence for urban areas Small and Verhoef (2007, p. 77, Section 3) conclude:

> Altogether, the evidence supports the likelihood of mild scale economies for the overall highway network in major cities. Scale economies are probably substantial in smaller cities in which one or two major expressways are important, and may disappear altogether in very large cities where expanding expressways is extraordinarily expensive due to high urban density.

Rail. Economies of traffic density in railways prevail except at very high densities that are rarely attained.[18] Train planning and operations costs are mostly fixed, especially since modern signaling systems are capital – rather than labour – intensive (Nash, 2005). Maintenance costs also exhibit substantial scale economies although the evidence is somewhat confounded with the cost of renewals. Evidence from the U.S. and western European railways indicates that the average cost curve is U-shaped with respect to network size (Preston, 1996). Small railways have advantages in terms of workforce flexibility and management responsiveness, but disadvantages for procurement and equipment utilisation. The evidence on (dis)economies of scope for passenger and freight services is mixed.

Urban public bus and rail transport. Despite a large empirical literature on urban public-transport cost functions, the results are somewhat inconclusive (Berechman, 1993). Any scale economies appear to be small, and are exhausted for fleets comprising more than 300–500 vehicles. Similar to intercity railways, public transport enjoys substantial economies of traffic density. Due to capital grants and political pressure to supply enough capacity to meet peak-period demands without major delays, transit operators typically maintain massive excess capacity. This contributes to their failure to recover costs.

4.3. Capacity Indivisibilities

A consensus has not been reached on the practical significance of indivisibilities in road capacity. Capacity is inherently discrete because the number of traffic lanes is integer valued. This indivisibility may be consequential in rural areas where the road network is sparse (Kraus, 1981a; Heggie & Fon, 1991), although capacity can still be adjusted by varying lane width, lateral clearance, horizontal and vertical alignments, etc. Railway capacity appears to be adjustable more smoothly by using traffic management systems.

In the presence of indivisibilities the LRAC curve has a saw-toothed shape (cf. Fig. 2.1). If the unit cost of capacity expansion is constant, then with growing demand the optimal investment schedule will be characterised by alternating periods of surplus and deficit – although as discussed in Section 3.4, cost recovery may still be achieved in present-discounted-value terms.

4.4. Implications for Cost Recovery of Roads

Despite general agreement about the theory, as well as about some pieces of the empirical evidence, assessments vary on the degree of cost recovery for roads under first-best pricing. Based on in-depth empirical estimates for major European countries in ECMT (2003) and UK Department for Transport (2004), Roy (2005, p. 8) concludes that "optimal pricing will generate revenues in excess of *both* current infrastructure costs *and* the costs of justified investments". Likewise, Gómez-Ibáñez (1992, pp. 357–358) reasons that, for the U.S., short-run marginal-cost pricing will yield large surpluses.[19] For other areas, however, deficits may result. Results from the Pricing European Transport Systems (PETS) project indicate that marginal social cost pricing (MSCP) will fail to cover investment costs in parts of the Nordic countries with low traffic volumes and low population densities, where external costs are correspondingly low (Sikow-Magny, 2003, p. 22). Also for developing countries Heggie and Fon (1991) conclude that due to substantial scale economies in both road construction and road usage, as well as capacity indivisibilities and the high proportion of maintenance costs that are fixed, a substantial deficit will result.

5. PRACTICAL AND MODELING COMPLICATIONS

In this section we briefly address some considerations that complicate application of the Cost Recovery Theorem.

5.1. Practical Complications

5.1.1. Accounting Practices

The Mohring–Harwitz model that underlies the CRT treats capital as a homogeneous and timeless input. The model bypasses practical complications such as technological change and accounting practices that may change over time or differ between levels of government or jurisdictions. The difficulties are aptly described by Ekelund and Hébert (1999, p. 184):

> Turvey has reminded us that perhaps too much has been made of the proposition that marginal-cost pricing necessarily involves a loss whenever decreasing costs (increasing returns) exist. Such a proposition is far too particularised because it requires some kind of imaginary long run in which ''capacity'' is not merely ''variable'' but plant is all newly built at today's prices using today's technology. Real-world problems surrounding publicly provided goods are more mundane but more complex. The important issue, according to Turvey, is whether forward-looking marginal costs, calculated from enhancing or reducing plans for expanding an *existing* system, meet or fall short of revenue requirements based on backward-looking accounting costs.

Accounting difficulties are also reviewed by Newbery (2005) in the context of establishing a regulatory asset base for a public roads authority in the UK.

5.1.2. Collection Costs

The Mohring–Harwitz model neglects the costs of collecting user charges and the costs borne by users in paying them. Collection costs can be incorporated into the model by treating them as a component of system-operating costs that are included in user charges. To the extent that there are scale economics in collecting charges a deficit will result. Naturally, collection costs are relevant for cost–benefit analysis of charging systems (Prud'homme & Bocajero, 2005; Mackie, 2005). For road-pricing schemes, the costs vary widely as a fraction of revenues[20] and the relatively high costs for the German HGV motorway charge have been an issue (see Chapter 10 of this volume).

5.1.3. Implementation Issues

In the cost recovery literature, decisions on pricing and revenue use are assumed to be made by benevolent agents acting for the public good. In practice, decision makers further their own interests. Politico-economical and contracting issues will be addressed in Chapters 3 and 4 of this volume. Attention is limited here to a few remarks about the scope for decentralising pricing, maintenance and investment decisions to infrastructure operators.

As Arnott and Kraus (1998b) and Arnott (2006) point out, it is desirable to establish pricing, maintenance and investment rules that are based on observables since this not only makes the operator's task easier, but also facilitates oversight. An example is the *Investment Rule* under constant returns to scale (see Section 2.2), whereby surpluses provide a signal to expand capacity and deficits a signal to let capacity depreciate. Unfortunately, investment rules become more complicated and less transparent in a dynamic environment with time-varying interest rates, technical change and adjustment costs (Arnott & Kraus, 1998b).

Implementation difficulties may also arise in determining parameter values relevant to cost recovery such as the degree of scale economies and price elasticities that must be estimated on the basis of imperfect information and limited data. As Arnott (2006) remarks, there is a danger that a numbers game will develop in which facility operators commission studies that point to low levels of cost recovery and corresponding high subsidies, and governments commission independent studies that reach contrary conclusions.

5.2. Modeling Complications

5.2.1. Congestion and Scarcity in Rail Transport

In the Mohring–Harwitz model, user costs $c(N,K)$ are assumed to be differentiable functions of usage and capacity, and in the homogeneous case they depend only on the ratio of usage to capacity. Rising user costs are ascribed to congestion-induced travel time delays. This formulation and interpretation are generally viewed to be suitable for roads. Applications to other modes of transport are not as clearcut, and they are problematic in the case of rail transport which is featured in the three interurban case studies (Chapters 9–11 of this volume).

Conflicts between rail users have a different character from road traffic congestion. To access the tracks, a train operator needs to obtain a ‘‘path’’ that specifies the departure time, speed, stopping pattern, etc. If the desired path conflicts either with paths that have already been allocated to other users, or with maintenance operations, then the path is unobtainable and ‘‘scarcity’’ is said to arise. Scarcity contrasts with congestion, which is caused by movements of other trains while the path is being traversed.[21] Timetables are designed to avoid congestion. But congestion does occur at high capacity utilisation rates and when unexpected delays experienced by one train have knock-on effects on other trains.

It is generally argued (e.g. Quinet, 2003; Nash, 2005) that user charges alone are not sufficient to support efficient usage of rail infrastructure. In practice, path allocation rules are the primary tool for regulating access. Moreover, both existing and recommended rail charging schemes differ qualitatively from Pigouvian taxation (Gibson, Cooper, & Bal, 2002).[22]

5.2.2. Capacity and Service Quality

In the Mohring–Harwitz model, supply is described by a single variable called "capacity". An increase in capacity reduces user costs for a given number of users or, equivalently, permits greater usage at the same user cost. But capacity does not affect free-flow travel speed or elements of user cost that are not related to travel time. As Larsen (1993) points out, road infrastructure embodies characteristics other than capacity as just defined. He refers to these characteristics as a "standard" which he defines (1993, p. 274) as "all the aspects of a road system that influence driving speeds, driving distances and operating costs of vehicles under non-congested driving conditions."

Some investments in road infrastructure enhance both capacity and "standard"; for example, replacement of a signalised intersection with a ramp junction (1993, p. 274). If capacity and "standard" are joint products, investment should be extended beyond the level dictated by the usual optimal capacity formula. And since MSCP calls for user charges to internalize congestion externalities, but not to capture the benefits from an improved "standard", the CRT no longer holds and a deficit will result under conditions where full cost recovery would occur without a change in "standard". Surprisingly, Larsen's (1993) ideas have not been followed up with empirical research to assess the importance of the "standard" effect, and the degree to which cost recovery may be affected.

6. CONCLUSIONS

For over a decade the European Commission has been promoting MSCP of transportation and sponsoring research projects such as REVENUE. MSCP supports allocative efficiency by assuring that the social benefits and social costs of using transportation infrastructure balance at the margin. Users are faced with the full marginal costs of their activities including external costs such as congestion and emissions. Traditionally, however, tolls and other types of user fees have been levied to cover the costs of building and operating transportation facilities. Private operators are forced to generate sufficient revenues to balance their books. And state-operated facilities are often obliged to defray at least some of their costs through user fees rather than relying wholly on the public purse.

In theory, cost recovery calls for average-cost pricing at some (often ill-defined) level of aggregation over modes and user groups. Self-financing appears incompatible with MSCP in general. The tension has been evident ever since Dupuit (1849). It played a role in the "marginal-cost pricing controversy" that was debated by Coase (1946) and others during much of the twentieth century.[23] Many economists at the time thought that the twin goals of supporting efficient usage and recovering costs could not simultaneously be met with any practical system of charges.

Enlightenment came with the seminal work of Mohring and Harwitz (1962) which showed that under plausible assumptions MSCP would, in fact, just cover costs. The simple version of the CRT goes as follows: assume user costs are homogeneous of degree zero in capacity and usage, capacity is perfectly divisible and can be expanded at constant marginal cost. Then, with MSCP and optimal capacity, infrastructure costs will be exactly recovered.

The CRT is notable in several respects. First, it is mathematically elegant as far as paucity of assumptions and ease of proof. Second, self-financing is consistent with the user-pays principle. There will be no need either for cross-subsidies within the transport sector, or for transfers between the transport sector and the rest of the economy. Third, the CRT is central to the theme of the REVENUE project. REVENUE was set up to examine how the revenues from transport user charges should be used. The CRT provides a simple answer for a first-best world: use the revenues to finance, operate and maintain the infrastructure on which the charges are levied. The revenues will be both necessary and sufficient to fund infrastructure of optimal size. Revenues will effectively be earmarked without any allocative inefficiency that typically accompanies earmarking in practice.

Mohring and Harwitz's work spawned a substantial theoretical literature that has focused on the generality of the CRT. The CRT has proved to be surprisingly robust to relaxation of assumptions. It holds, mutatis mutandis, with user heterogeneity, time-varying demand, growing demand, usage-dependent depreciation, adjustment costs and irreversibility, constraints on the timing of investments and on transport networks. Indivisibilities in capacity are the major exception.

Empirical studies of cost recovery have concentrated on the factors that determine the degree of cost recovery: scale economies in usage, scale economies in infrastructure and divisibility of capacity. The empirical evidence varies by mode.

For roads there appear to be mild economies of scale in usage, at least in expanding capacity from two lanes to four lanes, because multiple lanes facilitate passing and allow vehicles to travel closer to their preferred speeds. Roads are found to exhibit constant to moderate scale economies in infrastructure. There are substantial scale economies in durability with respect to pavement thickness, but these are largely offset by diseconomies of scope with respect to road width and pavement thickness. Road capacity is inherently discrete because the number of traffic lanes is integer-valued, although capacity can be adjusted incrementally in various ways. Indivisibilities are probably most important in rural areas where the road network is sparse and traffic densities are low. Studies that take these various factors into account differ in their assessments of cost recovery. Studies for major European countries and the U.S. conclude that MSCP will result in surpluses, whereas for low-density developed countries and developing countries deficits are predicted.

For intercity passenger and freight rail transport, substantial economies of traffic density appear to exist for both user costs and infrastructure. The theory therefore predicts that major deficits will result under MSCP. Similar conclusions are drawn for urban public transportation. Deficits are indeed common for public transport systems, although pricing and investment decisions in most countries do not adhere to the theoretical prescriptions that have been reviewed in this chapter.

The strong theoretical credentials of the CRT are not in serious dispute. Several practical considerations, however, limit its effectiveness for guiding policy. One is that the CRT is derived for a first-best world. Competing and complementary links or transport modes are assumed to be priced at marginal social cost. And related markets in other sectors of the economy (e.g. labour markets linked to commuting trips) are assumed to be distortion free. These assumptions are clearly violated. Infrastructure pricing and investment decisions should therefore be formulated in the second-best world of multiple distortions, with all the attendant analytical complications that this brings.

A second limitation of the CRT is that the underlying model includes only congestion externalities, and assumes that investments in infrastructure serve only to reduce congestion. This makes the model problematic for rail transport where congestion delays are secondary in importance to the "scarcity" of train paths. And the model ignores the fact that in addition to relieving congestion, infrastructure investments can raise free-flow speeds, enhance safety and improve travel conditions in other ways. A third limitation of the CRT is that it does not deal with such practical complications as accounting practices for valuing infrastructure, or implementation issues such as how to devolve pricing, maintenance and investment decisions to infrastructure operators.

These and other practical complications are endemic in the real world of second best. And they are manifest in the infrastructure projects dealt with in the case studies of Part II of this book. In practice, pricing and investment decisions should not rely wholly on simple Pigouvian tax formulas or the

Investment Rule (see Section 2.2). Rather, they should be based on cost–benefit analysis using sophisticated assessment models such as the MOLINO model described in Chapter 5 of this volume. Furthermore, the degree of cost recovery cannot reliably be deduced from rules of thumb such as the cost recovery formula. Cost recovery is an outcome of the welfare optimisation process, which may encompass myriad complications excluded from the theory. The CRT, the cost recovery formula and the Investment Rule should be viewed as useful benchmarks that guide expectations about cost recovery and aid understanding of results.

NOTES

1. Much has been written, of course, about optimal (second-best) pricing subject to a revenue constraint. European Commission policy has tended to favour non-linear pricing in the form of two-part or multi-part tariffs (e.g. ECMT, 1998). Such schemes are in fact widely used across the EU and elsewhere. For example, fuel taxes and tolls are variable charges for road transport while annual vehicle licence and registration fees and vignettes constitute fixed charges. An extreme case of high fixed charges are the import duties, registration fees and Certificates of Entitlement in Singapore that in aggregate increase the cost of owning a vehicle several-fold. Multi-part tariffs have several attractive properties for transport pricing: they are potentially Pareto-superior to linear tariffs (including Ramsey–Boiteux pricing); they provide an additional degree of freedom in pursuing distributional goals; and if there is a menu of tariffs user groups can self-select on the basis of their extent of usage.

2. General equilibrium considerations will be addressed in Chapter 3 of this volume.

3. See Small (1992, Sections 3.4, 3.5), Hau (1998, 2005a, 2005b), Gómez-Ibáñez (1999), Lindsey and Verhoef (2000), and Small and Verhoef (2007, Section 5.1).

4. Operating costs are assumed to be independent of usage. This assumption can be relaxed without affecting the results as long as operating costs are included in user charges. The implications of usage-dependent depreciation are examined in Section 3.5. Capital costs are expressed using the same time units as user costs and demand.

5. The conditions for self-financing are actually less restrictive than the theorem suggests because capacity can be defined in such a way that $\varepsilon = 1$. Indeed, self-financing is possible even if both conditions are violated but in an offsetting way. This is the case in Eq. (4) below if $h = s(\varepsilon^{-1} - 1)$.

6. Since user cost is assumed to be an increasing function of usage, $s + h > 0$ in Eq. (2) and $\rho > 0$ necessarily.

7. Such fiscal considerations are studied in Chapter 3 of this volume. Care is required: if there are other distortions in the economy, it is necessary to account for the effects of any investment and transport pricing decisions on these distortions. This may call for lower – rather than higher – transport prices; e.g. reducing the cost of using a commuting link may encourage labour supply that is over-taxed.

8. Under the assumptions of the CRT, profits are $\pi = \tau N - kK = (\tau \cdot N/K - k)K$, where τ is the user fee and k a constant. The *Investment Rule* can be proved by

showing that both τ and N/K are decreasing functions of K. This method of proof is suggested in Verhoef and Rouwendal (2004, p. 407).

9. In the case of rail transport infrastructure, user charges can be levied in various ways, and practice varies widely across the EU (Nash, 2005, Table 3). Charges are assessed by path – km, train, train – km, vehicle – km and tonne – km, and the levels of charges vary with characteristics of trains and track (train weight, vehicle type, axle load, speed, track quality, location and time of day). In some countries charges are also levied for the use of stations, depots, marshalling yards and other services such as passenger information.

10. This is an instance of the general result derived by Arnott and Kraus (1998b) (see Section 3.4) that cost recovery holds as long as capacity increments are optimal in size conditional on their timing.

11. This result follows immediately from the envelope theorem.

12. c_K is larger because N is larger and the assumption (cf. Section 2.1) that $c_{NK}<0$.

13. See D'Ouville and McDonald (1990) and Arnott and Yan (2000). If the toll is only marginally below τ^*, then second-best capacity exceeds first-best capacity (Wheaton, 1978; Wilson, 1983). Correspondingly, if the toll is marginally above τ^*, then second-best capacity is smaller.

14. To the extent that governments view user charges as a cash cow for generating revenue the analysis is also applicable to publicly operated facilities.

15. This is true with homogeneous users. As Spence (1975) shows, with heterogeneous users a monopolist that cannot price discriminate will under-invest in capacity if marginal users value service quality less on average than do inframarginal users (a plausible assumption), and over-invest in capacity if the opposite case holds.

16. Accident costs are borne in significant part by drivers, but a portion of the costs (e.g. to pedestrians and bicyclists) are not covered by insurance. Speed limits and other safety regulations are designed to reduce accident costs but they do not (except for speeding tickets) yield revenues that are directly related to usage. No consensus on the merits of charging users for the external component of accident costs has been reached. One significant complication is that the magnitude of the external cost of an additional trip is unclear.

17. Their long-run point estimates for scale economies are 0.92, 1.45 and 1.96, respectively for the three vehicle types. They attribute scale diseconomies for cars to two factors. First, cars are largely responsible for road capacity requirements since car trips are concentrated during peak periods. Second, infrastructure costs rise as easy options for expansion become exhausted, and as land becomes progressively more expensive to acquire.

18. Nilsson (2002) estimated a cost elasticity of 0.2, which would imply only 20 percent cost recovery under marginal-cost pricing. Pittman (2003) remarks that fixed costs may account for as much as 90 percent of total rail costs.

19. Amongst the reasons he identifies are the impossibility of expanding capacity in built-up areas, and the accounting practices of highway departments, which understate long-run highway capacity costs.

20. The proportion is 10–20 percent for the Norwegian toll rings (Ramjerdi, Minken, & Ostmoe, 2004), about 10 percent for Singapore (Goh, 2002, p. 33; Santos, Li, & Koh, 2004, p. 226) and about 50 percent (not including user compliance costs) in London (Transport for London, 2006, Table 9.1). According to an interim

assessment of the Stockholm congestion charge experiment (Algers et al., 2006), the revenues from congestion charges amounted to 760 million kroner at an annual rate. For maintenance and running costs the corresponding figure was 220 million kroner, which implies a cost/revenue ratio of 29 percent.

21. Scarcity and congestion are two forms of rationing that arise in what Kay (1979) refers to, respectively, as loss systems and delay systems. In a loss system, demand that exceeds capacity is rejected. In a delay system, excess demand is retained until capacity becomes available. Users who are denied immediate service in a loss system can be considered to incur a schedule delay, but not extra travel time or waiting time. By contrast, in a delay system they incur both schedule delay and waiting time costs.

22. Nash, Coulthard, and Matthews (2004) propose a method of setting capacity charges that entails measuring the capacity required by a given train run, and then estimating its opportunity cost in terms of other trains forced off the system. They emphasise that both steps are very complex. Measuring capacity is difficult because it depends not only on the characteristics of the rail line itself, but also on the characteristics of the trains running on it. And estimating the opportunity cost at all accurately can only be done after the timetable has been chosen and the set of potential users is known. Nash et al. (2004) conclude that a system of rail track charges can be calculated only after paths have been allocated.

23. See Blaug (1985) for an insightful review of this literature.

24. Second-order conditions and comparative static properties of the solution are derived in De Palma and Lindsey (2004).

ACKNOWLEDGMENTS

Financial support from the European Commission (EU Fifth Framework research project REVENUE (DG TREN, Contract No. GMA2-2002-52011)) is gratefully acknowledged.

REFERENCES

Algers, S., Brundell Freij, K., Eliasson, J., Henriksson, C., Hultkrantz, L., Ljungberg, C., Nerhagen, L., & Snidfeld Rosqvist, L. (2006). The Stockholm congestion charging trial – what happened? Expert Group Summary. http://www.stockholmsforsoket.se/upload/Rapporter/Expert_group_summary_060621.pdf

Arnott, R. (2006). Pricing, investment, and self-financing of congestible facilities (with application to transportation systems). Lecture presented at third international Kuhmo conference and Nectar cluster 2 meeting, July 13, Helsinki.

Arnott, R., & Kraus, M. (1995). Financing capacity in the bottleneck model. *Journal of Urban Economics*, *38*, 272–290.

Arnott, R., & Kraus, M. (1998a). When are anonymous congestion charges consistent with marginal cost pricing? *Journal of Public Economics*, *67*, 45–64.

Arnott, R., & Kraus, M. (1998b). Self-financing of congestible facilities in a growing economy. In: D. Pines, E. Sadka & I. Zilcha (Eds), *Topics in public economics: Theoretical and applied analysis* (pp. 161–184). Cambridge: Cambridge University Press.

Arnott, R., & Yan, A. (2000). The two-mode problem: Second-best pricing and capacity. *Review of Urban and Regional Development Studies*, *12*(3), 170–199.

Baumol, W., Panzar, J., & Willig, R. (1982). *Contestable markets and the theory of industry structure*. New York: Harcourt Brace Jovanovich.

Berechman, J. (1993). *Public transit economics and deregulation policy*. Amsterdam: North-Holland.

Bichsel, R. (2001). Should road users pay the full cost of road provision? *Journal of Urban Economics*, *50*, 367–383.

Blaug, M. (1985). Marginal cost pricing: No empty box. In: D. Greenaway & G. K. Shaw (Eds), *Public choice, public finance and public policy: Essays in honour of Alan Peacock* (pp. 15–30). Oxford: Blackwell.

Coase, R. H. (1946). The marginal cost controversy. *Economica*, *13*(August), 169–182.

De Palma, A., & Lindsey, R. (2004). *Relation between pricing, toll revenues and investment*. Task Report 2.1, Project MC-ICAM (implementation of marginal cost pricing in transport – integrated conceptual and applied model analysis: http://www.mcicam.net). Funded by the European Commission.

D'Ouville, E. L., & McDonald, J. F. (1990). Optimal road capacity with a suboptimal congestion toll. *Journal of Urban Economics*, *28*, 34–49.

Dupuit, J. (1849). Des Péages. *Annales des Ponts et Chaussées: Mémoires et Documents, 2nd ser.*, *17*(1), 207–248. [Translated by Elizabeth Henderson (1962). On tolls and transport charges. *International Economic Papers*, *11*, 7–31.

Ekelund, R. E., Jr., & Hébert, R. F. (1999). *Secret origins of modern microeconomics: Dupuit and the engineers*. Chicago, IL: University of Chicago Press.

European Conference of Ministers of Transport (1998). *Efficient transport for Europe: Policies for internalisation of external costs*. Paris: OECD Publications.

European Conference of Ministers of Transport (2003). *Reforming transport taxes*. Paris: OECD Publications.

Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). *Megaprojects and risk, an anatomy of ambition*. Cambridge: Cambridge University Press.

Gibson, S., Cooper, G., & Bal, B. (2002). Developments in transport policy: The evolution of capacity charges on the UK rail network. *Journal of Transport Economics and Policy*, *36*(2), 341–354.

Goh, M. (2002). Congestion management and electronic road pricing in Singapore. *Journal of Transport Geography*, *10*, 29–38.

Gómez-Ibáñez, J. A. (1992). The political economy of highway tolls and congestion pricing. *Transportation Quarterly*, *46*, 343–360.

Gómez-Ibáñez, J. A. (1999). Pricing. In: J. A. Gómez-Ibáñez, W. B. Tye & C. Winston (Eds), *Essays in transportation economics and policy: A handbook in honor of John R. Meyer*. Washington, DC: Brookings Institution Press.

Hau, T. D. (1998). Congestion pricing and road investment. In: K. J. Button & E. T. Verhoef (Eds), *Road pricing, traffic congestion and the environment: Issues of efficiency and social feasibility* (pp. 39–78). Cheltenham, UK: Edward Elgar.

Hau, T. D. (2005a). Economic fundamentals of road pricing: A diagrammatic analysis, part I – fundamentals. *Transportmetrica*, *1*(2), 81–115.

Hau, T. D. (2005b). Economic fundamentals of road pricing: A diagrammatic analysis, part II – relaxation of assumptions. *Transportmetrica*, *1*(2), 119–149.

Heggie, I. G., & Fon, V. (1991). *Optimal user charges and cost recovery for roads in developing countries*. The World Bank, Policy Research Working Paper Series: 780.

Jansson, J. O. (1997). Theory and practice of transport infrastructure and public transport pricing. In: G. de Rus & C. Nash (Eds), *Recent developments in transport economics* (pp. 74–134). Aldershot: Ashgate.

Kay, J. A. (1979). Uncertainty, congestion and peak load pricing. *Review of Economic Studies*, *46*, 601–612.

Keeler, T. E., & Small, K. A. (1977). Optimal peak-load pricing, investment, and service levels on urban expressways. *Journal of Political Economy*, *85*, 1–25.

Kraus, M. (1981a). Indivisibilities, economies of scale, and optimal subsidy policy for freeways. *Land Economics*, *57*, 115–121.

Kraus, M. (1981b). Scale economies analysis for urban highway networks. *Journal of Urban Economics*, *9*, 1–22.

Larsen, O. I. (1993). Road investment with road pricing? Investment criteria and the revenue/cost issue. In: A. Talvitie, D. Hensher & M. E. Beesley (Eds), *Privatization and deregulation in passenger transportation, second international conference on privatization and deregulation in passenger transportation* (pp. 273–281). Espoo, Finland: Viatek Ltd.

Levinson, D. M., & Gillen, D. (1998). The full cost of intercity highway transportation. *Transportation Research D*, *3D*(4), 207–223.

Lindsey, R., & Verhoef, E. T. (2000). Congestion modelling. In: D. A. Hensher & K. J. Button (Eds), *Handbook of transport modelling* (pp. 353–373). Oxford: Elsevier.

Mackie, P. (2005). The London congestion charge: A tentative economic appraisal: A comment on the paper by Prud'homme and Bocajero. *Transport Policy*, *12*, 288–290.

Mohring, H. (1972). Optimization and scale economies in urban bus transportation. *American Economic Review*, *62*, 591–604.

Mohring, H., & Harwitz, M. (1962). *Highway benefits: An analytical framework*. Evanston, IL: Northwestern University Press.

Nash, C. (2005). Rail infrastructure charges in Europe. *Journal of Transport Economics and Policy*, *39*(3), 259–278.

Nash, C., Coulthard, S., & Matthews, B. (2004). Rail track charges in Great Britain: The issue of charging for capacity. *Transport Policy*, *22*, 315–327.

Newbery, D. M. G. (1988). Road damage externalities and road user charges. *Econometrica*, *56*(2), 295–316.

Newbery, D. M. G. (1989). Cost recovery from optimally designed roads. *Economica*, *56*, 165–185.

Newbery, D. M. G. (2005). Road user and congestion charges. In: S. Cnossen (Ed.), *Theory and practice of excise taxation smoking, drinking, gambling, polluting, and driving* (pp. 192–229). Oxford: Oxford University Press.

Nilsson, J.-E. (2002). Pricing the use of Sweden railways: Are charges in line with marginal costs? Paper presented at second seminar of IMPRINT-EUROPE, May 14–15.

Pittman, R. (2003). *A note on non-discriminatory access to railroad infrastructure*. US Department of Justice Antitrust Division Economic Analysis Group Working Paper no. EAG03-5. Available at SSRN: http://ssrn.com/abstract=388382;DOI:10.2139/ssrn.388382

Preston, J. (1996). The economics of British rail privatization: An assessment. *Transport Reviews*, *16*(1), 1–21.

Prud'homme, R., & Bocajero, J. P. (2005). The London congestion charge: A tentative economic appraisal. *Transport Policy*, *12*, 279–287.
Quinet, E. (2003). Short term adjustments in rail activity: The limited role of infrastructure charges. *Transport Policy*, *10*(1), 73–79.
Ramjerdi, F., Minken, H., Ostmoe, K. (2004). Norwegian urban tolls. In: G. Santos (Ed.), *Road pricing: Theory and evidence, research in transportation economics* (Vol. 9, pp. 237–249). Amsterdam: Elsevier.
Roy, R. (2005). Optimal pricing: The route to a self-funding infrastructure – and more. *Institute of Economic Affairs*, *25*(1), 5–11.
Santos, G., Li, W. W., Koh, W. T. H. (2004). Transport policies in Singapore. In: G. Santos (Ed.), *Road pricing: Theory and evidence, research in transportation economics* (Vol. 9, pp. 107–131). Amsterdam: Elsevier.
Sikow-Magny, S. (2003). Efficient pricing in transport – overview of European Commission's transport research programme. In: J. Schade & B. Schlag (Eds), *Acceptability of transport pricing strategies* (pp. 13–26). Amsterdam: Elsevier.
Small, K. A. (1992). Urban transportation economics. In: J. Lesourne & H. Sonnenschein (Eds), *Fundamentals of pure and applied economics*. Chur, Switzerland: Harwood Academic Publishers.
Small, K. A., & Verhoef, E. T. (2007). *The Economics of Urban Transportation*. London: Routledge.
Small, K. A., Winston, C., & Evans, C. A. (1989). *Road work*. Washington, DC: Brookings.
Spence, M. (1975). Monopoly, quality, and regulation. *Bell Journal of Economics and Management Science*, *6*, 417–429.
Tisato, P. (1991). User costs in public transport: A cost minimisation approach. *International Journal of Transport Economics*, *18*(2), 167–193.
Tisato, P. (1998). Optimal bus subsidy and cross subsidy with a logit choice model. *Journal of Transport Economics and Policy*, *32*(3), 331–350.
Transport for London (2006). *Central London congestion charging: Impacts monitoring*. Fourth Annual Report. June. http://www.tfl.gov.uk/tfl/cclondon/pdfs/FourthAnnualReportFinal.pdf (accessed October 17, 2006).
UK Department for Transport (2004). *Feasibility study of road pricing in the UK: A report to the secretary of state for transport*. http://www.dft.gov.uk/stellent/groups/dft_roads/documents/divisionhomepage/029709.hcsp (accessed October 17, 2006).
Verhoef, E. T. (2006). *Self financing and transport networks*. Lecture presented at third international Kuhmo conference and nectar cluster 2 meeting. July 11, Helsinki.
Verhoef, E. T., & Rouwendal, J. (2004). Pricing, capacity choice and financing in transportation networks. *Journal of Regional Science*, *44*(3), 405–435.
Wheaton, W. (1978). Price-induced distortions in urban highway investment. *Bell Journal of Economics and Management Science*, *9*, 622–632.
Wilson, J. D. (1983). Optimal road capacity in the presence of unpriced congestion. *Journal of Urban Economics*, *13*, 337–357.
Yang, H., & Meng, Q. (2002). A note on 'highway pricing and capacity choice in a road network under a build-operate-transfer scheme'. *Transportation Research A*, *36A*, 659–663.

APPENDIX: PROOF OF THE COST RECOVERY THEOREM

Given a user cost function $c(N,K)$ and a capacity cost $F(K)$, the total social cost of usage is

$$TC(N,K) = c(N,K)N + F(K)$$

Social surplus, Ω, is measured by the area under the inverse demand curve, $P(\bullet)$, net of total social costs:

$$\Omega = \int_{n=0}^{N} P(n)\mathrm{d}n - TC(N,K)$$

For a given capacity, the short-run optimal level of usage is determined by the first-order condition

$$\frac{\partial\Omega}{\partial N} = P(N) - TC_N(N,K) = 0$$

or

$$p(N) = TC_N(N,K) = c(N,K) + c_N(N,K)N \tag{A1}$$

Given a toll of τ, the private (generalised) cost of usage is

$$p = c(N,K) + \tau$$

To support optimal usage *as per* Eq. (A1) the toll must be

$$\tau = c_N(N,K)N$$

The first-order condition for capacity[24] is

$$\frac{\partial\Omega}{\partial K} = -c_K N - F_k = 0 \tag{A2}$$

If $c(N,K)$ is homogeneous of degree h, then by Euler's theorem

$$c_N N + c_K K = h \times c$$

Multiplying by N:

$$(c_N N)N + c_K KN = h \times cN \tag{A3}$$

Using Eqs. (A2) and (A3):

$$\begin{aligned} R = \tau N = (c_N N)N = -(c_K K)N + h \times cN = F_K K + h \times cN \\ = \varepsilon F(K) + h \times cN \end{aligned}$$

The cost recovery ratio is therefore

$$\rho = \varepsilon + h\frac{c(N,K)N}{F(K)} \text{ QED}$$

This page intentionally left blank

CHAPTER 3

PUBLIC FINANCE ASPECTS OF TRANSPORT CHARGING AND INVESTMENTS

Stef Proost, Bruno De Borger and Pia Koskenoja

ABSTRACT

In this chapter we address three aspects of the pricing and financing of transport infrastructure. First, we analyse the optimal pricing and investment of transport infrastructure. We take a normative approach and propose cost benefit rules to assess investments together with the way they are funded. Second, we survey the specific problems that arise when local governments have responsibility for transport infrastructure and pricing: tax exporting, spillovers, etc. Third, we take a political economy stance and look at issues like lobbying, earmarking and bureaucratic influences.

1. INTRODUCTION

In this chapter, we address three aspects of the pricing and financing of transport infrastructure. Each aspect takes one section. In the first section, we analyse the optimal pricing and investment of transport infrastructure. We take a normative approach and use a general equilibrium context rather than a partial equilibrium context as is used in Chapter 2 in this volume.

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 59–80

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19003-X

This means that the role of government is one of an omniscient planner that maximises a social welfare function. The use of a general equilibrium model (a model for the full economy) allows an analysis of the efficiency trade-off between transport taxes and other taxes to pay for transport investments or for subsidies to particular modes. It also permits examination of the trade-off between efficiency and equity effects of paying investments by user charges or by general tax revenue.

In the next two sections we examine two potential sources of government failure. The first problem is the co-existence of several government levels. When local governments have responsibility for transport infrastructure and pricing, problems arise because each government is only interested in the welfare of the voters of its region. The specific problems discussed are spill-overs in investments, tax exporting, taxation of transport flows by several government levels, etc.

In the third section of this chapter we take a political economy stance. The assumption that governments are benevolent, and maximise the welfare of their voters, is a useful starting point for a normative analysis but it is incomplete. We all know that government behaviour is more complex and is the result of lobbying, vote maximising and bureaucratic behaviour. We analyse four aspects of this problem. First, can we expect marginal social cost pricing from governments? Second, what types of investments are most influenced by lobbying efforts? Third, can institutions such as earmarking of funds be explained by a political economy approach? Finally, does it make sense to decentralise decisions to agencies?

2. FUNDING INVESTMENTS AND OPTIMAL TRANSPORT PRICING BY A BENEVOLENT GOVERNMENT

2.1. Aims

Two reasons are often advanced for allowing deficits in the provision of transport infrastructures: increasing returns to scale (or decreasing average costs; see Chapter 2 of this volume) and equity. Let us first briefly examine the global (or economy wide) efficiency aspects of these deficits. Any deficit for a given transport mode must eventually be covered (with a delay if debt financing is used) by increasing the charges in the rest of the transport sector and/or by increasing general taxes. If prices in the other transport sectors are below marginal social cost, then it might appear that prices could be raised to fund the deficit without a deadweight loss. However, there is an opportunity cost of doing so because the extra revenue could also be used to reduce the general tax level. More particularly, reducing the labour tax level could bring important efficiency gains.

In order to know the equity effects of subsidies in the transport sector it is not sufficient to know that the subsidies go to low-income users. We also need to know who pays for the extra subsidies and this will depend on the type of taxes that are increased or other subsidy programmes that are skipped.

In this section, our aim is to present general cost benefit guidelines that can be applied to the choice of investments and pricing reform in the transport sector. Most existing cost benefit guidelines emphasise on the correct estimation of direct costs and benefits of the transport investments. Here we will emphasise the role of interactions with the rest of the economy and the equity aspects while keeping the transport sector representation as simple as possible.

Infrastructure supply can be seen as a congestible public good. The question of the appropriate supply of public goods in an economy with labour taxes is a classic in the theoretical public finance literature since Pigou (Atkinson & Stern, 1974; King, 1986; Sandmo, 1998). When we disregard equity, the answer looks simple but is not. The simple, but incorrect, answer is that the extra efficiency cost of raising tax money via labour taxes increases the cost of public funds above one and this means a smaller quantity of the public good. This reasoning is wrong in the case of transport investments because the transport investment will have a feedback effect on total tax revenues via the consumption of taxed commodities like transport or via reduced commuting costs. To include equity we also need to consider the income of the beneficiaries of the public good, as well as who pays for the extra supply.

The model used to develop the cost benefit rules is a generalisation of Mayeres and Proost (2001) and Calthrop, De Borger, and Proost (2007). It is a simple static model with passengers and freight transport imbedded in a general equilibrium model and where the capacity of transport infrastructure can be changed for a given rental price. There is no uncertainty and location is fixed. The aim is to derive general algebraic expressions that show the structure of the effects, and to demonstrate what effects can be calculated with transport models and what effects we need to import from more general economic models. The resulting cost benefit expressions are programmed in the MOLINO model (see Chapter 5 in this volume).

2.2. Model Structure

We assume N individuals, $i = 1, \ldots, N$, that differ in their productivity per hour of labour e_i. This productivity is unobservable by the policy maker and is the source of income inequality.[1] Every consumer maximises a utility function defined over a good C (the untaxed numeraire in the model) that requires no freight transport inputs, a consumption good that requires a lot of freight to produce (and called therefore "dirty" good D), a passenger transport good (T) and leisure.

Freight and passenger transport demand jointly produce an externality, interpreted as congestion. Since adding other external costs such as pollution and noise is straightforward, and does not affect the results, we leave them out. The individual faces two constraints: a budget constraint and a time constraint. The congestion function $\phi(\sum_{i=1}^{N} T_i + F, K)$ specifies the travel time incurred per unit of T; this depends on passenger ($\sum_{i=1}^{N} T_i$) and freight (F) transport demand as well as on transport capacity K.[2] For simplicity congestion is assumed to depend only on the sum of passenger and freight transport volumes (measured in commensurable units). We use a very general form of the transport cost function so that our expressions hold for all differentiable functions analysed in Chapter 2 in this volume.

The production structure of the economy is kept as simple as possible. We assume a linear aggregate production function that relates the production of passenger transport, the clean consumption good, freight transport, an intermediate input (X) and capacity and maintenance of infrastructure (K) to a single primary input, labour. Moreover, units are adjusted such that:

$$N(T + C) + F + X + K \leq \sum_{i=1}^{N} e_i L_i$$

We further assume that for a given level of the congestion externality the freight intensive consumption good is produced under constant returns to scale by combining freight transport and the clean intermediate good:

$$\sum_{i=1}^{N} D_i = CRS(F, X; \phi)$$

The higher the level of congestion $\phi(\sum_{i=1}^{N} T_i + F, K)$, the more other inputs are needed to produce one unit of D. We assume perfect competition for the whole production sector. Under these assumptions, the producer prices for T, C, F, X, K and L ($L = \sum_i e_i L_i$) all equal unity. A benevolent government is assumed to maximise welfare W, defined here as the unweighted sum of individual utilities. As all utility functions are identical, it will be the choice of the degree of concavity of the utility function that will determine the marginal social utility of one unit of income for the different individuals. We will denote the marginal utility of income of individual i as λ_i.

The benevolent government has two types of instruments: taxes (uniform lump sum tax G, proportional labour tax τ_T, tax on passenger transport τ_T and on freight transport τ_F) and infrastructure capacity, K. We assume the government is required to maintain a balanced budget every year. The cost of capacity K (here cost is normalised to one) is represented as a rental price:[3]

$$\left(\tau_T \sum_{i=1}^{N} T_i + \tau_L \sum_{i=1}^{N} e_i L_i + [\tau_D + \tau_F F_{ND}] \sum_{i=1}^{N} D_i\right) - K = -NG$$

where $F_{ND}\sum_{i=1}^{N} D_i = F$ represents the total demand for freight services in the economy and where F_{ND} stands for the freight input needed per unit of good D.

We will use this model to derive cost benefit rules for a small increase in infrastructure capacity ($\mathrm{d}K$), that can be funded by increasing different types of taxes and charges: a uniform tax on all households ($\mathrm{d}G$), an increase in the proportional labour tax ($\mathrm{d}\tau$), an increase in the tax on passenger transport ($\mathrm{d}\tau_T$), and an increase in the tax on freight transport $\mathrm{d}\tau_F$.

We use here a marginal tax reform approach[4] starting from a situation where labour taxes are not necessarily optimised.[5]

2.3. Capacity Expansion Financed by an Increase in the Head Tax

We examine first the effect of funding the capacity expansion via a change in the head tax G:

$$\begin{aligned}
\mathrm{d}W = \sum_i & \text{equity weight}_i \,\{\text{passenger transport benefit} \\
& + \text{freight transport benefit}\}\,[\mathrm{NRC}_G] \\
& - \Gamma_G\,\{\text{rental cost capital} - \text{total induced transport tax revenue}_{\mathrm{NRC}_G} \\
& - \text{total induced labour tax revenue}_{\mathrm{NRC}_G}\}
\end{aligned}$$

where the marginal cost of public funds collected by an increase in the head tax is approximated[6] as (where R stands for total government revenue (excluding $\mathrm{d}G$) and where the level of congestion is kept fixed at the level $\bar{\phi}$):

$$\Gamma_G = \frac{1}{1 + \left.\dfrac{\partial R}{\partial G}\right|_{\bar{\phi}}}$$

The NRC stands for the "net reduction in congestion" as a result of the increase in capacity of the infrastructure by one unit. The NRC is computed taking into account three effects. First, the effect of larger capacity for given traffic volume is taken into account. Second, the increased transport volume that is induced by the lower congestion level. Finally, the specific income and substitution effects of raising the head tax on the consumption of transport. Because of the latter effect, we need to index the NRC by the type of tax used to finance the expansion of infrastructure.

The benefits of a transport capacity expansion are a sum of individual benefits that are equity weighted. The individual passenger transport benefit of a capacity expansion will consist of the benefits in time saved (valued at individual values of time), and the individual freight transport benefit will be the decreased production cost of the household's consumption bundle. The equity weight equals the marginal utility of income of the household divided by the average marginal social utility of income. For high degrees of aversion to inequality (here high concavity of the utility functions in the social welfare function), the equity weight of poor income groups can be very high (much higher than the often used inverse of the income after tax). When transport benefits accrue mainly to income groups that have a high equity weight, this increases strongly the benefit term.

The third line in the above equation equals the marginal cost of public funds raised via a head tax (Γ_G) times the net amount of funds that needs to be raised via an increase in the head tax. The net amount to be raised equals the rental cost of capacity expansion minus the transport tax revenues induced by the improved infrastructure quality and minus the induced labour tax revenues. Both changes in tax revenues amounts are net changes. Those are computed over the whole transport network since traffic volumes are interdependent across links.

According to most studies, the marginal cost of funds raised by a head tax is expected to be lower than one because a higher head tax tends to increase the supply of labour and therefore tax revenues. The third line of $\mathrm{d}W$ related to the funding of the capacity expansion is not equity weighted because all households pay the same head tax.

Important additions from general equilibrium models into the cost benefit rule are

– The welfare weights, which can be inferred from revealed preferences of policy makers by examining the structure of the current tax system (see Mayeres & Proost, 1997, for an application to transport).
– The induced labour tax revenues: this may be important when the transport project affects mainly commuting traffic; if labour supply is a function of the net wage after deduction of commuting costs, the elasticity of labour supply can be used to estimate this term.
– The marginal cost of public funds of a head tax.

2.4. Capacity Expansion Financed by an Increase in the Proportional Labour Tax

The net welfare effect of a small increase in the transport capacity now equals:

$$\begin{aligned}\mathrm{d}W = \sum_i & \text{equity weight}_i \, \{\text{passenger transport benefit} \\ & + \text{freight transport benefit}\} \, [\mathrm{NRC}_{\tau_L}] \\ & - \Gamma_{\tau_L}\left(\frac{\sum_i^N \frac{\lambda_i}{\lambda^*} e_i L_i}{\sum_i^N e_i L_i}\right)[\text{rental cost capital} \\ & - \text{total induced transport tax revenue}_{\mathrm{NRC}_{\tau_L}} \\ & - \text{total induced labour tax revenue}_{\mathrm{NRC}_{\tau_L}}]\end{aligned}$$

The marginal cost of public funds collected by an increase in the proportional labour tax is approximated as:

$$\Gamma_{\tau_L} = \frac{1}{1 + \dfrac{\tau_L}{\sum_i e_i L_i} \dfrac{\partial\left(\sum_i e_i L_i\right)}{\partial \tau_L}\Bigg|_{\bar{\phi}}}$$

The benefits of a transport capacity expansion per unit of reduced congestion are identical to the case with a head tax. The NRC may be different because of other income and substitution effects.

The third line related to the funding of the capacity expansion is now equity weighted because it matters who will pay the net increase in labour taxes that is needed. The marginal cost of public funds collected via a proportional labour tax depends on the weighted sum of labour supply elasticities. One expects this sum to be negative so that Γ_{τ_L} is larger than one.

There exist many definitions and estimates of the marginal cost of public funds. The most common definition is the efficiency cost of raising one unit of tax revenue, given that the tax revenue is spent on a public good that does not affect the consumption of taxed commodities. If we can disregard secondary effects of the labour or head tax on congestion (keeping capacity constant) we can use existing estimates in the literature. This is the way we define our marginal cost of public funds. Kleven and Kreiner (2003) provide order-of-magnitude estimates of the marginal cost of public funds that account for effects on labour force participation decision and hours worked. They also account for the replacement payments for unemployed, count social security contributions on labour as taxes, and the average indirect tax level as a proportional labour tax. They provide estimated values of the marginal costs of funds for different countries using a standarised set of labour supply elasticities. We report here two extremes: Belgium (high labour taxes) and the UK (rather low labour taxes) for a regressive tax (head tax) and for a proportional labour tax.

The 2.52 value means that every euro that needs to be raised by labour taxes to balance the budget after a transport investment has an efficiency cost of 2.52 euros. Table 3.1 illustrates clearly that the way the capacity is financed plays an important role in the cost benefit assessment. The marginal costs of public funds, as defined in the previous equations, are available for all OECD countries and can be used in the MOLINO model (see Chapter 5 in this volume) and in the case studies reported in this volume.

Table 3.1. Examples of Marginal Costs of Public Funds Values.

	Belgium	UK
Regressive tax (head tax)	1.11	1.09
Proportional labour tax	2.52	1.37

Source: Kleven and Kreiner (2003).

2.5. Capacity Expansion Financed by an Increase in the Passenger Transport Taxes or Freight Transport Taxes

The net welfare effect of a small increase in the transport capacity now equals:

$$\begin{aligned}
\mathrm{d}W = \sum_i & \text{equity weight}_i \,\{\text{passenger transport benefit} \\
& + \text{freight transport benefit}\}\,[\mathrm{NRC}_{\tau_T}] \\
& - \Gamma_{\tau_T}\left(\frac{\sum_i^N \frac{\lambda_i}{\lambda^*} e_i T_i}{\sum_i^N e_i T_i}\right)[\text{rental cost capital} \\
& - \text{total induced transport tax revenue}_{\mathrm{NRC}_{\tau_T}} \\
& - \text{total induced labour tax revenue}_{\mathrm{NRC}_{\tau_T}}]
\end{aligned}$$

where the marginal cost of public funds collected by an increase in the passenger tax is approximated as:

$$\Gamma_{\tau_T} = \frac{1}{1 + \frac{\tau_T}{\sum_i T_i} \frac{\partial\left(\sum_i T_i\right)}{\partial \tau_T}}$$

The benefits of a transport capacity expansion are identical to the previous cases except that the NRC is now expected to be much larger for the same capacity expansion because the increase in passenger transport tax reduces congestion too.

The third line related to the funding of the capacity expansion is again equity weighted because it matters who will pay the increased passenger taxes. The marginal cost of public funds collected via a passenger tax is a function of the direct effect of the passenger tax plus the feedback effect of the passenger tax on congestion and so on total number of trips.[7] This marginal cost can be computed via a transport market model.

One can rewrite the cost–benefit expression and reformulate it in terms of deviations between the marginal external cost on the different transport markets and the unit tax on the different markets. This alternative formulation makes clear that transport taxes closer to the marginal external cost will in general improve efficiency and welfare. There are, however, two additional elements that need to be taken into consideration and that can justify a deviation from pure marginal social cost pricing. A first deviation is equity and this can point to higher or lower prices depending on the welfare weight of the most frequent users. The second is the need to raise public revenues and this can justify setting higher prices for modes with low-price elasticity and setting lower prices for modes that are complements to highly taxed activities. An example of the latter is a commuting train service whose price affects labour supply.

We can use the same type of cost–benefit analysis for the welfare effect of investments financed by an increase in the freight transport taxes. The major difference will be that the benefits come predominantly via lower consumption prices for freight intensive goods and that this benefit may have a different equity impact. Also the marginal cost of public funds will be different as price elasticities of freight transport may be different.

2.6. Investment Rules in a Growing Economy

We have so far assumed a static economy that repeats itself indefinitely. Capital taxes and uncertainty are absent and all discount rate issues are implicit in the rental cost of capacity. As long as we disregard risk and uncertainty we can extend the previous investment assessment rules to a growing economy. Following Liu (2003), we can define the investment rule for an economy with a capital tax and a proportional labour tax and identical individuals that have an infinite lifetime. An infrastructure investment project can now be defined as a stream of investment and maintenance costs $\Delta I_t,\ t = 1, \ldots, \infty$; a stream of benefits for the households $\Delta B_t,\ t = 1, \ldots, \infty$; and a stream of induced tax receipts ΔR_t, $t = 1, \ldots, \infty$. The investment rule now becomes:

$$\mathrm{d}W = \sum_t \frac{\Delta B}{(1+r_n)^t} - MCF_{\tau_L} \sum_t \frac{\Delta I_t - \Delta R_t}{(1+r_g)^t}$$

where $r_n = r_g(1-\tau_K)$, and

$$MCF_{\tau_L} = \frac{\sum \dfrac{w_t L_t}{(1+r_n)}}{\partial\left[\sum \dfrac{R_t}{(1+r_g)^t}\right] \Big/ \partial \tau_L}$$

where τ_K is the tax on capital income in the economy.

We see that the benefits are discounted at the consumers' rate of discount net of taxes, r_n, but the government uses a gross interest rate, r_g, that is equal to the gross return of capital in the economy. This procedure is more rigorous than the use of a weighted average of the net and gross interest rates of the social discount rate.

Extending the framework to an economy with risk and uncertainty is complicated as risk and uncertainty are difficult to include in a general equilibrium model. For a cost benefit analysis we recommend to use the market rate of interest for private suppliers of capital. These contain a risk premium. The magnitude of the risk premium will be a function of the risk class of the transport investment (Chapter 4 of this volume surveys these estimates).

2.7. Summing up the Investment Rules

We have shown that a cost benefit assessment of capacity expansion needs to take into account the type of taxes used to finance the capacity expansion. The type of funding matters for both the equity and the efficiency effects. The cost benefit assessment can be operationalised using a combination of transport models and information from general economic models. We can apply the same framework to assess any budget preserving change in taxes. The application of these cost benefit rules is modelled in the MOLINO model (see Chapter 5 of this volume).

3. TAX AND INVESTMENT RULES IN AN ECONOMY WITH SEVERAL GOVERNMENT LEVELS

3.1. A Taxonomy of Problems

How can efficient pricing and investment by local and regional governments be assured? What are the prescriptions to handle spillovers in benefits (cross-border traffic and transit traffic), and what risks do we run with horizontal and vertical tax competition issues? Is federal matching of regional investment needed, and how should the matching schemes be designed?

We distinguish three different cases. The first distinction is between horizontal tax and vertical tax competition (see Fig. 3.1). In vertical tax competition two different government levels (regional and urban, or federal and regional) set taxes on the use of the same infrastructure.[8] In horizontal competition, two local governments are competing for the same tax

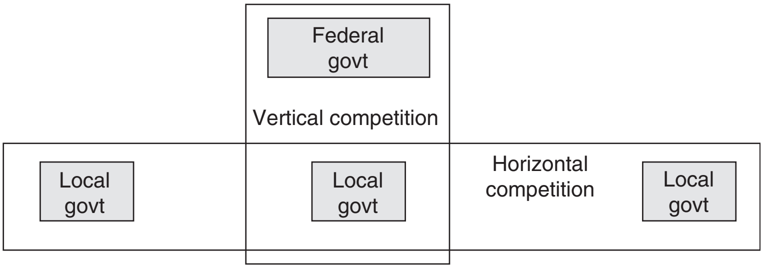

Fig. 3.1. Vertical and Horizontal Tax Competition.

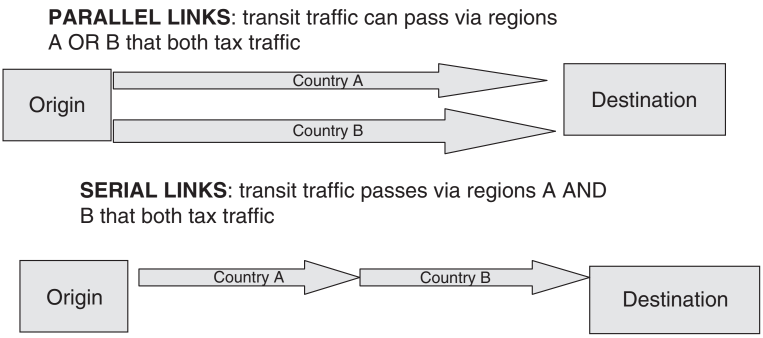

Fig. 3.2. Two Types of Horizontal Tax Competition.

revenues. Within the horizontal tax competition category, we need to distinguish between two more cases (see Fig. 3.2). The first setting is where two countries are non-cooperatively competing for the same transit traffic ("parallel tax competition"): say Switzerland and Austria competing for German trucks going to Italy. The second setting is where transit traffic has to go through two or more countries before it reaches its destination ("serial competition").

3.2. What Pricing Problems to Expect

In the parallel horizontal competition model we know from numerical illustrations that, if infrastructure use is priced, this pricing guarantees a more or less efficient use of existing infrastructure even if the tolls charged are expected to be higher than the marginal external cost. This is due to the Bertrand type of competition for transit that drives taxes down. When we allow for investments by the two competing governments it is expected that investment levels are suboptimal too.[9]

So what matters is to have charging for infrastructure use in place. It is only when infrastructure use cannot be priced that there will be overuse of the infrastructure and underinvestment. The underinvestment problem stems from two reasons: the need for capacity expansion will probably be higher because prices are too low, and second there will be insufficient incentives to invest since the benefits to transit are not taken into account by local governments.

So when the infrastructure cannot be tolled, and when transit traffic accounts for a large proportion of the total use, federal funding may be needed. Federal funding may also be called for when the local tax base is small and the marginal cost of funds for the local government is higher than that for the federal government. This will depend on the tax structure and revenue sharing mechanisms in place (Hoyt, 2001).

In the second case (serial horizontal tax competition), we need to distinguish again the case where traffic can be charged and where it cannot. When two local governments can tax transit traffic that passes through the two countries, this will result in charges whose sum is even higher than regular monopoly charges. In fact, each government, when setting its charges, disregards the loss of profits for the other government.[10] However, investment levels by local governments will in this case not fall too short of the optimum for the (suboptimal) level of traffic.[11]

Again if traffic cannot be charged at all, there will be underinvestment for the very same reasons as in the parallel competition case. In this case, federal funding of the matching type that is in line with the share of transit traffic benefits can be an important instrument to achieve a better capacity level. A high local marginal cost of funds may be a justification for an even higher matching of investment funds.

In the third case (vertical tax competition), we have several government levels that tax or subsidise the same car or rail traffic. The result may again be overcharging and probably a reduction of funding options for capacity expansions for the local governments. Take as example a federal government that sets a high fuel tax for the whole country. In this case it is suboptimal to rely on local governments to fund their road infrastructure via extra tolls as the federal government has already creamed off the market. The high fuel taxes on car use also bias the investment criteria of local public transport. One of the costs of local public transport expansion will be the loss of fuel tax revenues at the federal level but this will not be considered as a cost by the local government as fuel tax revenues are usually not returned to the local government. This is an example where local governments may actually over invest rather than under invest in local public transport. There exist several types of solutions to this problem. Revenue sharing schemes can be one option.

An obvious question is why we would ever decentralise transport decisions to local governments as this generates inefficient behaviour. Decentralisation can be seen as the result of a historical political process, but in the background there may be two important economic forces at work. The first is that better information may be available at the local level. This means that we have a procurement problem where a central government tries to control the behaviour of a local government via pricing regulation or via matching subsidies. This regulation problem between a principal (central government) and a better-informed agent is studied in more detail in Chapter 4 in this volume. The second important force is that transport infrastructure that mainly serves local traffic may be an important source of lobbying at the federal level if the federal level pays for the infrastructure. The problems raised by lobbying are discussed in more detail in the following section.

4. POLITICAL ECONOMY OF PRICING AND INVESTMENT RULES

4.1. From Normative to Positive Economics

The optimal pricing, investment and revenue use rules, discussed in the previous sections, are derived from normative economics. This means that a policy maker maximises some weighted sum of individual utility functions and derives optimal policies under all types of constraints on instruments and incentives. This can serve as a benchmark but remains a naïve policy prescription. We need a richer model that accounts for the way in which public decisions are made.

In this richer model we need voters, agencies, administrations and lobby groups. As we are dealing with a complex game between many groups where the rules of the game differ across countries, insights are not very clear-cut. In our survey of the theory we focus on three different political economy models and look for insights on transport pricing and investment. Political economy is a positive approach: one tries to understand the outcome of the policy process as a function of its structure. The lessons one can derive are therefore lessons about the likely outcomes of particular institutions.

We start by discussing the traditional static common-agency lobbying model. Next we analyse a dynamic model to explain earmarking. We end by discussing the role of agencies in political economy. Each of the three models illustrates one dimension of the public decision process, and it should be kept in mind that these dimensions will often be present simultaneously.

4.2. The Common Agency Lobbying Model

In the common agency model (see Dixit, Grossman, & Helpman, 1997; Grossman & Helpman, 1996) there is a policy maker who is influenced by the voting process and by a lobbying process. The voting process is kept unspecified and results in policies that maximise a weighted sum of voters' utility functions. Lobby groups propose to the government a menu of (truthful) lobbying contributions. The lobbying contributions proposed to the policy maker are a function of the policy proposed by the policy maker: the better the proposed policy matches the preferences of the lobby group, the higher the lobbying contributions. Different groups compete to influence the policy maker.

The formulation of the lobbying game is simplistic in the sense that neither the lobby formation process, nor the political and bureaucratic processes are described. The main advantage of this common agency approach is that an explicit solution can be derived for the lobbying game. It can be shown that the equilibrium to this lobbying game is a policy that maximises a weighted sum of voters' utilities plus a weighted sum of the utility of the lobby groups. The weights of the lobby groups reflect their respective strengths.

We can illustrate the model by discussing two simple transport issues. The first is the introduction of a toll on a motorway. There are only two groups in this society with sizes n_1 and n_2. Their numbers of trips on the motorway as a function of the toll are x_1 and x_2 and their shares in the redistribution of the toll revenues are s_1 and s_2. Toll revenues are redistributed by reducing other taxes, and the shares s_1 and s_2 are taken as exogenous. Imagine that group 2 is a very effective lobby (say truckers) and group 1 has no lobby power at all. Assume that lobby group only can influence the level of the Pigouvian toll but not the shares s_1 and s_2. The level of the toll, τ, preferred by each lobby group will be a function of its relative use of the motorway and its share in the redistribution of the toll revenues. Table 3.2 spells out a few polar cases where the preferences of the second group for the toll level can be determined unambiguously. If the use of the motorway is identical

Table 3.2. The Equilibrium Toll in Relation to the Pigouvian Toll as a Function of the Relative Use and Revenue Share of Lobby Group 2.

	$s_2 = 1/n_2$	$s_2 = 1/(n_1+n_2)$	$s_2 = 0$
$x_2 > 0$, $x_1 = 0$		$\tau <$ Pigouvian	
$x_2 = x_1$	$\tau >$ Pigouvian	$\tau =$ Pigouvian	$\tau <$ Pigouvian
$x_2 = 0$, $x_1 > 0$		$\tau >$ Pigouvian	

for both groups then only the share in the road toll revenues matters for the toll preference. The lobby group 2 will favour a toll beyond the Pigouvian level when it does not use the motorway but uses shares in the revenues, etc. From Table 3.2 it appears very clearly that the redistribution of toll revenues and the relative use are the key parameters in the lobbying behaviour; and when this lobby group is powerful, the resulting tolls may be higher than, equal to, or lower than the Pigouvian taxes.

The second example is the decision whether to build a specific road or public transport investment that cannot be tolled. This can be a bridge or a metro to a remote location. This type of investment is called a specific public good.[12] As Persson (1998) and others have pointed out, the supply of specific public goods paid by a general tax is a very common way to favour a lobby group. The problem is that the benefits are for a very small group, while the costs are for a large group of taxpayers. This implies that lobbying may have a particularly high payoff when the use of the infrastructure is not priced and the benefits are only shared by a small group (region or one specific mode).

These two simple examples show clearly where lobbying activities may be most rewarding. Of course, the lobby group activities can be modelled in much greater detail taking into account important factors such as their control of certain information media.

4.3. Earmarking in a Dynamic Political Model

There are many examples of earmarking in the field of transport pricing and investment. Earmarking of tax revenues for a particular purpose can be motivated for several reasons:

- To gain a political majority in a heterogeneous population by building in compensation for losers. This motive can be studied using the static common agency model presented above.
- To limit spending by bad politicians. Brett and Keen (2000) develop a model that will be discussed later in this section.

– To increase the utility of a favoured lobby group in periods beyond the political incumbency by allocating resources to long-lived infrastructure (Glazer, 1989). This idea bears some resemblance to the specific transport investment example discussed above.
– Buchanan's idea of using a separate tax for each type of expenditure to allow voters to decide over each type of public good.
– Bös' (2000) model of competing expenditure departments and a taxing department.

Earmarking is in general not favoured by finance ministers because it constrains general tax and expenditures policy. Normative public finance theory tends therefore to discourage earmarking.

A particularly relevant model is the one of Brett and Keen (2000) that we present in some more detail here. They use a model with identical individuals. There are two types of politicians: the good ones and the bad ones. The good ones choose an election platform that maximises the utility of the citizens. The bad ones maximise tax revenues and then waste them in projects they like.

In terms of transport policy we could translate the model setting into a very simple road-pricing problem with fixed road capacity where the level of the externality (congestion) needs to be contained by using congestion charges. The congestion charges can be transferred back to the citizens either by lowering other taxes or through increased transfers. This would be the policy of the good politician. The bad politician is interested in using the revenues from the congestion charges for other projects that are of no value to the citizens.

The model contains two periods. In the first period, the incumbent proposes a tax policy for the next period. The voters do not know the type of the incumbent politician or the type of the challenger, but they do know the probability that the incumbent and challenger are good politicians. The incumbent and the challenger both have to present election platforms in the first period. Voters try to infer the type of politician from the proposed policy platforms and vote at the end of the first period for a politician. In the second period, the elected politician reveals his/her type. A good politician executes his/her promises, and a bad politician wastes the money.

In addition to choosing the externality tax, in the first period the politicians can also commit to earmark the congestion tax revenues to a useful cause. Earmarking will preclude the possibility that a bad politician uses the revenues in a bad way. But earmarking creates an efficiency loss for each euro that is earmarked. This loss could be the result of inflexibility built in by the earmarking, or it could result from the lack of incentives for the receiving agencies to use the expenditure budgets optimally. If there were no inefficiency loss from earmarking, a political programme that consists of the optimal congestion tax combined with the earmarking commitment would always win the election.

It is interesting to analyse first the tax policies that would result if earmarking were impossible or not credible, and the externality damage is known. There are a few immediate insights that follow from this formulation. First, when there is a risk that bad politicians win the elections, voters favour tax rates that are lower than the Pigouvian rate because less revenue is potentially wasted. Second, no politician will ever propose a tax rate higher than the Pigouvian tax, since by doing so he would reveal his (bad) type immediately.

What does the option of earmarking add to this game? First, a sufficient condition for a good politician to propose earmarking is that the inherent efficiency loss of earmarking is relatively low compared to the risk of bad politicians taking over.

Second, when the good politician chooses to earmark, the tax rate proposed will be higher than the tax rate proposed in the absence of earmarking but still lower than the Pigouvian tax. The proposed tax will be higher with earmarking because the risk of wasting the revenue has disappeared. On the other hand, the tax rate proposed under earmarking will be lower than the Pigouvian tax because every euro collected has an efficiency loss.

Brett and Keen have some more results when the true level of the externality is known by the politician, but will only become known to the voters in the second period. One could imagine that the voters lack good traffic forecasts and do not know the future level of congestion. This raises interesting questions because we know that the bad politician has an interest to overstate the probability of large externality damages as he can justify higher taxes and more revenues. In this case, there may be equilibria where the good politician proposes earmarking when the optimal tax rate is high and chooses not to earmark when the optimal tax is low, while the bad politician always chooses to set a low tax and not to earmark.

Brett and Keen's framework helps us explain why earmarking exists despite its inherent inefficiency. Good politicians propose taxes below the Pigouvian level for two types of reasons. First, they want to minimise the losses in case a bad politician wins the election; and second, in the case of earmarking, the tax has to be lower because the revenues of the tax are suboptimally used. Earmarking is more likely in their model the lower the share of good politicians and the lower is the efficiency loss of earmarking.

When the extent of the expected externality problems is uncertain, earmarking can again help to signal the good politicians.

This model is simple and explains why good politicians can favour earmarking. It is obviously incapable of explaining the precise use of the earmarked funds for transport as any good use is acceptable in this model. The earmarking of funds for specific transport purposes has therefore to be explained by other motives. The model presented here gives insights, but needs empirical validation.

4.4. The Role of Agencies and Politicians

Most political economy models do not represent the administration and agencies as separate players in the game. Yet, it is clear that these parties do influence the decisions because their information advantage helps them restrict the potential choices of politicians. No simple theory of the internal organisation of government has yet been developed (Tirole, 1994) but there is an increasing interest in a positive theory of why politicians take on certain tasks and delegate other tasks to agencies.

Alesina and Tabellini (2004) have studied this question in a general setting. They first look for an ideal distribution of tasks that is based on the characteristics of politicians and administrations. They find that politicians take best responsibility for tasks where:

- Differences in performance are due to effort rather than to technical ability.
- The preferences of the public are unstable and uncertain so that flexibility is desirable.
- Time inconsistency is unlikely to be a relevant issue.
- Politicians cannot favour short-term objectives over long-term objectives.
- The stakes for organised lobby groups are small.
- Bundling of different policies is important for efficiency reasons or to compensate losers of a particular policy.

When the allocation of responsibilities is left to the politicians, a different delegation of tasks may result. Politicians:

- want to keep tasks that could generate large rents and large campaign contributions;
- prefer to delegate more "risky" tasks to agencies as they can then blame the administration; and
- refuse to ever delegate redistributive policies as these allow politicians to form winning coalitions.

Delegation of tasks to investment agencies because of their technical ability to assess projects is therefore not easily accepted by politicians. Chapter 4 in this volume discusses the potential advantages of transport investment agencies in more detail.

5. CONCLUSIONS

This chapter has surveyed three public finance aspects of the use of transport revenues. The first one is the dependence of the optimal transport investments on the way the investment is financed. There are many costs of public funds. The cost of public funds depends on the type of tax used to finance the investment (head tax, labour tax, transport tax or charge) and it should not be forgotten that the transport investment project itself may have important side effects on public revenues that need to be taken into account. This chapter has suggested cost benefit rules that contain these effects.

The second public finance issue discussed is the presence of different layers of government. It was shown how the non-cooperative behaviour of local governments can lead to inefficient pricing and investment of transport in serial and in parallel networks. When we add a federal government level that also taxes the same traffic, things may become even worse.

The chapter finished with a political economy approach and looked for the real determinants of low transport taxes and peculiar practices like earmarking of transport tax revenues. There does, as yet, not exist a satisfactory explanatory theory for these phenomena.

NOTES

1. This is the standard way of introducing income inequality (see Mirrlees, 1971).
2. If the passenger transport good corresponds to road travel then K denotes road capacity and congestion takes the form of travel delay. Alternatively, if the passenger transport good is a public transport mode then K denotes the capacity of the relevant public transport link and congestion may take the form of delays (e.g. flight delays with air travel) or a decline in service quality (e.g. crowding on commuter trains).
3. We present this constraint without the option of borrowing. We can include borrowing but the government budget holds for all periods so this implies higher taxes later. In this model, we represent a stationary economy so that the capacity cost is a rental price that includes investment, maintenance and interest charges.
4. A marginal tax reform approach only holds for small changes in taxes and investments. The main analytical and practical advantage of a marginal tax reform is that much less information is needed to determine the sign and magnitude of ΔW. As the transport sector is small compared to the whole economy, tax changes will be small.

5. Even if all existing taxes are optimised, the equity weights will not all be equalized because redistributing income via distortionary taxes is costly. The same model can also be used to derive so-called Pareto-improving tax reforms that generate net utility gains for all individuals. With many different individuals and no lump sum instruments, such reform may be impossible to design (Guesnerie, 1977).

6. This is an approximation as we have neglected the direct effect of the change in the head tax on total transport, congestion and revenues for the initial capacity level. This term is probably very small, and neglecting this term allows us to use estimates of the marginal cost of public funds in the public economics literature. The same observation applies to the funding via a labour tax considered below. For the funding of investments via transport taxes, we cannot and do not neglect this term.

7. It is written as a total derivative but capacity is kept constant in this expression.

8. More detailed argumentation and models can be found in De Borger and Proost (2004) and in De Borger, Proost, and Van Dender (2005).

9. See De Borger, Dunkerley, and Proost (2006).

10. This problem is known in the Industrial Organization literature as the double marginalisation problem (see Tirole, 1988).

11. See discussion of investment behaviour by a monopolist in Chapter 2 in this volume and De Borger, Dunkerley, and Proost (2007).

12. In the case of transport infrastructure that is built with a large capacity but has almost no users, the marginal cost per user is very low as there is no congestion and this can therefore be considered as a public good that is non-rivalrous in nature.

ACKNOWLEDGMENTS

Financial support from the European Commission (EU Fifth Framework research project REVENUE (DG TREN, Contract No. GMA2-2002-52011)) is gratefully acknowledged. We are grateful to R. Lindsey and E. Quinet for useful comments on a previous draft.

REFERENCES

Alesina, A., & Tabellini, G. (2004). *Bureaucrats or politicians?* NBER Working Paper no. 10241, January.

Atkinson, A. B., & Stern, N. H. (1974). Pigou, taxation and public goods. *Review of Economic Studies*, *41*, 191–228.

Bös, D. (2000). Earmarked taxation: Welfare versus political support. *Journal of Public Economics*, *75*(3), 439–462.

Brett, C., & Keen, M. (2000). Political uncertainty and the earmarking of environmental taxes. *Journal of Public Economics*, *75*, 315–340.

Calthrop, E., De Borger, B., & Proost, S. (2007). Externalities and partial tax reform: Does it make sense to tax freight (but not passengers) transport. *Journal of Regional Science*, Forthcoming, Aug 2007.

De Borger, B., Dunkerley, F., & Proost, S. (2006). *The interaction between tolls and capacity investment in serial and parallel transport networks.* Discussion paper 06, 7, Center for Economic Studies. www.econ.kuleuven.be/ces/
De Borger, B., Dunkerley, F., & Proost, S. (2007). Strategic investment and pricing in a congested transport corridor. *Journal of Urban Economics*, Forthcoming.
De Borger, B., & Proost, S. (2004). Vertical and horizontal tax competition in the transport sector. *Reflets et Perspectives de la Vie Economique*, *XLIII*(4), 45–64.
De Borger, B., Proost, S., & Van Dender, K. (2005). Congestion and tax competition in a parallel network. *European Economic Review*, (Nov), *49*, 2013–2040.
Dixit, A., Grossman, G. M., & Helpman, E. (1997). Common agency and coordination: General theory and application to government policy making. *Journal of Political Economy*, *105*(4), 752–769.
Glazer, A. (1989). Politics and the choice of durability. *American Economic Review*, *79*(5), 1207–1213.
Grossman, G., & Helpman, E. (1996). Electoral competition and special interest politics. *Review of Economic Studies*, *63*, 265–286.
Guesnerie, R. (1977). On the direction of tax reform. *Journal of Public Economics*, *7*, 179–202.
Hoyt, W. H. (2001). Tax policy coordination, vertical externalities, and optimal taxation in a system of hierarchical governments. *Journal of Urban Economics*, *50*, 491–516.
King, M. (1986). A Pigouvian rule for the optimal provision of public goods. *Journal of Public Economics*, *30*, 273–291.
Kleven, H. J., & Kreiner, C. T. (2003). *The marginal cost of public funds in OECD countries: Hours of work versus labor force participation*. CESifo Working Paper no. 935.
Liu, L. (2003). A marginal cost of funds approach to multi-period public project evaluation: Implications for the social discount rate. *Journal of Public Economics*, *87*, 1707–1718.
Mayeres, I., & Proost, S. (1997). Optimal tax and investment rules for congestion type of externalities. *Scandinavian Journal of Economics*, *99*(2), 261–279.
Mayeres, I., & Proost, S. (2001). Tax reform for congestion type of externalities. *Journal of Public Economics*, *79*, 343–363.
Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, *38*(114), 175–208.
Persson, T. (1998). Economic policy and special interest politics. *Economic Journal*, *108*, 310–327.
Sandmo, A. (1998). Redistribution and the marginal cost of public funds. *Journal of Public Economics*, *70*, 365–382.
Tirole, J. (1988). *The theory of industrial organization*. Cambridge, MA: MIT Press.
Tirole, J. (1994). The internal organization of government. *Oxford Economic Papers*, *46*, 1–29.

CHAPTER 4

THE CONTRACTING OF INVESTMENT AND OPERATION, AND THE MANAGEMENT OF INFRASTRUCTURE FUNDING BODIES

David Meunier and Emile Quinet

ABSTRACT

Public authorities can make a wide variety of institutional arrangements for the provision and operation of infrastructure. What is the best organisation or, to put it more pragmatically, what are the advantages and drawbacks of the various options? This is the question that the present chapter will address, building on the lessons of principal-agent theory and on the analysis of various situations pertinent to the European context. The creation of an infrastructure agency is discussed, and then various modes of transport are reviewed: motorways, railway tracks and seaports. Some numerical values are derived for parameters used in cost-benefit assessments of infrastructure projects.

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 81–109

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19004-1

1. INTRODUCTION

Public authorities can make a wide variety of institutional arrangements for the provision and operation of public utilities. In the case of electricity, for instance, power is supplied either by national agencies or by private firms under various types of regulation. Telecommunications are operated by regulated firms, private or public, in some cases under limited franchises. There are even more varied types of provision for transport infrastructures, from direct command and control by the public service to free market structures, operating through various kinds of franchises.

These alternatives are encountered in many situations, from the national or international scale to the city scale. In general, the choices actually made are conditioned by history, unforeseen circumstances, and the shape of the political and administrative organisation of each country. Each alternative has advantages and disadvantages related to the peculiarities of the situation such as the characteristics of the mode, the structure of information flows between the participants and the structure of the downstream market.

What is the best organisational form or, to pose the question more pragmatically, what are the advantages and drawbacks of the various options? In what specific situations should each of them be used? These are the questions that the present chapter will address, building on the lessons of principal-agent theory[1] (Laffont & Tirole, 1993), and on the experience in Europe (Vickers & Yarrow, 1988; Gómez-Ibáñez & Meyer, 1993; Estache, Guasch, & Trujillo, 2003).

The ideal way to achieve this goal would be to establish a complete typology of all possible organisations, to assess the virtues and drawbacks of each of them, and to conclude with recommendations on when and how to use them according to the specific situations. It would be a huge task to produce such an exhaustive review. The scope of this text is more limited, both in terms of the possible organisations and in terms of the modes, which are analysed. The choices are inspired by the most frequent issues and situations that arise in organisation for economic co-operation and development (OECD) and European countries.

Among the possible organisations, the present analysis will focus, for each of two types of tasks, on two contrasting choices to be made.

Management of investment: Should the public authority do the job itself (as is the case in many countries for roads and motorways) or should it delegate part of its powers to an infrastructure agency (a solution often found for air traffic control)?

Management of operations: Should either the public authority or the delegated infrastructure agency manage the operation of a given infrastructure by itself (as is the case for German motorways), or should it franchise operations to a public or private firm (as it happens in France and in the UK, through various types of franchise)? This question will be answered by the analysis of three types of infrastructures of particular importance in Europe: motorways, rail tracks and seaports. European transport has some particular features: many infrastructures are cross border, or are used by transit traffic, and some kind of European regulation is superimposed on national regulations. In order to provide insight into the problems that arise in such situations, some attention is devoted to the organisational arrangements and regulation in an international framework.

The chapter is organised as follows. Section 2 analyses the arguments for and against a transport infrastructure agency. Section 3 summarises the main arguments in choosing between public and private management as exemplified by public–private partnerships. Section 4 is devoted to the examples of motorways, rail tracks and seaports. Section 5 examines the case frequently encountered in Europe where international projects involve several regulators or several principals acting together. Section 6 presents a rough typology of empirical studies that have derived estimates for the values of some key parameters that play a major role in the previous sections. Section 7 concludes with a summary of the main results.

2. THE CASE FOR A TRANSPORT INFRASTRUCTURE AGENCY

This section addresses the question whether a public authority should delegate part of its powers to a transport infrastructure agency (TIA). We will first outline what a TIA consists of, and then discuss some considerations about its scope, its resources and its powers. Finally, we will consider the efficiency implications of a TIA, and its acceptability to the main stakeholders.

2.1. Mandate of a Transport Infrastructure Agency

Theoretically, a public authority could create a TIA responsible for some or all of the stages in developing and operating transport infrastructure: medium-term and long-term planning, financing, construction of infrastructure, management and maintenance. In practice, direct public involvement can hardly be avoided in the planning and programming stage, due to both legal and social constraints and because of the usual market failures. Indeed, building new transport infrastructure typically requires expropriation of private property through the exercise of powers of eminent domain. The infrastructure's "public interest" needs to be assessed, and considerations beyond market concerns have to be taken into account; for example, the effect on regional development or the environment. The political impact of new infrastructure and the intense media coverage that it draws are such that public authorities adhere firmly to their power of decision in this regard. Usually, decisions about what type of infrastructure should be built, and when, are taken by the public authorities alone.[2]

From a theoretical point of view, really independent planning might lead to decisions nearer to a collective optimum than decisions made through the political process. Nevertheless, though a TIA may be deprived of such a planning function, it could play an indirect role within the process of decision. Indeed, using terms taken from incentives theory, a TIA may help reduce information asymmetries among actors by generating and disseminating useful new information. For instance, such information may relate to the financial or external costs of an infrastructure project, to its distributional impact, to the demand function or to the nature and magnitude of the uncertainties linked to the project.

Henceforth, we will consider a TIA that is given some official direct role in financing, and possibly the implementation, management and maintenance of the infrastructures within a certain scope. As with any functional entity, a TIA has to be defined by its specific mission and its scope, and endowed with specific resources and powers of decision. In this chapter, we will focus primarily on infrastructure financing, and consider the interactions between such a TIA and diverse stakeholders.

2.2. Taking into Account Interactions between a Transport Infrastructure Agency and Stakeholders

2.2.1. Scope of a Transport Infrastructure Agency

The scope of a TIA has to be sufficiently precise. If not, uncertainty about infrastructure implementation due to political vacillations or counter-orders is not reduced by the creation of the TIA, and the field is open to lobbying games that influence the projects financed by the TIA. In this case, worthwhile projects may be unduly delayed, the TIA may be used to achieve self-interested short-term political aims (see Chapter 3 of this volume) and incompatible infrastructures may even result.

As regards the geographical and network limits of the TIA's scope, an extension of the agency's mandate should in theory enhance its performance. But it would need more information in order to act efficiently, therefore incurring higher costs. It would also have to respect more rules: the more extensive the powers that are given, the larger are the risks of collusion,[3] and the TIA will need careful observation and regulation.[4] The case of the Swiss multimodal fund (FINÖV) illustrates a broad but precise application domain, together with a transparent follow-up process (see Chapter 9 of this volume).

2.2.2. Resources of a Transport Infrastructure Agency

In order to reduce the uncertainty about infrastructure implementation, and to give the TIA sufficient independence and credibility, the TIA has to be given resources commensurate in magnitude and duration with the TIA's scope and lifespan. These resources can include regular financial inflows, exceptional transfers or assets such as land property. This is essential for the TIA to be able to borrow enough money and exert its financial leverage of public funds.

We will not discuss here the optimality of earmarking resources to a TIA. We will merely note that, for instance, even the International Monetary Fund has agreed that, in some cases, earmarking through road funds created in Africa and Latin America may be warranted.[5] The case studies of the German HGV (heavy goods vehicle) and of the Swiss multimodal fund (see Chapters 10 and 9 of this volume) show that considerations of the cost of public funds are also quite important in this respect.

The TIA's funds are either diverted from other uses, or are new resources coming from some actors' pockets. In any case, this is a potential source of conflict. The same is true of powers given to the TIA and of the resulting lobbying, which may result in the premature disappearance of the TIA. For instance, with its resources, the TIA may develop the part of the network within its authority, whereas the other parts of the network, directly under the authority of the ministry of transport, may decline, due to a parallel reduction in budget decided by the ministry in charge of the budget. If the final result is more or less the same allocation of resources as without any TIA, but with higher co-ordination costs, welfare decreases.

Interestingly, TIAs are sometimes managed by boards that include the main interest groups concerned.[6] This is a way of improving strategic and operational co-ordination between the TIA and stakeholders, as well as between stakeholders.

2.2.3. Powers of a Transport Infrastructure Agency

We will mainly focus here on the powers of a TIA that affect its overall financing ability. A TIA's powers of negotiation with banks and other sources of financing essentially depend on the quality of resources available to the TIA, on the risks that it faces, and on the public backing it receives.[7] In this context, "quality" refers not only to the amount of funding, but also to whether funding is likely to be sustained or increased over time, and whether the TIA has the power to influence its funding by varying prices or by other means. The actual independence of the TIA from the yearly lottery of budget allocation, and from the fluctuations in policy decisions, are key factors in the stability and credibility of a TIA.

If funding arrangements between the public authorities concerned and the TIA are not fixed, the TIA may gain bargaining power with the help of constraining rules. For instance, the bylaws of Réseau Ferré de France (RFF) stipulate that stakeholders requesting new infrastructure must contribute financially so that RFF's accounts are not adversely affected. Some power may be granted through exclusivity rights to the TIA on part of the transport network. Relative to other public co-financing entities, the TIA also has some power on the scheduling of the projects, thus having an indirect influence on the implementation of the plans and this gives it some indirect bargaining power.

Other important powers may be the power of contracting with concessionaires or of sub-contracting for the implementation of various tasks on the transport network, and some power over land use. Such a power of (sub-)contracting may include design and implementation of the contracts, and the selection of the concessionaire. A balance has to be found regarding the autonomy left to the TIA through these powers: they increase the risks of collusion or abuse of monopoly power, for instance, but, however, they allow greater financial leverage of public funds through securities on loans and may improve the economic use of assets left to the TIA.

2.3. Acceptability of a Transport Infrastructure Agency

Political bodies that are considering whether to create a TIA are, of course, influenced by the above considerations on conflicting powers or resources. Depending on the sensitivity of the public authorities to the lobbying activities that result, the decision to create a TIA may be abandoned or the TIA's authority may be reduced.

But the main hurdle for a TIA is to get approval from the delegating authority itself. Even though decision powers on planning and programming are not delegated, it may be in the general interest to delegate other powers but, at the same time, contradictory to the short-term political interest to do so. One key issue is the question of independence: we have seen in Section 2.2 that independence was essential for the credibility and long-term continuity of the TIA's actions. But as long as discretionary power exists with respect to transport infrastructure, the TIA would be at the same time more efficient but also less likely to be created by the delegating entity.

So the conditions that make a TIA socially desirable also reduce or destroy the motivation of the public authorities to create it. Indeed, most TIAs have been created either under international rules or pressures (from the World Bank, the European Union), or in countries with very developed and open democratic processes such as Switzerland. And the classical organisation through government departments has often proved to be quite efficient, as the high-quality and well-developed motorway networks in Western Europe demonstrate.

Nevertheless, strong incentives to create a TIA may exist when the classical organisation does not work properly. A TIA may speed up construction of an infrastructure plan that would be unattainable within a reasonable time limit due to budget constraints. Some authorities have seen a TIA as a simple way to outsource an important debt without jeopardising infrastructure development, though this matter is much more complex in reality. Other authorities have hoped that a TIA would protect a long-term investment policy against political changes and pressures. In practice, beyond the theoretically oriented reasons above, a TIA may be created just for pragmatic reasons. For instance, politicians may create a TIA for outsourcing specific functions that politicians feel to be quite technical, or because these functions are believed to be more of a source of problems than a source of good news for media coverage.[8]

In any case, the conditions that lead to the creation of a TIA are of utmost importance. For instance, creating a TIA does not guarantee that decisions will be made in the public interest. Capture by some lobby, or biased action, may still occur with a TIA unless democratic processes are imposed, for example, on the choice of the TIA's board, the transparency of its internal rules, the publicity given to its actions and their justification, or the regulation scheme of the TIA.

In light of the foregoing analysis one can identify four possible advantages of a TIA vis-à-vis public authorities in managing transport infrastructure.

Independence: The credibility of mid-term and long-term commitments on transport infrastructure development would be strengthened by a TIA – at least if it is granted sufficient financial autonomy to insulate it from changes in short-term political objectives.

Borrowing conditions and financial leverage: In some developing countries, an agency may have better borrowing conditions than a public entity, but if the country's risk premium is too high, or contrarily if the country's economic situation is quite good, this will not be the case. Financial leverage may be obtained more easily by the TIA, especially if it is allowed to use innovative financial schemes, with due public regulation, and has adequate resources.

Reduction of information asymmetries: An agency may help to reduce informational asymmetries, whatever its precise role may be. This is all the more important as, considering the great size of the infrastructure plan, agency and regulator must contract with large firms that have informational advantages and considerable market power. The creation of a TIA could be justified for the sole purpose of gathering information if doing so allows it to regulate more effectively in the long run. On the other hand, creating an agency deprived of this role would make no sense if the aim is to optimise the operation of the overall system.

A better contracting ability: The activity of the TIA may consist of the design, organisation, procurement and monitoring of incentive contracts. This activity of "contract engineering" may give the TIA useful experience for developing more efficient management of infrastructure development and use.[9]

On the whole, strengthening the mid-term and long-term continuity of infrastructure development, and performing efficient contract and financial engineering, through well-designed "independent" outsourcing, may be strong elements in favour of a TIA whenever traditional organisations are too limited in these matters. Nevertheless, the transparency and credibility of the TIA must be assured, and the reluctance of authorities to create such a TIA will be greater if the current political system is not transparent and gives opportunities for discretionary decisions.

3. PUBLIC–PRIVATE PARTNERSHIPS

Public–private partnerships can take many forms. An exhaustive analysis will not be provided here. We will only compare two contrasting cases: one in which the public authority builds and operates the infrastructure, and one in which these tasks are assigned to a private firm through a concession (build operate transfer – BOT). Several considerations are analysed: financial aspects, divergence of goals between the principal and the agent, effectiveness of the principal's control over the agent and effectiveness of regulation.

Financial aspects: From a macro-economic point of view, private financing is often presented as a way to circumvent Public Sector Borrowing Requirements (PSBR)[10] without avoiding the crowding out effects on private investment.[11]

From a micro-economic point of view, many arguments are presented in favour of the lower cost of public funds compared to private funds and the risk premium which they bear. But many of these arguments are questionable. One is that the public authority can diversify the risks.[12] This is true, but the risks may not be fully diversified. Indeed, public projects are usually concentrated in just a few sectors. Sectoral risks thus remain, and the risk diversification argument is diluted. Furthermore, the risk portfolio of large firms is also quite diversified. A second argument is that the state can spread the risks among many taxpayers, whereas in private firms stockholders are concentrated. That is true too, but stockholders may also be large collective investment funds and they may share their risks with other entities. A third argument is that the public loans bear no risk. This is also true, at least in European countries (in developing and transition countries, public loans may bear a large risk), but it is at the expense of the taxpayers, and the excess burden of taxes may outweigh the zero risk premium.

Divergence of objectives: Public and private organisations are governed by different objectives, usually described to a first approximation as general welfare[13] for public organisations and profit for private ones. The presence of either market power or externalities may exacerbate the difference. In fact, these objectives are a rough description of reality. In each of these organisations, the actors are driven by different objectives, for instance the politicians are worried by the re-election process whereas the managers of a firm wish to avoid dismissal while earning as much for themselves as possible.

Effectiveness of controls: In each organisation, control is achieved through a chain of relations. In the public sector, the chain goes from the elector to those who are elected to office, and then to the public service; in the private sector, the chain goes from the shareholder to the managers, then to staff within the firm.

The effectiveness of these chains is debatable, but it is generally acknowledged that private control is more effective than public control. Elected public officials are rather loosely controlled by the electors. In the case of privatisation of the management of a public service, the decrease in taxes may not be perceived by taxpayers, but the effort to achieve this decrease may harm some categories of workers, who will strongly resist the action. The politician has a strong information asymmetry vis-à-vis the public service, which may either oppose any change from the status quo or attempt to increase its budget instead of decreasing its costs.

The private sector also has failures. Although shareholders are interested in the firm's performance, they have little incentive to acquire detailed information about the firm or to exercise control unless their individual holdings are large. The control of managers is exerted by the stock market through the risk of bankruptcy and takeover. This control is rather loose, and may leave the manager with excessive freedom. Furthermore, the actors have divergent goals: the shareholder is interested in the profit, the manager in his salary.

The conclusion is that as long as externalities or market power are low, private management should in general be preferred, and public management in the opposite case. The precise choice between the two types of organisation depends on the quality of regulation, which in turn depends on factors such as asymmetry of information, uncertainty and market competition. In any case, each specific situation should be analysed individually and no absolute rule can be invoked.

4. THREE REFERENCE CASES

This section reviews institutional issues regarding transport infrastructure that arise in the European Union with respect to motorways, rail tracks and seaports.

4.1. Motorways

Suppose that a public authority or an infrastructure agency has to build a motorway. Should it build and run the motorway on its own or should it franchise it? If it franchises it, how should the concession be designed? Before addressing these questions it is helpful to highlight some defining characteristics of motorways.

- Motorways generate large external costs, both on the road (congestion on the motorway) and external (congestion on other network links, environmental damage).

- Traffic volumes and, to a lesser extent, costs are uncertain and difficult to forecast, especially in countries with little or no experience of (toll) motorways.
- There is not much information asymmetry between the regulator and the agent, at least ex-ante. Ex-post, once the motorway is operating, the agent gains an informational advantage about operating conditions such as the timing and magnitude of traffic peaks.
- The downstream market (mainly the trucking market, and in some sense too the private car traffic) is a competitive one in which no user accounts for an appreciable fraction of total traffic.
- Motorways serve various types of demands: local short-distance journeys and long international journeys, passengers and freight, etc. But they have little control over traffic other than through the structure and levels of tolls.[14] It is almost impossible to differentiate the product at the level of a given road link except through minor devices such as restriction of certain lanes and differential speed limits for heavy goods vehicles.[15]

4.1.1. Problems Related to Uncertainty

Apart from cost uncertainty, which is encountered in some specific local conditions, the main source of uncertainty for motorways arises with traffic forecasts. Under traditional auction mechanisms used to award concession contracts, uncertainty about traffic levels creates uncertainty about revenue. Some authors[16] argue that endogenous duration contracts should be used whereby firms would be invited to submit offers with claims for total revenue. This mechanism assures the firm that its declared costs will be covered, and therefore reduces the risk premium of the concession. However, possible side-effects, for instance on cost reduction efforts or bidding strategies, have to be controlled by the principal.

4.1.2. Problems Related to Efficiency

Geographical scope: Because of congestion, accidents and environmental spillovers, externalities may exist between a motorway and neighbouring links. These externalities have consequences for social efficiency if the motorway is privately operated. In the case of complementary links in series that are controlled by different operators, the outcome is double marginalisation and tolls that exceed the optimal level by a potentially large mark-up. By contrast, if the links are substitutes (e.g. parallel links between a common origin and destination[17]), competition constrains the mark-up charged by each operator.

These prototypical serial and parallel road network results are insightful, but their applicability to actual road networks in Europe is limited. Except in some urban situations, it is rare to find two or more parallel roads carrying traffic with a common origin and destination. Rather, the traffic has multiple origins and multiple destinations, and, except for some specific situations, a motorway is in competition with another motorway only for a small part of its total traffic. Generally speaking, a majority of motorway traffic is short-distance traffic; i.e. less than 200 km.

These considerations lead to three conclusions. First, control of market power of the concession is necessary. Second, to decrease the gap between the objectives of the concessionaire and social welfare, it is preferable to have complementary (serial) links provided by the same concession. And third, substitute links should be operated by different concessions since competition is welfare enhancing. Efficiency will be promoted if the franchise's network covers a market such that, in relation to this sub-network, the rest of the network has a minimum of complementary links and a maximum of substitute links. Of course, this kind of design also has to take into account the financial viability of the sub-network.

Tariff regulation: The previous sub-section highlights the need to regulate tolls in order to contain a concessionaire's market power, even if this market power is reduced by the presence of substitute links.[18] This toll regulation could be implemented through a price-cap, which would leave the risk to the concession, once the traffic risk is covered by the endogenous duration device and the risk on costs is estimated to be rather low.

This price-cap should be understood as an average price-cap, leaving open the possibility of varying tolls over time in order to cope with peak traffic loads. It is not sufficient to regulate average prices, leaving the concessionaire free to vary tariffs over time as it pleases. It can be shown that a profit-maximizing concessionaire would not choose the socially optimal peak/off-peak toll differential.[19] Unfortunately, the regulator has an informational disadvantage on real-time operations that leaves it in a bad position to stipulate peak-period tariffs. Information on traffic congestion is more easily available (e.g. through drivers' complaints or data from inductance loops under the road) and less costly to gather for the principal than the information needed for direct implementation of peak-pricing regulation.

Given these constraints a better form of regulation would be to impose a minimum quality of service level (for instance, a minimum speed), to be achieved through toll modulation. This is done using scheduled tolls on State Route 91 in Riverside County, California, and Highway 407 in Toronto. And it is done using responsive (real-time) tolling on Interstate 15 in

San Diego County, California, and on Interstate 394 in Minneapolis-St. Paul. However, given the multi-dimensional nature of service quality (e.g. average speeds and travel time reliability for each user group and on each link of the network), and the difficulty of defining and measuring them, great care is required in such a regulation and the flexibility given to the concessionaire should not become a blank cheque.

Regulation of the quality of service: In addition to the time variation required for congestion pricing, the question arises how to induce the concessionaire to provide a good technical quality of service in terms of smoothness of the road surface, speed of reaction to accidents and so on. It can be shown that a profit-maximizing operator will provide an optimal quality of service under reasonable assumptions.[20] But in practice some quality dimensions may be unobservable to users and even resented (for instance, preventive maintenance that temporarily disrupts traffic). In these cases, the price regulation should include a bonus that covers the cost of providing the requisite quality of service. This mechanism is used in Italy and Spain. It raises the question as to how quality can be measured for infrastructure operations and also about the measure of expenditures on quality (Muren, 2000).

Efficient investment incentives: Generally speaking, the infrastructure manager has no decision to take regarding large investments, which are decided by the public authority. It is nevertheless useful to consider the infrastructure manager's incentives to invest. One reason is that the manager may have many small investments to decide on such as widening a traffic lane or creating a new interchange. Another is that the manager may influence the public authority's decisions through lobbying.

The first point to note is that the principal and infrastructure manager generally disagree on the optimal timing of investments. Because the manager does not capture all the benefits of users, the socially optimal time to invest may be earlier than the timing that yields maximal financial returns. But if there is some form of competition between candidate concessionaires, the time of implementation may be advanced up to the time when the private benefit of the motorway is zero. It follows that the regulator, who controls the implementation date, may be subject to pressures from potential concessionaires to advance this date. The outcome of these two opposing biases is not clear.

Principal and agent may also disagree on the optimal size of investment once the choices of concessionaire and timing of investment have been made. In the case of a simple rate-of-return regulation,[21] the Averch–Johnson effect leads to over-investment. But if the regulation imposes short-run marginal social cost pricing, the concessionaire has an incentive to under-invest because this results in more congestion and considerably higher congestion tolls. A two-part rule that would avoid this distortion would be to:

- Impose a price-cap on the average tariff, and allow for time variation in the toll to achieve a minimal level of service (for instance, a minimum hourly speed).
- If demand is high so that the average tariff exceeds the price-cap, allocate the excess revenues to a fund managed by the principal.
- The optimal design of such a rule could depend on factors such as the conditions concerning the end of the contract, and would need to be analysed in depth.

4.1.3. Conclusions on Motorways

Optimal regulation of motorway concessions is fraught with several difficulties:

- Motorway traffic creates large externalities both for road users themselves (notably congestion and accident externalities on the link, and on substitute and complementary links) and those outside the system (local and global emissions and noise).
- Traffic forecasts are subject to large margins of error, which creates revenue uncertainty for prospective concessionaires.
- The regulator has an informational disadvantage about real-time traffic operations vis-à-vis the operator.
- Optimisation of real-time operations on the whole road network is difficult to accomplish through regulations.

These difficulties loom especially large in areas with high population densities and complex road networks. In such circumstances, the classical solution of building and operating motorways by the public authority may be preferable to granting concessions. Operation of the network could be sub-contracted to a private operator who would be rewarded not from the charging revenues, but by a fee linked to the achievement of some predetermined travel time goal.

If a "concession" is granted, consideration should be given to incorporating in the design an endogenous duration, time variation of tariffs with a cap on average tariffs, and a minimal quality-of-service level. If the tariff required to achieve a predetermined level of congestion exceeds the price cap, the excess revenues would be distributed to a fund managed by the public authority. A bonus on the price-cap could be granted depending on the technical quality of the service provided.

4.2. Rail Tracks

European railways have several distinguishing features. First, it is European Commission policy to implement separation between infrastructures and operations. In many countries, this reform has led to a clear separation of the infrastructure management vis-à-vis operations, and to market-type relationships between the infrastructure manager and the operators. Second, extensive programs of new, mainly high-speed, rail tracks have been devised by several member countries. At the European level, a large network of major transport infrastructures called the TEN-RTE network has been launched. This network is composed of 30 major corridors for which a majority of schemes are railway schemes.

In the European rail infrastructure management, the downstream market is a monopoly or at best an oligopoly, with an incumbent which is, or was in the past, merged with the infrastructure manager. It follows that, contrary to the cases where infrastructures and operations are unbundled, there is a problem for path allocation[22] to the operators. This path allocation is achieved not only through tariffs but also through a hierarchical process, based on grandfather rights and priority rules.

Apart from these European peculiarities, rail infrastructure management exhibits the same features as in other parts of the world, due to the characteristics of the industry:

- The operators have more options available to deal with demand than in the case of motorways (product and price differentiation).
- The environmental, congestion and scarcity[23] externalities are lower per passenger – kilometer or tonne – kilometer than for motorways.
- Multiple uses: Similar to motorways, railways serve diverse user segments except that the ‘‘short distance’’ segment represents a small proportion for freight, and is limited to large conurbations for passengers.

The challenges of designing railway concessions will be addressed by considering first information asymmetries and uncertainty, and then questions of efficiency.

4.2.1. Information Asymmetries and Uncertainty

Railway concessions usually provide a low rate of return, at least in Europe. A solution to financing would be to grant a monopoly to the incumbent when the profitability is low and to allow entry when the profitability is high. Unfortunately, as Caillaud and Tirole (2004) have demonstrated, the principal can hardly extract private information about demand from the incumbent firm through an auction of monopoly rights on rail service, because the more the incumbent is willing to pay for the franchise, the higher is the profitability and consequently the social benefit from not granting monopoly rights.

Though there is no evidence, it is reasonable to assume that there is a high degree of information asymmetry in favour of the incumbent; for example, uncertainty about demand is likely to be greater for potential entrants. This fact will be important for the rest of the analysis.

4.2.2. Efficiency

Tariff regulation: First let us note that, as downstream markets are monopolies or oligopolies, under the assumption of profit maximisation final transport prices are too high vis-à-vis the optimum. In order to correct this distortion, the regulator can lower the access price so as to lower the cost to the downstream firms. The usual tariff structure whereby charges per train – kilometer vary by type of train and track is non-optimal when – as is very likely – the infrastructure manager faces a binding budget constraint. In that case, efficient prices also depend on demand; i.e. on the origin and destination of the journey.

A further consideration is that tariffs alone are not sufficient to support efficient operator's services. Path allocations are implemented through procedures for solving conflicts based on priority rules that are consistent with the procedures used in the past by the historical incumbents. It is likely that these procedures, designed by an integrated monopoly, are not adequate in the presence of new entrants and of an independent infrastructure manager. Implementation of these allocation rules is open to a great deal of subjective interpretation, and an infrastructure manager may exploit this ambiguity to boost profits, thereby departing all the more from welfare maximisation. Strict regulation is therefore required. At the extreme, a regulator may implement the rules directly, especially if the rules are elaborate as is the case for complicated networks.

Quality of service regulation: The considerations reviewed above for motorways also apply to rail, and argue for an infrastructure fee based on observable characteristics of service quality. However, notable difficulties arise from the multiplicity of railway uses. Indeed, it is often difficult to simultaneously provide the different service quality levels required by different segments of demand (e.g. regional passenger trains and long-distance freight trains). Safety is an especially important and difficult issue, where there is a potential conflict between the profit of the infrastructure manager and welfare. This is the reason why public status for the infrastructure manager is recommended. Access charges should also include performance regimes, based on the reliability of the service in adhering to the timetable. This would need a clear definition and process for allocating responsibility since delays may originate not only from the operator in question, but also from the infrastructure manager, from other operators, and from interactions between all these players.

Proper inducement to invest: Unlike motorways, where investments are all-or-nothing, railway investments can vary in a more continuous way, through traffic management systems or relative to connection infrastructure (rail yards, intermodal transfer points, etc), for instance. Choosing appropriate levels of investment is important because it affects both product diversity and capacities. The difficulties of controlling the behaviour of the infrastructure manager are increased by the need for good investment decisions as well as good tariff and track allocation decisions, especially if the manager is in a weak financial situation. This is an argument in favour of direct management by the public authority.

Geographical scope of the concession: The geographical scope of the concession must take into account the typical length of journey, and thus should be greater for rail than for motorways. This conclusion is reinforced by the fact that, compared to roads, competition between parallel links is more rare and complementarities between adjacent links are more frequent. Private concessions should not be subject only to "local" regulation:[24] for instance, the regulation of allocation rules, of pricing and quality rules, should be appropriate to the multi-purpose uses of the link. The problem of the leeway given to the local regulator (or concessionaire) appears here: it may be in the local interest to favour a local segment of traffic, ignoring other segments (such as long-distance low-value traffic, or traffic transiting through the geographical area) although they may be of national importance, due to the external costs of alternative road transport, for instance. Some considerations on multiple principals are addressed in Section 5.

From this discussion, it is tempting to consider two sizes for the concession: first a limited number of concessions covering a small geographic area, granted to private firms, and easy to regulate, such as tunnels or specific links; secondly, very large areas, granted to public firms or managed directly by the public authority.

4.2.3. Conclusions on Rail Tracks

Compared to motorways, the management of rail operations is much more complex, as it relies not only on tariffs, but also on path-allocation rules and allows for more differentiated products with higher information asymmetry (traffic levels, cost, congestion level, etc.). It is more difficult to regulate the right level of investment, especially on safety. These facts militate in favour of implementing private concessions only for short and simple links, while operating large networks through public firms or directly by the public authority.

4.3. Port Infrastructure

The high degree of competition between European ports, combined with traditionally high public involvement in ports, leads to complex situations as regards the distribution of roles between public and private entities. Areas traditionally the prerogative of governments, such as safety, health and customs, are locally combined with services to the ship and services to freight, in interaction with ship owners confronted by fierce international competition, local authorities concerned with land use and land transport operators.

Compared to roads and railways, port infrastructures in Europe embody the following characteristics:

- Infrastructure is often dedicated to some specific uses (e.g. container traffic) or to some specific users, whether end users such as factories, or intermediary users such as ship owners.
- Demand is uncertain. High volatility may be caused by fierce competition between ports, and it may be subject to one or a few strategic actors' decisions.
- Port operators compete mainly on quality of service and price; product differentiation appears as regards organisation of the whole transport chain. Indeed, port infrastructure is but one of many elements necessary to provide complete transport services.
- Access of users to the infrastructure is regulated not only through tariffs but also through a hierarchical process of capacity allocation.
- Environmental externalities are on average lower than for motorways.
- Multiple principals: There is a potential problem of over-investment in ports that may be advantageous at the local level, but not globally due to cannibalisation of business between competing ports.

Similar to the treatment above for railways, the appropriate design of port concessions will be addressed by considering information asymmetries, uncertainty and efficiency.

4.3.1. Information Asymmetries and Uncertainty

Due to the degree of competition, port concessions often yield a low rate of return. This, together with considerations about risk diversification and management of scarce resources (water and land infrastructures, land use[25]), may explain why ports are usually local monopolies with high involvement of the public sector. The public central authority often has little knowledge about the demand for port services and the strategies of key stakeholders. Knowledge is better at the local level, but this does not eliminate uncertainty due to the high external risk mentioned above.

4.3.2. Efficiency

Tariff regulation: Due to fierce competition between ports, prices for port services tend to be low except for low-elasticity traffic in the short term.[26] Under budget constraints, Ramsey–Boiteux-type tariffs are appropriate. The usual tariff structure indeed has Ramsey–Boiteux features that reflect the bulk of the ship and the nature of the goods transported. Non-linear tariffs that feature quantity discounts based on ship bulk are commonly employed to attract larger ships.

The case for regulation is strengthened when safety is a serious concern, or when the economic impact of some types of goods and the financial interest of the port are incompatible. For instance, independent port operators may set high tariffs for goods that are essential to sustain local economic activities. Additionally, capacity allocation in peak hours or peak days may be an issue, mainly due to the congestion externalities involved.

Taking these elements into account, the regulator may play a direct role in capacity allocation, especially when safety is at stake, or may set some limits to the freedom of allocation by the operator. An example may be public intervention to increase the priority of access by barges compared to access by deep-sea ships, so as to facilitate the development of inland transport alternatives to road transport.

Quality of service regulation: The arguments developed for railways apply to ports, all the more so for historical ports that have high potential impacts on city life or tourism due to environmental or safety risks.

Proper inducement to invest: Port investments are usually expensive and long-lived, and are often dedicated to some specific demand. Therefore the risk and stakes are high, favouring solutions that imply public financing. Furthermore, demand evolves much more quickly than the time scale required to design, negotiate, approve and build port infrastructure projects. This is especially true for container infrastructure investment.[27] Added to the importance given to the expectations of key transport actors, this may lead ports to launch investment projects very early, possibly resulting in over-investment in some cases. The problem of the regulator is then that his role is often limited to negative actions such as delaying projects that may then come too late when demand rises. But the reverse is also possible for investments that are beneficial to the local economy but with low financial returns for the port operator, such as cruise-ship investment. This may explain why local interests are usually well represented within the governance structures of ports.

Geographical scope of the concession: Ports have functional and geographical characteristics with natural economies of scope (e.g. terrestrial and nautical infrastructure, land use, safety management) that favour integrated monopoly. But within a port some types of sub-concession by function may be possible such as dedicated or semi-dedicated terminals.

4.3.3. Conclusions on Port Infrastructure

Compared to railways, the management of port operations is arguably even more complex if the port traffic is dominated by products with volatile demand or a few concentrated users. It is difficult to assure appropriate levels of investment in infrastructure, and time scales are often difficult to co-ordinate. Consequently, private concessions may be appropriate only under limited circumstances, and duly regulated.[28] Otherwise, ports should be operated either through public firms or directly by the public authority with, as the case may be, private sub-concessions for functions such as operation of a specialised terminal.

5. QUESTIONS RAISED BY THE EXISTENCE OF SEVERAL PRINCIPALS

Regulation or control of transport infrastructure in Europe falls under the jurisdiction of multiple principals in two circumstances: for national transport when the regions and the state co-operate to finance infrastructure, and also for large projects when the national regulator and the European Commission both intervene. We present below several cases in which two principals interact, either to finance a project or to manage neighbouring infrastructures.

In the case where two regulators finance an investment project, imperfect information on costs or benefits for one of them may well lead to a loss of welfare. Indeed, some projects are not implemented even though they would increase welfare: imperfect information leads to a loss of welfare, and the fact that there are two regulators increases this sub-optimality.

The infrastructures of two adjacent countries carry both domestic traffic and international traffic. It is easy to show that, if each regulator is concerned only by the welfare of its citizens, and not by the welfare of the citizens of the other country, the outcome for infrastructure pricing is higher than the Ramsey–Boiteux tariff level. Bassanini and Pouyet (2000, 2003) analyse a setting where the infrastructure managers of two bordering countries are in charge of pricing access to their networks for downstream transport firms that provide international services. They analyse the virtues of various schemes of financing (with or without subsidies) and co-ordination processes for pricing the infrastructure. Among other results, they highlight the importance of the role of a supra-national authority facilitating international coordination, but at the same time emphasise that the welfare effects of policy measures aimed at favouring international services depend on the financing system adopted in each country as well as on the features of final demand.

On the whole, the examples identified above show that the existence of several principals may create inefficiencies. This point is developed in Martimort (1992).

These considerations change when moral hazard and incentives are introduced; see for instance Martimort (1996) and Caillaud, Julien, and Picard (1996). The main conclusions are the following: as far as incentives are concerned, decentralisation is superior to centralised management insofar as the decentralised regulators have better knowledge of the actions of the agent. Furthermore, decentralisation has advantages when communication of information from the local level to the central level is either costly or imperfect. Decentralisation, and more generally the existence of several principals, allows for better commitment vis-à-vis the operators.

6. EMPIRICAL ESTIMATES

The most important considerations developed in the previous sections can be summarised with a few numerical parameters, which characterise the outcome of the procedure chosen to manage the infrastructure:

1. The *cost of public funds* in the case of public funding;
2. The *risk premium* in the case of private finance;
3. The *gain in cost efficiency* – both for infrastructure construction and maintenance – in the case of concession;

4. The *quality of regulation,* i.e. the degree to which the private incentives of the concessionaire and the public objectives of the regulator can be aligned.

Let us elaborate on each of these parameters.

6.1. The Cost of Public Funds

The marginal cost of public funds (MCPF) has been extensively analysed in the economic literature, and the subject is reviewed in Chapter 3 of this volume. Let us just recall here that the MCPF measures the excess burden or loss of welfare caused by an extra euro of tax. The value of the MCPF depends on the type of tax (direct or indirect, on labour or on income), and on numerous other structural aspects of the economy. Estimates of the value of the MCPF derived from general equilibrium models typically fall in a range of 1–2 or more, with the most common values around 1.3.

6.2. The Risk Premium

The risk premium is relevant for investment funding. It adds to the interest rate without risk (R_f), which is the rate on Government bonds. The risk premium can be estimated for a given firm through the stock market, using the capital asset pricing model (CAPM) and its well-known hypotheses:

$$R_u - R_f = \beta_u(R_m - R_f)$$

where

R_u is the return on the equity;
R_f, the return on risk-free Government bonds;
R_m, the average return of the stock market; and
β_u, the volatility or beta of the equity.

The risk premium is the difference between the return on the equity and the return on the risk-free bond.

It can be shown that:

$$\beta_u = \frac{\mathrm{Cov}(R_u, R_m)}{\mathrm{Var}(R_m)}$$

Estimates of these parameters have been made by many authors for many economic sectors, but specific estimates for public utilities industries are rare. The most comprehensive estimates are found in Alexander, Estache, and Oliveri (2000).[29] They have estimated the values of risk premiums for public utilities all over the world. It appears that ($R_m - R_f$) ranges from 5 to 8 per cent, and R_f from 2 to 3 per cent.

β_u is calculated by these authors according to various situations. The estimates depend on the region, the sector and the type of regulation. For Europe, studies obtain mean values of 0.58 for equity and 0.43 for assets, which can be differentiated according to the mode (0.59 for airports, 0.44 for roads, 0.52 for rail[30]) and according to the regulatory type (0.49 for high-powered regulation, 0.46 for medium-powered regulation and 0.40 for low-powered regulation[31]).

6.3. The Gain in Cost Efficiency

Several studies show that concession to a private firm induces cost efficiency, both for infrastructure construction and for maintenance and operations. Evidence can be found in Vickers and Yarrow (1988) or Gómez-Ibáñez and Meyer (1993), as well as in the experience of technical bodies which had to compare motorway concessions and the same kind of motorways built by public services. No unique figure comes out of this body of results, and many of them are purely qualitative. Nevertheless it can be estimated that the gain in efficiency in motorway building and maintenance is around 5–10 per cent. This percentage should be understood as the reduction in cost provided by an average concession compared to the costs of a publicly run infrastructure.

There is a large literature on the efficiency of the rail industry, which shows that the gain in efficiency in the rail industry can be much higher, from 20 to 30 per cent, and depends on the regulatory process. But these figures should be applied to European case studies with caution since they pertain to experience in Japan, the USA and South America, and to firms that operate infrastructure and services rather than to construction of infrastructure. For the purpose of introducing values of parameters in a numerical model, we recommend the average figure of 20 per cent for maintenance of rail infrastructure. For infrastructure construction there is no direct evidence, and the values chosen for motorways can also be adopted for rail.

6.4. The Quality of Regulation

In classical incentive theory, regulation leaves an information rent to the operator that is linked to the operator's level of effort. It induces a welfare loss vis-à-vis the welfare which would be reached under situations of perfect information. The incentive theory provides formulae to calculate these losses. The most tractable formulae are based on several assumptions: there is no imperfection in the regulation process, there is a benevolent welfare-maximising regulator, there is no collusion or bribes, the operators are rational profit maximisers, and the probability distributions of unknown information are known and common knowledge. The real world is far from these assumptions, and dramatic failures can occur such as the well-known experiences of Railtrack and the French motorway TEO.[32] But these experiences have not been translated into hard figures for welfare losses; furthermore it is difficult ex-ante to forecast these bad situations.

A paper by Gagnepain and Ivaldi (2002) is specifically devoted to assessing the performance of incentive schemes in the operation of public transit systems (data from French cases). Cost-plus contracts are shown to perform less well than even fully uninformed[33] second-best incentive schemes; but price-cap contracts call for sufficiently informed regulators in order to be correctly implemented. However, this analysis relates to operations, not to infrastructure management.

Some studies[34] analyse the global level, trying to assess the impact of a comprehensive deregulation process, a privatisation process or the creation of new types of competition such as for air transport. Many studies address the gain in cost efficiency due to privatisation, depending on the type of regulation. Among other sources, the World Bank has provided data and empirical analyses on various sectors.[35] Transport sector problems have been addressed by Vickers and Yarrow (1988), Gómez-Ibáñez and Meyer (1993), Friebel, Ivaldi, and Vibes (2003) and Estache et al. (2003) *inter alios*. From these studies coefficients can be estimated for each specific situation or case study. A stream of studies[36] has examined the effect of the regulatory process in the rail industry. They find efficiency gains of a few percentage points for high-powered regulations. However, the transaction costs and rent/risk premium may counterbalance the intrinsic advantage of high-powered regulation, when overall welfare is considered. Furthermore, these results only provide hints for our problem as, unfortunately, they relate to services and not to infrastructures.

Taking stock of these results, it appears that the unique way to provide quantitative estimates for use in a model such as MOLINO (see Chapter 5 of this volume) is to make expert guesses about the cost of information rent and imperfect control of the effort of the infrastructure operator (say, for instance, a magnitude of 10 per cent of the infrastructure cost). Furthermore, careful attention should be paid to the design of the contract. As shown in the above-mentioned examples, poor design can lead to dramatic losses. Due to the lack of empirical studies, it is unfortunately difficult to be more precise about this crucial problem.

7. CONCLUSION

Contracting policy and institutional arrangements in transport infrastructure have been analysed here so as to give guidelines useful to public choice. The reader will have noted that, apart from some clear-cut assessments, no automatic choice is proposed: each specific situation should be considered individually. The main items and parameters mentioned in this chapter will be helpful for designing a regulatory scheme appropriate for public goals.

Though no automatic rule emerges, a meta-method can be outlined here, making use at each step of the corresponding parts of this chapter:

- First, clarify and formalise what the public goals are for the situation encountered, and identify possible conflicts (Section 5).
- Among the whole range of possible functions for a transport infrastructure agency, consider which ones could be useful in this context, as compared to an internal public entity.
- If some functions appear to be potentially important, estimate in more depth under which conditions of scope, resources, powers and constraints they could be successfully undertaken by an infrastructure agency, in the short and in the long run.
- Concerning the diverse possibilities of intervention by the private sector, after a preliminary review on a general level (Section 3), the list proposed in Section 4 (uncertainty and information asymmetry, efficiency: scope, pricing regulation, quality of service, level of investment) may be helpful, whether directly if the situation considered fits into the modal cases treated in Section 4, or as guidance for public reflection in all cases.
- Some hints are given in Section 5, relative to the recurrent problem of multiple public authorities.
- At each step, the indicative empirical estimates mentioned in Section 6 may be used to assess quantitative elements.

As a whole, this ad hoc process is quite consistent with a key message of this chapter: the quality of design and effectiveness of operation of the regulatory scheme are of paramount importance, whatever solution of internal organisation or outsourcing is adopted. The higher the degree of outsourcing, the more crucial this issue of quality becomes.

NOTES

1. This theory considers a ‘‘principal’’ and an ‘‘agent’’. The principal designs a menu of possible contracts in order to select the best agent and/or to have him act in the best way relative to the principal’s objectives. At the same time, the choice of contract made by the agent gives valuable information to the principal, thereby reducing the asymmetry of information between principal and agent.

2. One exception is Japan where the councils that choose basic construction plans include scholars and researchers.

3. Moreover, enlarging the field may reduce the degree of competition in the market for transport services encountered by the agency and allow it to exercise more market power.

4. It is not too difficult to assess the direct performance of a motorway company on an isolated link, but it is far less easy to do so on the scale of a network. And it is all the more difficult to assess the performance of a multi-modal agency.

5. The goal underlying earmarking is to assure funding for activities that are socially rewarding in political environments where funds might otherwise be diverted to other uses. Earmarking is briefly discussed in Chapter 3 of this volume.

6. For example, the organisational structure of many seaports in Europe includes representatives of the port users, ship owners, local authorities, local firms and staff members too.

7. A TIA performs tasks of public interest, and it is normally a public entity. Consequently, it is unlikely that the public authority would not intervene if the TIA gets into major difficulties.

8. An example would be air traffic control.

9. Such effects may also occur within a classical government office, but may be limited there by the weight of administrative processes that impede innovative schemes, by the limited scope of administrative contracts, by rigid hiring rules and so on.

10. PSBR are constraints imposed on public sector borrowing by international agreements (such as the Maastricht agreements) or constitutional regulations.

11. This argument does not hold for shadow tolls. In the shadow toll system, which is used for motorways in the UK, the loans made by the franchisee are reimbursed by payments from the public authority, which depend on traffic volumes. The toll received by the franchisee is paid by the public authority rather than by users.

12. For a discussion of risk see Abelson (2005, Section 3.2).

13. More precisely, a weighted sum of consumers’ surplus and firms’ profits, the weights taking into account possible distributional concerns.

14. Some additional control can be exerted by choosing the level of maintenance and provision of paid and unpaid services (e.g. road signs, variable message signs, rest stops, petrol stations).

15. The balance between the various uses of a motorway may be an issue, notably when co-financing or co-supporting of the project involves entities with different objectives (multiple principals). This topic will be considered in Section 5.

16. See, for instance, Engel, Fischer, and Galetovic (2001) and Nombela and de Rus (2004).

17. Oligopolistic competition between parallel toll roads is analysed in Engel, Fischer, and Galetovic (2004). See also De Palma and Lindsey (2000, 2002), De Borger, Proost, and Van Dender (2005) and De Borger, Dunkerley, and Proost (2007).

18. Incentives theory indicates that, even if price-cap regulation is highly powered, it is not the optimal form of regulation. Offering a concessionaire a menu of contracts would help to improve oversight by inducing the concessionaire to reveal part of its information. Nevertheless, we will use the term "price-cap" from now on since we will try to go more deeply into specific issues, and also because price-caps are frequently used.

19. Depending on the values of the price elasticities, it may even happen that the concessionaire lowers the tolls at peak hours.

20. See Tirole (1988) or Quinet and Vickerman (2004).

21. Under the Averch–Johnson type of rate-of-return regulation, a ceiling is imposed on the firm's realised rate of return on its physical capital. See Guthrie (2005) for a more elaborate discussion of rate-of-return regulation.

22. For railway operations, since two trains have to be kept well apart for safety reasons, the use of a track during a precise time interval has to be scheduled and allocated for each train of each operator.

23. In rail operations, congestion occurs when one train impedes another train from running at its optimal time (the second service has to be postponed for instance). Scarcity is manifest if the introduction of a new train service forces an existing service to be cancelled (the simplest case being two trains running at the same time in opposite directions on a single track).

24. Meaning a regulation that takes into account only limited local parameters.

25. The coastal areas available for port activity are shrinking. They represent high values for many conflicting uses, and they experience increasing environmental concerns and protection issues.

26. For instance local factories with high relocation costs and long delays, such as refineries.

27. Note for instance that the so-called "Port 2000" project in Le Havre became operational only in 2006.

28. For example, regulation may be needed in order to avoid having a concessionaire acting as monopoly landlord instead of developing transport services.

29. Many papers deal with more specific issues, such as Estache and Pinglo (2004), who analyse returns earned in infrastructure in developing countries since the Asian crisis.

30. The values for rail pertain to firms, which run both the infrastructure and the services, as is the case in all parts of the world except Europe.

31. Regulation regimes are ranked from "high powered" down to "low powered" along with the decreasing incentives for cost reduction they offer companies. A high-powered regime, such as price-cap or revenue-cap, has significant incentives, whereas a low-powered regime is basically some kind of rate-of-return.

32. This project was a franchised toll motorway in the area of Lyons, France (see Raux & Souche, 2004). The initial concession would have reduced overall welfare, therefore jeopardising the acceptability and viability of the project.

33. The regulator is said to be fully uninformed when he knows only the minimum and maximum possible levels of efficiency of the contractor.

34. See, for example, Morrison and Winston (1989) and Maillebiau and Hansen (1995).

35. For instance, Guasch, Laffont, and Straub (2003) tackle water and transport concessions and regulatory design.

36. For instance, Pestieau and Tulkens (1990), Oum and Yu (1994), Oum, Waters, and Yu (1999), Cantos, Pastor, and Serrano (1999) and Friebel et al.(2003).

REFERENCES

Abelson, P. (2005). Financing transport infrastructure: Public finance issues. In: K. J. Button & D. A. Hensher (Eds), *Handbook of transport strategy, policy and institutions* (Vol. 6, pp. 359–373). Amsterdam: Elsevier.

Alexander, I., Estache, A., & Oliveri, A. (2000). A few things transport regulators should know about risk and the cost of capital. *Utilities Policy*, *9*(1), 1–13.

Bassanini, A., & Pouyet, J. (2000). Access pricing for interconnected vertically separated industries. In: C. N. E. Niskanen (Ed.), *Helsinki workshop on infrastructure charging on railways* (pp. 41–71). VATT-Discussion Papers 245, August. http://en.vatt.fi/file/vatt_publication_pdf/k245.pdf

Bassanini, A., & Pouyet, J. (2003). Strategic choice of financing systems in regulated and interconnected industries. Communication to the *first conference on railroad industry structure, competition and investment*, Toulouse, November 7–8.

Caillaud, B., Julien, B., & Picard, P. (1996). National versus European incentive policies: Bargaining, information and coordination. *European Economic Review*, *40*, 91–111.

Caillaud, B., & Tirole, J. (2004). Essential facility financing and market structure. *Journal of Public Economics*, *88*(3–4), 667–694.

Cantos, P., Pastor, J., & Serrano, L. (1999). Productivity, efficiency and technical change in the European railways : A non-parametric approach. *Transportation*, *26*(4), 337–357.

De Borger, B., Dunkerley, F., & Proost, S. (2007). Strategic investment and pricing decisions in a congested transport corridor. Under press in *Journal of Urban Economics*.

De Borger, B., Proost, S., & Van Dender, K. (2005). Congestion and tax competition on a parallel network. *European Economic Review*, *49*(8), 2013–2040.

de Palma, A., & Lindsey, R. (2000). Private toll roads: Competition under various ownership regimes. *Annals of Regional Science*, *34*, 13–35.

de Palma, A., & Lindsey, R. (2002). Private roads, competition and incentives to adopt time-based congestion tolling. *Journal of Urban Economics*, *52*, 217–241.

Engel, E. M., Fischer, R., & Galetovic, A. (2001). Least-present-value-of-revenue auctions and highway franchising. *Journal of Political Economy*, *109*, 993–1020.

Engel, E. M., Fischer, R., & Galetovic, A. (2004). Toll competition among congested roads. *Topics in Economic Analysis & Policy*, *4*(1), Article 4, Berkeley Electronic Press. http://www.bepress.com/bejeap/topics/vo14/iss1/art4/

Estache, A., Guasch, J. L., & Trujillo, L. (2003). *Price caps, efficiency payoffs, and infrastructure contract renegotiation in Latin America*. World Bank Policy Research Working Paper 3129.

Estache, A., & Pinglo, M. E. (2004). *Are returns to private infrastructure in developing countries consistent with risks since the asian crisis*. World Bank Policy Research Working Paper, No3373.

Friebel, G., Ivaldi, M., & Vibes, C. (2003). *Railway (de)regulation : A European efficiency comparison*. IDEI Working Paper 221. Updated version (2005) available at SSRN: http://ssrn.com/abstract=505302

Gagnepain, P., & Ivaldi, M. (2002). Incentive regulatory policies : The case of public transit systems in France. *RAND Journal of Economics*, *33*(4), 605–629.

Gómez-Ibáñez, J., & Meyer, J. (1993). *Going private: The international experience with transport privatization*. Washington: Brooking.

Guasch, J. L., Laffont, J. J., & Straub, S. (2003). *Renegotiation of concession contracts in Latin America*. World Bank Policy Research Working Paper 3011.

Guthrie, G. (2005). *Regulating infrastructure: The impact on risk and investment*. Victoria University of Wellington. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=757014

Laffont, J.-J., & Tirole, J. (1993). *A theory of incentives in procurement and regulation*. Cambridge, MA: MIT Press.

Maillebiau, E., & Hansen, M. (1995). Demand and consumer welfare impacts of airline liberalisation: The case of the North Atlantic. *Journal of Transport Economics and Policy*, *29*(2), 115–136.

Martimort, D. (1992). Multiprincipaux avec anti-sélection. *Annales d'économie et de statistique*, *28*, 2–37.

Martimort, D. (1996). The multiprincipal nature of government. *European Economic Review*, *40*, 673–685.

Morrison, S. A., & Winston, C. (1989). Enhancing the performance of the deregulated air transportation system. Brookings Papers on Microeconomics.

Muren, A. (2000). Quality assurance in competitively tendered contracts. *Journal of Transport Economics and Policy*, *34*(1), 99–112.

Nombela, G., & de Rus, G. (2004). Flexible-term contracts for road franchising. *Transportation Research Part A*, *38*, 163–179.

Oum, T., Waters, W., II., & Yu, Ch. (1999). A survey of productivity and efficiency measurement in rail transport. *Journal of Transport Economics and Policy*, *33*(1), 9–42.

Oum, T., & Yu, Ch. (1994). Economic efficiency of railways and implications for public policy: A comparative study of the OECD countries' railways. *Journal of Transport Economics and Policy*, *28*, 121–138.

Pestieau, P., & Tulkens, H. (1990). *Assessing the performance of public sector activities : Some recent evidence from the productivity efficiency viewpoint.* November. Programme d'actions de recherches concertées n 87–92/106 with CORE and n 90–94/141 with Université de Liège, University of Liège.

Quinet, E., & Vickerman, R. (2004). *Principles of transport economics*. Cheltenham: Edward Elgar.

Raux, C., & Souche, S. (2004). The acceptability of urban road pricing: A theoretical analysis applied to experience in Lyon. *Journal of Transport Economics and Policy*, *38*(2), 191–215.

Tirole, J. (1988). *Theory of industrial organisation*. Cambridge, MA: MIT Press.

Vickers, J., & Yarrow, G. (1988). *Privatization: An Economic Analysis*. Cambridge, MA: MIT Press.

This page intentionally left blank

CHAPTER 5

COMPARING ALTERNATIVE PRICING AND REVENUE USE STRATEGIES WITH THE MOLINO MODEL

André de Palma, Robin Lindsey, Stef Proost and Saskia Van der Loo

ABSTRACT

Cost–benefit analysis plays a central role in planning and investment decisions related to transportation. Yet, this process is often rather obscure and difficult to control and check by an outsider. We propose here a new engineering-economic-based tool, MOLINO, to perform cost benefit analysis of transport projects and regulations in a network and multi-period context. MOLINO performs cost–benefit analysis for different transport modes and types of freight and/or passenger traffic, peak and off-peak time periods, diverse market structures (private or public monopoly or duopoly, regulated or unregulated) and various financing schemes. Congestion levels are computed endogenously. MOLINO computes costs and benefits over multiple periods and the length of the time horizon is flexible. Outputs include equilibrium values of user and social benefits, financial flows and measures of effectiveness such as congestion delays.

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 111–131

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19005-3

1. INTRODUCTION

The goal of this chapter is to integrate the theoretical prescriptions for optimal pricing, investment and revenue use as developed in Chapters 2, 3 and 4 in this volume into one simple quantitative assessment model. A quantitative model is needed because the theoretical guidelines contain elements of contract theory, public economics, political economy and transport economics. Often, the best solution is not obvious and one will need to compute numerically the net benefits of the different options. Moreover, the optimal solution may depend on the normative preferences of the policy maker who may favour a particular equity/efficiency trade-off, or – for historical reasons – place more trust in a particular type of institution.

We propose therefore a general model MOLINO to assess alternatives. Because the alternatives are more complex than a single investment or a simple change in pricing regime or revenue-use rule, the comparison will be between alternative regulation schemes. By regulation schemes we mean a complete description of the market context, and the pricing, revenue use and investment rules that are used by the different players. The MOLINO model should be seen as an assessment scheme. In principle every problem can be studied with a specific transport model. However, it may be easier and more consistent to use the same simple model to assess very different projects.

The core of the model is a representation of the transport market with two alternatives. These alternatives can be of two different modes or two parallel routes for the same mode. Each alternative can be used for freight and passenger transport, and a distinction is made between peak and off-peak periods. The user cost of each alternative is determined by its generalised cost, which is endogenous. The time cost component depends on the ratio of volume to capacity, and the money price component depends on the market regime in which the two operators function and the taxes and tolls are set by local and federal governments. The transport market model computes a user equilibrium and an equilibrium for the price-setting game between operators or infrastructure managers. The core of MOLINO is completed with a financial fund module and with welfare functions for local and federal governments that include external costs and public finance variables.

The MOLINO model can be used to study diverse transport policies ranging from a cordon toll in a city to the pricing of port services. In this chapter, we describe the MOLINO model and provide some details about its implementation. In Section 2 we provide an overview of the model. In Section 3 we summarise its different components. Sections 4 and 5 discuss respectively the investment module and the financial accounting and funds module. Section 6 presents the welfare assessment module, Section 7 discusses the different possible regulation schemes, and Section 8 concludes with a brief discussion of the software implementation.

2. OVERALL STRUCTURE OF THE MOLINO MODEL

2.1. Overview

Fig. 5.1 provides an overview of the MOLINO model. There are two types of input data. The first type of data, depicted in the top-left box of Fig. 5.1, is needed to calibrate the model to the case study. The data must suffice to describe the baseline equilibrium flows, speeds and prices over the time horizon, $t = 1, \ldots ,T$, as well as the infrastructure stock and the initial financial structure.[1] The second type of input data, depicted in the top-right box of Fig. 5.1, are the policy inputs that define a regulation scheme over the time horizon, $t = 1, \ldots ,T$. These inputs include rules for pricing, investment and revenue use, as well as the types of contracts used for the different transport alternatives. To calibrate the model, it is necessary to define a

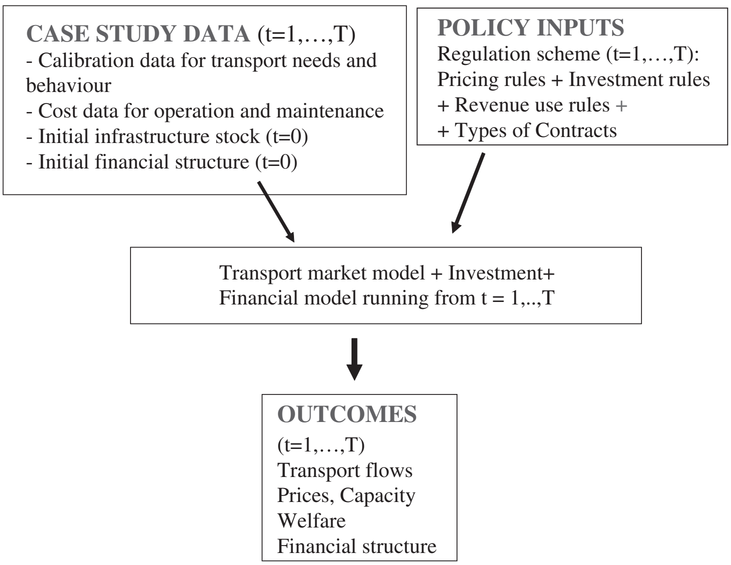

Fig. 5.1. General Structure of the MOLINO Model.

baseline input set that specifies both types of input data for a reference case: demand- and cost-input data and a baseline regulation scheme. Once the model is calibrated, one can assess alternative regulation schemes by changing only the policy inputs.

The middle box in Fig. 5.1 is a calibration and simulation module for the two transport alternatives. It calibrates the model in the reference case. Once calibrated, it can be used to compute the transport equilibrium for a given regulation scheme. This module computes both user equilibria and transport prices – where the prices are themselves Nash or Stackelberg equilibrium prices, if the operators of the alternatives set prices non-cooperatively.

The output box at the bottom of Fig. 5.1 reports measures of effectiveness for the regulation scheme including welfare, revenues, financial structure, etc.

2.2. The Dynamics of the Model and the Role of the Different Modules

The MOLINO model includes several categories of agents. There are two types of passenger users (e.g. poor and rich users) and two types of freight users (e.g. local and transit users). For each of the two transport alternatives (different modes, or different routes) there is an infrastructure manager who takes decisions on infrastructure (capacity) and a transport service operator who uses the infrastructure to deliver transport services to the users and sets the charges that users pay. Table 5.1 illustrates these agent types for rail, road and inland waterway modes. In addition there are two types of government agents: one local government (which disregards the benefits of transit users) and one federal government.

To study revenue use and transport investments, a dynamic approach is required that specifies the use of capacity extensions, revenue streams and financial structures over the time horizon. The simplest model approach is a recursive structure in which investment decisions are taken every period on the basis of some form of expectations. It is assumed that investments

Table 5.1. Illustration of Role of Different Non-government Agents in MOLINO.

Mode	Infrastructure Manager	Service Operator	User
Rail	Rail infrastructure manager	Rail operator	
Road	Road authority	Bus company	Car user Truck user Bus passenger
Inland waterway (IWW)	IWW authority		

initiated in period t become available in period $t+1$, and that the financial structure variables of period $t-1$ determine the investment options in period t.

The dynamics of the MOLINO model are shown in Fig. 5.2. A modelling period can be defined as a year, or a period of several years represented by a single modelling period. As indicated at the top of Fig. 5.2, the regulation scheme consists of rules for pricing and contracting, operations, rules for investment and contracting of investments, regulations for financial structure, and cross-subsidy rules and break-even constraints. The regulation scheme affects how the model works throughout the time horizon $t = 1, \ldots, T$.

For each modelling period MOLINO makes use of four modules: the transport module, the investment module, the financial reporting module and the infrastructure fund module.

The **transport market module** is the most important. It describes, for a given period, a given infrastructure, and a given regulation scheme the demand for and supply of each transport service. Supply is chosen by

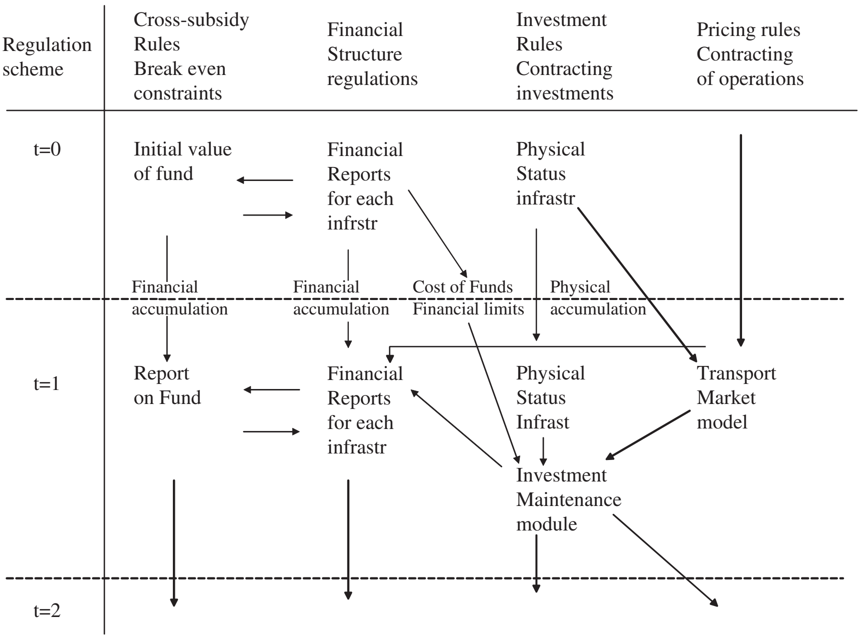

Fig. 5.2. The Modules of the MOLINO Model and its Dynamics.

infrastructure operators (e.g. a rail company determines price and frequency of service). Demand results from decisions of passenger and freight users. Pricing rules affect this module via their effects on volumes and prices on the transport markets, contracting of operations affects the operation costs and these are an input into the price-setting process.

The **investment module** (one for each of the two transport alternatives considered) keeps track of the physical capacity that is available as well as its quality. In the investment module an infrastructure manager decides on investments as a function of user benefits, expected profits, financial constraints and cost of capital.

The **financial reporting module** (one for each of the two transport alternatives considered) reports the incomes and expenditures, as well as how investments are financed and the resulting status of assets and liabilities. The financial reporting module also records the subsidies received from or given to other modes or operators via transport funds.

The **infrastructure fund** module records the operation of the funds (one for each alternative but they can also be merged in some cases), including its accumulation over time, its income received (if any) from each mode and the disbursement of subsidies (if any) to each mode.

To illustrate the functioning of the model over the time horizon, we will briefly describe the inputs and outputs of the different modules for the periods 0–1. In period 0, a regulation scheme is specified as defined in the top row of Fig. 5.2. Initial values are required for all the stock variables: the physical status of the infrastructure, assets and liabilities for each infrastructure manager, and initial balances for the infrastructure funds. We use Fig. 5.2 to guide us and within each period we move from right to left.

In period 1, each infrastructure manager inherits from period 0 a physical infrastructure, and puts it at the disposal of the operator for a user's fee. The operator implements a pricing rule and sets quality (say frequency of rail service) for the use of the infrastructure by the final users (passengers and freight). In setting prices, the operator makes use of the information on the cost parameters associated with the type of contracting for the operations. The pricing rules embody the market behaviour (non-cooperative, cooperative) as well as possible break-even constraints and an objective function (e.g. profit maximisation). The users of transport infrastructure take prices and quality of infrastructure as given. The behaviour of infrastructure managers, operators and users jointly determine the transport market equilibrium for period 1.

In period 1 each of the infrastructure managers has inherited physical infrastructure and financial stock variables (debt, financial reserves, grants from the transport fund, etc.). The infrastructure managers choose how much to charge the operators of transport services, and they make decisions on investment and possibly maintenance. New investments become operational in the next period. The investment rules incorporate three elements: expectations about future market conditions, inherited physical infrastructure and financial stock variables. The transport market equilibrium, together with the investment decisions, will determine the financial results for period 1. The resulting financial structure may determine the cost of capital and financial constraints that affect investment possibilities.

In period 2, the infrastructure managers make new pricing decisions taking into account the new infrastructure capacities, and so on for each period until the end of the time horizon is reached. Important overall modelling assumptions are that the model is deterministic and all agents have myopic expectations.[2] A stochastic model with learning over time about demand and cost parameters may be more realistic to analyse such aspects of transport infrastructure financing as public–private partnerships and risk taking. Both features are interesting avenues for further research.

3. COMPONENTS OF THE TRANSPORT MODULE

The transport module described here has a very simple structure that can be given different interpretations depending on the case study at hand. Table 5.2 identifies variables that are included in the transport module for each year.

The structure of the transport module is illustrated in Fig. 5.3. There are four types of users (two types for passengers, and two for freight) and two competing transport options, "Supply 1" and "Supply 2", that can stand for various mode combinations; e.g. road–road, road–rail, rail–air, intermodal-road, etc. Transport demand and supply interact through prices and service quality for given capacity to reach equilibrium in each period. The regulation scheme specifies pricing, investment and revenue use as well as the cost parameters that result from the type of contracting of operations (see Section 3.2).

The equilibrium of the transport market for period t determines two types of stock variables: the financial stock variables (including "infrastructure funds") and the infrastructure capacity. Only these two stock variables are carried forward as state variables into period $t+1$. The capacity decisions of final transport infrastructure users, such as the vehicle stock, are assumed to be optimal conditional on the flows and are not accounted for in the model.

Table 5.2. Components of the Transport Model.

	Passenger	Freight
4 User categories	2 Types of users (e.g. poor & rich)	Local and transit freight
2 Modes	• Free highway & toll highway • Road & rail • Rail & air	• Free highway & toll highway • Road & rail ....
Time periods	Peak and off-peak	Peak and off-peak
Elasticity of total trip demand	Elastic	Elastic
Service quality	Dimensions of quality can include: • congestion delay • smoothness of road surface • reliability • ease of toll payment	

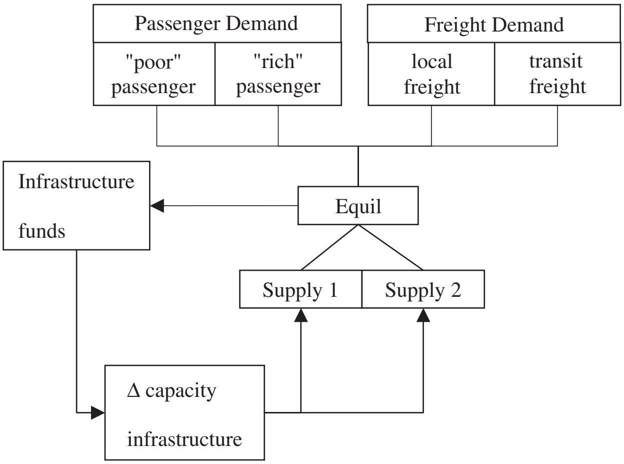

Fig. 5.3. Structure of Transport Module.

3.1. Demand for Transport

Passenger transport preferences are modelled using nested CES-type utility functions (see Keller, 1976). The nested utility structure is represented in Fig. 5.4; this structure applies for each type of consumer and each modelling period. At each level, consumers choose between two options based on the

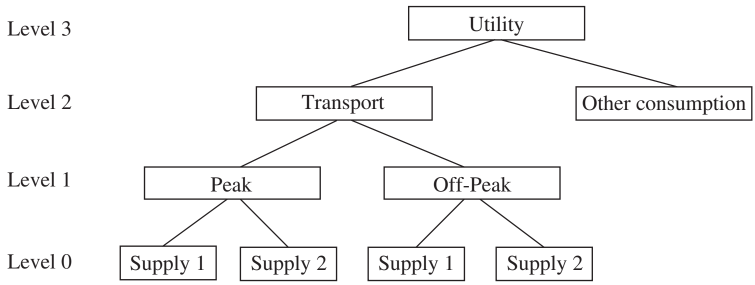

Fig. 5.4. Utility Tree for Passenger Transport.

relative prices, their incomes and their preferences (captured in the elasticities of substitution at the different levels and their initial expenditure shares).

We distinguish two categories of consumers, e.g. "poor" and "rich" consumers as in Fig. 5.3, who can differ in preferences and incomes. Whereas both types of users face the same set of choices (i.e. the utility trees corresponding to the two consumer types are identical), their elasticities of substitution may differ and they can respond to prices in different ways. The distinction between the two types of consumer is important from a welfare economics perspective when computing the equity impacts of alternative policies.

The main advantage of CES utility functions is that they are easy to calibrate to a case study. All that is needed for every period are modal shares, total income and four elasticities of substitution. The four elasticities required are: one capturing the ease of substitution between transport and the consumption of other goods (top branch of the tree in Fig. 5.4), one for the substitution between peak and off-peak travelling (second branch in Fig. 5.4) and finally two elasticities of substitution reflecting preferences between the two modes during peak and off-peak hours (lowest pair of branches in the tree). The utility functions should be interpreted as aggregates of many individual preferences; they do not represent the preferences of any individual user.[3]

Nested CES-type functions are also used to model freight transport, this time for the cost functions of the producers (see Fig. 5.5). This structure applies for each type of freight user and each modelling period. An important assumption is that the total production of the firms that use freight transport is fixed. This avoids the need to specify a general equilibrium model while retaining a variable demand for freight because firms can alter

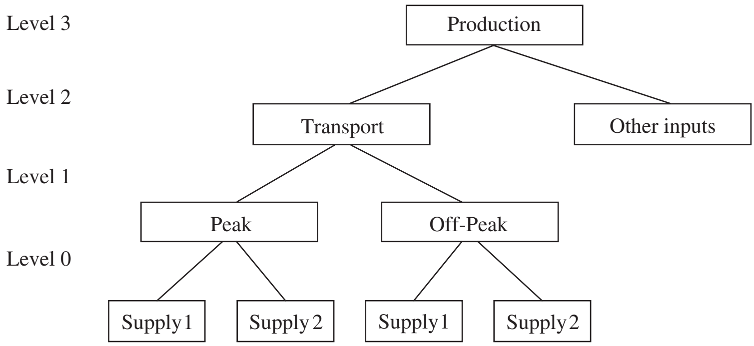

Fig. 5.5. Representation of Freight Transport Demand.

the input mix between freight and other inputs to minimise the cost of a given output. There are two types of freight transport users, transit and local, which may have different elasticities of substitution. This can be important as local governments may favour local freight and charge transit freight more if they can.

The main advantage of the nested CES functional form to model demand for freight transport is again its ease of calibration. As in the case of passenger transport, the only information required for each period are modal shares, factor share of freight transport, total production costs and four elasticities of substitution. The MOLINO model provides default values for elasticities of substitution that can be replaced by case-specific information whenever better information is available.

3.2. Cost Functions

The different types of costs and their characteristics are summarised in Table 5.3. The MOLINO model employs rather simple default functional forms that can easily be changed if this is warranted by the information available. The cost function information from Chapters 2 and 4 in this volume can be used to choose default values. As the model is not stochastic, procurement issues are modelled in a very simple way: as a fraction by which costs of infrastructure, maintenance and operation are increased if no incentives are used to decrease costs; e.g. by selecting private suppliers by tender instead of using in-house production.

Table 5.3. Cost Functions used in the MOLINO Model.

	Type of Cost	Assumptions used in MOLINO	Does Procurement Matter?
Investment in infrastructure	Investment cost	Function of investment and existing capacity	Yes, tendering versus non-tendering
Maintenance of infrastructure	Maintenance	Function of existing capacity and of total use by type of user	Yes, tendering versus non-tendering
Operation cost	All operation costs (building, vehicles and other)	Fixed cost + Variable cost that is a function of total use by type of user	Yes, tendering versus non-tendering
User cost	Time costs	Bottleneck formulation: function of volume over capacity	No
User cost	Resource costs	Proportional to volume of transport by user	No
External cost (other than congestion)	Air pollution, accidents, noise, etc.	Constant per trip, depends on type of user	

3.3. Structure of Taxes, Charges and Prices

In the transport module, the passenger and freight transport users make decisions on the basis of generalised prices that include time costs (endogenous in the case of congestion) and money prices. The money price is the sum of tolls, charges, tickets and taxes as illustrated in Fig. 5.6. The taxes are aggregated into two categories: local and federal. All tolls and charges are aggregated into one payment per unit of transport service provided by the operator. The arrows in Fig. 5.6 represent payments.

3.4. Market Structure and Behaviour of Suppliers

The operators of the two transport services may cooperate or not, and their goals may be to maximise profits, local welfare or global welfare. Table 5.4 illustrates a few possible market structures for the operators. Duopoly is modelled as Nash behaviour in prices. This means that each operator takes the prices of the competitor as given when maximising its objective function (profit for a private firm or local welfare for a local government). Prices are assumed to be the choice variables rather than quantities since capacities are given.[4] Other market structures can be considered such as a mixed oligopoly

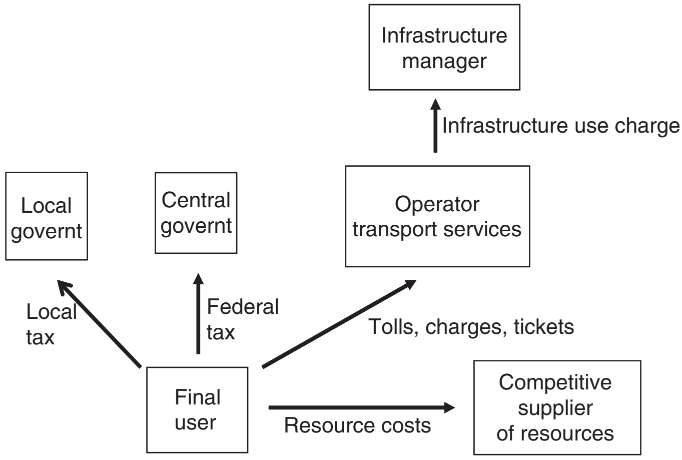

Fig. 5.6. Structure of Prices in Transport Module.

Table 5.4. Market Structure and Behaviour of Transport Service Suppliers.

Regime	Supply 1	Supply 2
Traditional private duopoly (two competing airlines, air versus private rail, …)	Profit max Nash behaviour in prices	Profit max Nash behaviour in prices
Mixed duopoly	Profit max Nash behaviour in prices	Welfare maximising government with Nash behaviour in prices
Welfare maximum	Welfare maximisation by jointly optimising prices of both transport service suppliers	

with the government as a Stackelberg leader, and Nash competition between two competing local governments.

Market structure can also be analysed for the infrastructure managers.[5]

In addition to the market structures for operation and infrastructure provision, it is possible to analyse vertical integration between service provision and infrastructure management, competition between integrated and non-integrated suppliers, and so on.

4. THE INVESTMENT MODULE

Section 3 focused on the equilibrium of the transport market in a single period. This section builds on Section 3 by considering investments in infrastructure capacity, the structure of the financial reporting and pooling of revenues into funds.

4.1. Investment Decisions and Physical Capacity

Each of the two infrastructure managers can make investment and maintenance decisions as well as pricing decisions. The analysis begins in a reference year, year 0, and investments are evaluated over a time horizon $[0, \ldots, T]$. Let $K(t)$ denote the physical capacity of infrastructure in period t for a given mode. And let $I(\tau)$, $\tau = -1, \ldots, -LT$ denote past investments in infrastructure, where LT stands for the age of the oldest capacity still operational in year 0, and $I(t)$ represents investments for $t = 0, \ldots, T$. The capacity in year t is then

$$K(t) = \sum_{\tau=-LT}^{t-1} (1-\delta)^{t-\tau} I(\tau)$$

where δ is the rate of depreciation.

Investments prior to the reference year are exogenous, whereas investments during the model period can be exogenous or endogenous. This modelling approach is a simple one in which capacity decays exponentially and there are no explicit maintenance decisions.

Many investment principles can be modelled in MOLINO. In the current version only two investment rules have been programmed:

A. *Exogenous investments*

This is useful to represent investments that have already been decided by a political authority. In our case studies (see Chapters 7–11 in this volume) most investments studied are of this type.

B. *Optimal investment rule when prices are optimised*

As explained in Chapter 2 in this volume, if prices have been set optimally (be it to maximise welfare or maximise profits), the optimal level of investment is determined by the condition that the marginal cost of investment matches the present-discounted value of reductions in user costs. There are two problems in modelling this endogenous investment rule. First, as investments affect all future periods, optimal investments require the simultaneous solution of the model for the whole horizon. Second, investments and pricing need to be solved simultaneously.

We therefore simplified the implementation of the endogenous investment rule by adopting two approximations. First, expectations are assumed to be myopic as discussed in Section 2.2. This means that, in order to compute the present-discounted future savings in users' costs of an additional investment now, we extrapolate the current savings using a growth rate specified by the modeller. Second, since we first optimise prices and keep these prices fixed in the current period, our procedure only holds for small investments (envelope theorem).

4.2. Cost of Capital and Interest Rates

4.2.1. Cost of Capital for Private Suppliers

For the investment decisions of private infrastructure managers we can use a cost of capital. The cost of capital is the discount rate at which the decision maker is ready to trade-off cash flows over time. Information on the cost of capital can be used in the investment decision rules of the private suppliers. An investment or maintenance decision is justified if the expected present-discounted cash flow of this investment is positive when the cash flows are discounted at the cost of capital. The cost of capital is a function of the risk class of the investment, and default estimates are taken from Chapter 4 in this volume.

4.2.2. Interest Rate for Governments

For governments, the interest rate is in principle the gross interest rate (before capital tax). Some sources also recommend the addition of a risk premium.

5. THE FINANCIAL REPORTING MODULE

Before we describe the accounting for infrastructure managers, it is useful to examine very briefly the flow of funds in MOLINO. The default flow is shown in Fig. 5.7; it can easily be adapted to the needs of a case study.

For the infrastructure managers, the accounting module provides inputs for the financial structure and this may determine the cost of capital and financial constraints for investments. For suppliers of transport services, an accounting module is necessary to compute break-even constraints. Since the equipment used by transport operators is not modelled, financial accounting is simpler for the transport operators than for the infrastructure managers.

The financial accounting information used for private companies[6] contains two parts: the income statement and the balance sheet. The income statement reports, for a given year, the main categories of expenditures and the main sources of revenues. The balance sheet reports, at the end of every year, the origins of the funds used (liabilities) and the uses of these funds (assets).

Important exogenous inputs for the financial reporting are:

- The financial structure: ratios of debt over equity, etc. at time $t = 0$.
- The financial policy: share of financial needs funded by new debt and new equity.
- The legal depreciation rates.
- The interest rate (a function of risk class and capital structure).
- The outputs of the investment module and the transport market module.
- Structure of subsidies from transport sector funds.
- The regulation scheme that may impose constraints on financial structure.

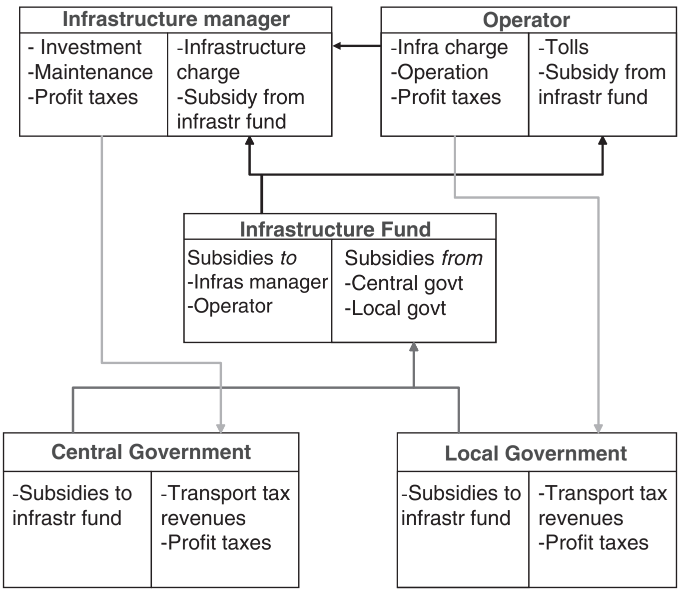

Fig. 5.7. Default Flow of Funds of MOLINO.

Important outputs of the financial reporting include:

- Subsidies needed or funds transferred to transport funds.
- Information on cost of capital for new investments.

Accounting rules and constraints are very case-study specific, and for that reason we do not provide here a default set of financial accounts for MOLINO.

6. WELFARE ASSESSMENT OF ALTERNATIVE REGULATION SCHEMES

Regulation schemes are assessed by analysing a present-discounted and weighted sum of benefits and costs. Table 5.5 describes the components of this objective function, their weighting within every period and their discount rate. The various symbols are defined in the text that follows.

This table is derived from the cost–benefit criteria described in Chapter 3 in this volume. Three factors make the cost benefit assessment complex.

1. Equity. The policy maker needs to weight the different benefits and costs by the relative weights given to the various population groups (here poor and rich income groups): the w_p, w_r coefficients. These coefficients affect all costs and benefits since every type of cost or benefit must be allocated to the different income groups. To allocate costs and benefits properly, it is necessary to determine who pays the residual tax payments $f(w_p, w_r)$. This function $f(\,)$ is a weighted average of w_p and w_r, where the weights are determined by the share of labour taxes paid by the poor and the rich. The same type of weighting function is used to take into account the distributional effect of who receives the profits of transport suppliers $w_{OP1}(w_p, w_r)$, etc., who benefits from lower freight costs $g(w_p, w_r)$, who bears the external costs $w_{EC}(w_p, w_r)$, etc.
2. Welfare effects in the rest of the economy. Some transport operations or infrastructure investments may be part of the public sector, and in this case taxes outside the transport sector must be adjusted to fund investments or provide subsidies for the transport sector. The welfare effects of these changes are captured in each period by the marginal cost of public funds of labour taxes Γ_{TL}, or an equivalent if other taxes are used to balance the budget (see Chapter 3 in this volume).
3. The discount rate. In order to compute present-discounted values, one needs to use an interest rate, net of capital tax, for households (r_n), a

Table 5.5. Components of the Welfare Assessment Function.

Component	Content	Weight within a Period	Intertemporal Discount Rate
Utility of "poor" household user of transport	Generalised consumers' surplus expressed in money equivalent	w_p	r_n
Utility of "rich" household user of transport	Generalised consumers' surplus expressed in money equivalent	w_r	r_n
Cost of local firms using freight	Generalised cost function of production	$g(w_p, w_r)$	r_f
Cost of foreign firms using transit freight	Generalised cost function of production	w_T	r_f
Tax revenue Central government	Net tax receipts including feedback effects on all taxes of transport infrastructure and pricing changes	$f(w_p, w_r)\,\Gamma_{TL}$	r_g
Tax revenue Local government	Net tax receipts including feedback on all taxes of effects of transport infrastructure and pricing changes	$l(w_p, w_r)\,\Gamma_{TL}$	r_g
Profit operator of infrastructure 1	Net revenues from tolls and user charges	$w_{OP1}(w_p, w_r)$ or $f(w_p, w_r)\,\Gamma_{TL}$	r_{OP1} or r_g
Profit operator of infrastructure 2	Net revenues from tolls and user charges	$w_{OP2}(w_p, w_r)$ or $f(w_p, w_r)\,\Gamma_{TL}$	r_{OP2} or r_g
Profit infrastructure Supplier 1	Net revenue of charges to operators after deduction of infrastructure investment and maintenance costs	$w_{IN1}(w_p, w_r)$ or $f(w_p, w_r)\,\Gamma_{TL}$	r_{IN1} or r_g
Profit infrastructure Supplier 2	Net revenue of charges to operators after deduction of infrastructure investment and maintenance costs	$w_{IN1}(w_p, w_r)$ or $f(w_p, w_r)\,\Gamma_{TL}$	r_{IN1} or r_g
External costs (other than congestion)	Air pollution, noise, accidents	$w_{EC}(w_p, w_r)$	r_n

gross interest rate for governments (r_g), an interest rate for firms using freight (r_f), an interest rate for providers of transport services that may be in the public sector (rate r_g) or the private sector (rate r_{OPi}, $i = 1,2$) and an interest rate for the infrastructure investments (r_g or r_{INi}, $i = 1,2$). We provide the option to choose different discount rates for different agent types. Liu (2003) advocates the use of different discount rates for consumers' surplus and government revenues. Adjustments for risk may also call for different discount rates.

When the components of this objective function are weighted differently, one can use the model to represent other objective functions to be optimised and in this way simulate the behaviour of other agents.

The behaviour of local governments can be analysed by using as objective function the weighted sum of the local user benefits (excluding transit by putting $w_T = 0$), the local net revenues from taxes and net income from local transport operations. To take a political economy perspective, one can analyse the impacts of policies on lobby groups such as a particular income group, a group of infrastructure suppliers, etc.

7. POSSIBLE REGULATION SCHEMES

Many regulation schemes can be envisaged. To assess them it is necessary to identify for each mode the infrastructure manager, the transport service operator and how the manager and operator take their decisions. This leads to seven questions:

1. Who decides levels of investment and maintenance of infrastructure?
2. Who executes the investment decisions (who builds with what type of contract)?
3. Who sets the charges paid by operators for the use of the infrastructure?
4. How are the deficits of the infrastructure managers financed (or who receives the surpluses)?
5. How is an operator that uses the infrastructure organised?
6. Who sets prices for the final users?
7. How are any deficits from operation financed?

As for the WHO questions, we need to distinguish at least the following four types of agents:

a. The central government (that takes into account the welfare of all citizens).
b. The local government (that disregards transit traffic, and may be concerned with only one of the two modes in a setting with two competing regions).
c. Private suppliers.
d. Competitive external suppliers of services (when competitive tendering is organised).

Table 5.6. Possible Regulation Schemes for One Mode.

Type	Investments				Operation			Objective
	Residual finance	Who decides investment	Who builds the infrastructure	Type of pricing infrastructure	Residual finance	Who sets prices	Service provider	
M11	Labour tax	CG	Public company (no tender)	Ad hoc	Labour tax	CG	Public comp (no tender)	Welfare max
M12	Labour tax	CG	Tender	MSCP	Labour tax	CG	Tender	MSCP
M13	Head tax	LG	Public company (no tender)	Ad hoc	Head tax	LG	Public comp (no tender)	Ad hoc
M14	Head tax	LG	Tender	MSCP	Head tax	LG	Tender	MSCP
M15	N/A	PS	Tender	Profit max	N/A	PS	Tender	Profit max

Abbrev.: CG, central government; LG, local government; PS, private supplier; MSCP, marginal social cost pricing; N/A, not applicable.

Naturally, the answers to the seven questions and the identities of the relevant agents will be case-study specific. One can envisage a few standard specifications for each mode. In Table 5.6, M1j refers to the jth specification for mode 1.

M11 is a combination that suffers from several inefficiencies since both infrastructure and operation are organised within the government, residual funding is via labour taxes and pricing is ad hoc rather than optimised. M12 is another polar case in which all elements are optimised, but it includes a distortionary labour tax for both investment and operation. M13 and M14 are run by the local government, which has fewer resources for funding and optimises only the welfare of its own citizens. M15 is a standard private case. Other cases such as public infrastructure and private operation can also be considered.

A complete regulation scheme specifies the structure of both the modes. For example, M15 + M25 is a wholly private scheme. M12 + M22 is a fully efficient scheme except for the funding by a labour tax. In addition, we need to specify whether there is any cross subsidisation between the modes, and if so whether this is mediated via an infrastructure fund with specific operating rules. Again, many variants can be imagined.

8. SOFTWARE IMPLEMENTATION OF MOLINO

A research version of MOLINO has been programmed in Mathematica 5.0 with input and output via Excel worksheets.[7] This version was used for the case studies in Chapters 10 and 11. For Chapter 9, a slightly different version of MOLINO was programmed. Finally, for Chapters 7 and 8 dedicated existing urban modelling software packages were used. After the case studies were performed, the MOLINO model was reprogrammed in a more user-friendly way and has been extended to take on board more general network structures as well as to take into account risk and uncertainty in the demand and cost parameters.

NOTES

1. By financial structure we mean a set of financial ratios for a firm that determine its capacity to attract capital and the cost of this capital. We elaborate somewhat on the financial structure in Section 5.

2. This means that the expectations of the different agents are not necessarily consistent.

3. Discrete choice functions would clearly be more appropriated better but there were no disaggregated representative samples available for any of the case studies.
4. Quality of service variables could be added in an extension of the model. Travel time is already included in the generalised price.
5. Not modelled in the current version of MOLINO. At this level one can make the same distinction between private or public duopolies as in Table 5.4. Competition between infrastructure managers can be assumed to take place in capacities where infrastructure managers anticipate the pricing game played later by the operators.
6. We use here as guide the international Anglo-Saxon tradition that is taught in international business schools. We do not do accounting in a narrow sense but rather use accounting information for business decisions. The schemes used in this section are described in Brealey and Myers (1996).
7. A full-scale application to the building, financing and pricing of a tunnel can be found in Proost, Van der Loo, de Palma and Lindsey (2005). A manual is available on-line.

ACKNOWLEDGEMENTS

Financial support from the European Commission (EU Fifth Framework research project REVENUE (DG TREN, Contract No. GMA2-2002-52011)) is gratefully acknowledged.

REFERENCES

Brealey, R., & Myers, S. (1996). *Principles of corporate finance*. New York: McGraw Hill.

Keller, W. J. (1976). A nested CES type utility function and its demand and price-index functions. *European Economic Review*, *7*(2), 175–186.

Liu, L. (2003). A marginal cost of funds approach to multi-period public project evaluation: Implications for the social discount rate. *Journal of Public Economics*, *87*, 1707–1718.

Proost, S., Van der Loo, S., de Palma, A., & Lindsey, R. (2005). A cost benefit analysis of tunnel investments and tolling alternatives in Antwerp. *European Transport/Transporti Europei*, *31*, 83–100.

This page intentionally left blank

PART II:
REAL WORLD EXPERIENCES

This page intentionally left blank

CHAPTER 6

THE ROLE OF THE CASE STUDIES AND THE MAIN POLICY QUESTIONS ADDRESSED

Heike Link

ABSTRACT

This chapter introduces Part II of the book. It provides an overview of the five case studies of revenue use in Europe that are presented in Chapters 7–11. The first two are urban case studies that deal with the use of toll revenues and investment in Oslo, and with the rejected congestion-charging proposal for Edinburgh. The other three case studies focus on investment funds in interurban road and rail transport in Germany, Switzerland and France. The main goal of this chapter is to describe the central policy questions addressed in the case studies.

One of the most widely debated contemporary transport policy questions is how to use the revenues from transport user charges efficiently while satisfying organisational and acceptability constraints. The question applies not only to existing user charging schemes but also to schemes that have been proposed in particular countries and schemes that are stipulated in EU policy directives and regulations (e.g. European Commission, 1998). There are a variety of established and proposed methods of using revenues including allocating revenues to the general budget, reducing labour taxes, earmarking

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 135–142

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19006-5

revenues for the transport sector, cross-subsidising other modes, and financing investment funds that pool revenues over regions, infrastructure types and modes. The EU White Paper "European Transport policy for 2010: Time to decide" (European Commission, 2001) calls for revenues either to be fed into national or regional funds to finance measures to reduce external costs (for a more detailed discussion of the underlying double dividend issue see Chapter 3 of this volume), or to be used to finance new infrastructure or alternative modes. Funding practises for both purposes exist in Switzerland, Germany and France. The revenues from the Swiss HGV (Heavy Goods Vehicles) charging scheme are channelled to fund the new Alpine rail tunnels. For the German eco-tax, which is imposed on all fuel consumption in Germany, revenues are used to reduce labour-related social insurance costs. In the case of the Zürich airport noise fund, which is financed from emission-dependent aircraft landing fees, the money is used to finance noise-protection measures and to compensate noise-affected populations. In France, an investment fund (AFITF) has been created to cross-finance new motorway projects using revenues derived from road charges on existing motorways.

Transport policy formulation and implementation must respect acceptability constraints to be feasible. Various studies (Jones, 1991; Bartley, 1995; Steg & Vlek, 1997; Link & Polak, 2003 inter alios) have determined that in addition to charge levels, the use made of the revenues is critical for making a scheme acceptable. Most of this research indicates that earmarking of revenues to the transport sector is necessary to make transport-pricing measures acceptable, even though earmarking may not be efficient.

Against this background, Part II of this book presents the findings from five case studies of using revenues from transport user charges in Europe. Each of the case studies deals with a specific problem of infrastructure investment and revenue use in a different country. The case studies are motivated by a range of factors: experience with established or newly implemented policies, policy proposals or ideas that are currently under discussion in the scientific community. All the case studies describe how revenues are currently used, identify what changes if any have been proposed in charging, investment and revenue use, and ask whether better alternatives exist.

Two case studies deal with revenue allocation from urban road charging schemes. The Oslo study (Chapter 7) describes the evolution of tolling policy in Oslo. It focuses on the policy shift from using revenues to finance roads specified in Oslo package 1 to the use of toll ring fees for public transport investments in Oslo package 2. The study analyses the welfare effects of this shift towards cross-subsidisation, and deals with regional distribution effects and inter-modal impacts. The Edinburgh study (Chapter 8) focuses on the unsuccessful proposal to introduce congestion charging in Edinburgh. It explores the welfare-distributional effects of the scheme on Edinburgh residents and non-residents, the merits of investing in particular tram lines and the general question whether revenues should be used to reduce general taxation or earmarked to the transport sector.

Three case studies deal with investment funds in interurban road and rail transport. The Swiss study (Chapter 9) investigates the efficiency and equity impacts of the railway investment fund FINÖV, which was created to finance the construction of new rail infrastructure with revenues from the distance-related heavy vehicle fee. The study deals specifically with the new Alpine rail tunnels where severe problems caused by the rapid growth of transit traffic have to be tackled. The case study assumes that the tunnels will be built and focuses on whether it is efficient to cross-finance the tunnels from the HGV charging revenues. The background of the German case study (Chapter 10) is the introduction of a distance-related motorway charge for HGVs in January 2005. Eighty per cent of the toll revenues are earmarked for investments in the transport sector, with shares to roads (50 per cent), rail (38 per cent) and inland waterways (12 per cent). A state-owned Infrastructure Agency (VIFG) is responsible for the allocation of revenues to specific projects. The study compares the welfare effects of this scheme and alternatives, and assesses acceptability issues such as a potential trade-off between charge level and use of revenues. The French case study (Chapter 11) deals with the welfare impacts of the French multi-modal investment agency AFITF, which was created to help finance major transport infrastructure investment projects. AFITF has an independent budget that is funded from motorway tolls, and contributions from Public–Private Partnerships, local governments and the European Union. The study assesses the efficiency of cross financing through AFITF from existing motorways to new motorway or rail projects as an alternative to direct public subsidies.

While each of the case studies considers different specific policy proposals and issues, they address a common set of general questions that can be summarised as follows:

What are the welfare effects of alternative user charging and revenue use schemes?

To take up this question, each study analyses the following schemes:

- A status-quo pricing scheme. In the interurban case studies this scheme is based on average cost pricing (average infrastructure costs in Germany and France, average infrastructure costs plus environmental and accident costs in Switzerland).

- Variants of the status-quo scheme that differ with respect to the range of infrastructure and set of vehicle types charged.
- Alternative schemes, which contain elements of marginal social cost pricing (MSCP). In addition to pure MSCP, the French study considers MSCP with mark-ups as a way to raise supplementary revenues to cross-subsidise rail investments.

Should the revenues from transport user charges be earmarked to the transport sector or should they feed into the general budget?

Earmarking is a controversial practice, as is evident from the debate over earmarking HGV charging revenues related to the amendment of the European HGV charging directive. Three of the status-quo charging schemes in the case studies feature some form of earmarking. The HGV charging revenues in Germany and Switzerland are earmarked for investments in the transport sector. And the revenues from the Oslo toll ring were earmarked to road construction under Oslo package 1, and are currently used to finance public transport investments under Oslo package 2. The main argument for earmarking is that it can help to prevent political abuse of funds (see Chapter 3). But earmarking creates inflexibility in the allocation of funds, and may result in excessive allotments to some modes at the expense of other modes. All the case studies explore the welfare impacts of earmarking revenues versus allocating them to the general budget.

Should user charge revenues collected from one mode be used to cross-subsidise investments in other modes?

A majority of existing and planned charging schemes not only earmark revenues to the transport sector as a whole, but also channel them into cross-subsidies to other modes. This is the case for the Oslo toll ring package, the German and Swiss HGV charging schemes and the French Lyon–Turin rail link. Just as earmarking is contentious as a general practice, so is earmarking in the form of cross-subsidisation. Attitudes vary.[1] Some studies report an increase in the acceptability of road user charges if the revenues are used to improve travel alternatives such as rail and urban public transport. Other studies reveal opposition to using revenues for any purpose other than for the mode on which the charges are levied. With this diverse and confusing set of results as a backdrop the Oslo, Swiss and German studies set out to assess attitudes towards cross-subsidisation in their specific case-study settings.

Should road user charges collected on one type of road be used to cross-subsidise other types of roads?

In parts of Europe, tolls are levied on some motorways but not on other roads. One question that rises in national transport policy formulation, as well as at the EU policy level, is whether surplus revenues from heavily travelled tolled motorways should be used to cross-subsidise roads that generate too little (or no) toll revenues to cover their costs. The French case study on motorways undertakes a welfare analysis of this issue. In some regions of Europe motorway tolls induce significant traffic diversion onto untolled alternative roads, which creates congestion and safety problems on the roads involved as well as noise and disruption in nearby towns. This raises another question whether toll revenues should be used to expand and improve the tolled motorways themselves, or whether the money should instead be spent to reduce bottlenecks on toll-free alternatives.

What are the implications of different ownership and procurement rules on pricing- and revenue-spending options?

There is growing interest both in the scientific community and at the political level in harnessing the private sector to provide transport infrastructure. Doing so will entail changes in infrastructure ownership, and implementation of procurement options such as tendering of services. Existing ownership and procurement options differ in their rules for charging, cost recovery and use of revenues. The rules affect the efficiency of infrastructure provision, maintenance and operation as well as the acceptability of the charging schemes.

Two of the case studies analyse the welfare effects of different regulatory frameworks (ownership, Public–Private Partnership schemes, procurement options) as well as the question to what extent they influence charging schemes and the use of revenues. The German tolling study includes an acceptability survey of German road haulage companies that explores to what extent overall acceptability of the HGV charging scheme is influenced by the institutional set-up for toll collection, the question who decides on revenue spending and specific options for revenue use. The French case study considers a scenario in which the private motorway operators are replaced by a public agency, and assesses what impact this institutional change would have on the charge level to be set, and the level of public subsidies necessary to break even (Fig. 1).

The case studies use a common methodological approach. Each constructs a set of three or more alternative schemes defined by combinations of pricing, revenue use and investment rules. See Chapter 5 of this volume for

	Scope	Pricing	Revenue use & financing	Invest-ment
Rules	*What sectors / sub-sectors are covered?*	*Which pricing rule?*	*What use of revenues, what financing?*	*Which invest-ment rule?*
Regulatory framework	*What actors are involved, with what functions?*	*Who sets prices?*	*Who decides on revenue use and financing?*	*Who makes investment decisions?*
Procurement & implementation	*Private or public provision?*	*Payment? Enforcement? Exceptions?*	*Revenue collection & management?*	*Tenders? Contracts?*

Fig. 1. Elements of Alternative Pricing, Procurement, Investment and Revenue use Schemes. *Source:* Suter et al. (2004).

more details on the way the different dimensions enter a case study in the MOLINO model. The set of schemes includes:

1. The status quo.
2. One or more proposed schemes if they exist.
3. A scheme that is either first best or theoretically superior to the status quo.
4. Additional schemes as applicable.

All the schemes are assessed with respect to their efficiency, equity, technical and organisational feasibility and acceptability, with an emphasis on efficiency and equity. Because the scope and nature of the problems investigated in the studies were driven by political proposals and discussions in the respective countries, the scenarios vary widely. To provide a common analytical basis, and to facilitate comparison of results, the MOLINO model described in Chapter 5 was used for the three interurban case studies.

Although the studies use a common methodological approach, they also differ in various ways. All studies are motivated by country-specific policy proposals and/or policy measures that have been approved for implementation. As a consequence, they differ with respect to the modes of transport involved and the complexity of the networks, the types of investment made, the institutional settings and the type of questions analysed. The results of the studies have to be interpreted in the case-specific contexts as defined by congestion levels, the marginal cost of public funds, public trust in local governments and so on. Rather than providing generalised and transferable recommendations, the case study results illustrate a range of possibilities that are useful to keep in mind when formulating policies for a new situation.

For example, the studies show that marginal social cost pricing may be inferior to alternative policies with higher prices if there is a significant premium from additional revenues because of a high marginal cost of public funds. Another finding is that the merits of earmarking depend very much on the benefits from allocating funds to the earmarked infrastructure vis à vis spending the money for maintenance or capacity enhancements in other modes or sectors. This is demonstrated by the results of the three interurban case studies. The Swiss and the French studies support cross-subsidisation from road to rail, while the German case study concludes that revenues should be retained within the road sector. These diverging results are attributable to differences between the countries in the road and rail networks as well as to differences in the exact questions that are asked.

A question of overriding importance in formulating charging and revenue-use schemes is the potential trade-off between efficiency and acceptability. This aspect is examined using attitudinal surveys in the Oslo and Edinburgh urban case studies as well as the German interurban study. Although the surveys covered different types of actors, and dealt with different types of road pricing, several common findings are apparent such as a wide acceptance of the user pays principle, a general support for earmarking of revenues, the need for policy packages as a combination of charging and revenue use, and the importance of public transport improvements to achieve acceptability of urban road pricing schemes.

Finally, all the case studies find that the welfare impacts of pricing, earmarking and investment decisions are strongly interrelated, and depend on the levels of taxes (in particular fuel taxes), on the configuration and quality of infrastructure and service provision in the competing modes, and on other local circumstances. The general lesson to draw from this is that pricing and revenue use packages must be tailored to the particular objectives and setting in question.

NOTE

1. See for example Jones (1991), Thorpe, Hills, and Jaensirisak (2000) and Link and Polak (2003).

REFERENCES

Bartley, B. (1995). Mobility impacts, reactions and opinions. Traffic demand management options in Europe: The MIRO project. *Traffic Engineering and Control*, *36*, 596–603.

European Commission (1998). *Fair payment for infrastructure use: A phased approach to a common transport infrastructure charging framework in the EU*. White Paper COM(98) 466 final, Brussels.

European Commission (2001). *European transport policy for 2010: Time to decide.* White Paper COM(2001), Brussels.

Jones, P. M. (1991). Gaining public support for road pricing through a package approach. *Traffic Engineering and Control*, *32*(4), 194–196.

Link, H., & Polak, J. (2003). The acceptability of transport pricing measures amongst public and professionals in Europe. *Transportation Research Record*, *1839*, 34–44.

Steg, L., & Vlek, C. (1997). The role of problem awareness in willingness-to-change car-use and in evaluating relevant policy measures. In: T. Rothengatter & E. C. Vaya (Eds), *Traffic and transport psychology*. Amsterdam: Pergamon.

Suter, S., Springer, U., de Palma, A., Lindsey, R., Ricci, A., Fagiani, P., Moilanen, P., van der Hoofd, M., Carmona, M., & Baker, J. (2004). *Case studies specification, REVENUE project deliverable 3*. Funded by 5th Framework RTD Programme, ISIS, Rome.

Thorpe, N., Hills, P., & Jaensirisak, S. (2000). Public attitudes to TDM measures: A comparative study. *Transport Policy*, *7*, 243–257.

CHAPTER 7

USE OF TOLL REVENUES AND INVESTMENT IN OSLO

Jon-Terje Bekken and Bård Norheim

ABSTRACT

This chapter investigates issues of efficiency and acceptability in the Oslo toll-financed investment packages. We show that the current and planned packages are far from optimal, due to restrictions on fares, revenue use and contents of the packages. These restrictions, however, seem to be important for political acceptability. The packages are also moving in the "right" direction. The main focus of the policy process in Oslo has been to find a compromise that is agreeable to all stakeholders, rather than to select the policy packages that are most economically efficient.

1. BACKGROUND

Norway has a long tradition of financing public road infrastructure through tolls, dating back to 1929. Traditionally, tolls were related to isolated projects, such as bridges and tunnels. However, with the introduction of the first European toll cordon around a city centre, in Bergen in 1986, the tide shifted. Today, three of the four largest cities in Norway, as well as several smaller cities, have toll cordons. More than one-third of the investment in

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 143–160

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19007-7

public roads is currently financed by tolls. Most of this comes from the urban toll cordons.

The Oslo toll ring (Oslo package 1) started in 1990. It was initially planned as an ordinary toll to finance tunnels under the city centre. However, before it was established, the municipality of Oslo joined forces with the neighbouring county, Akershus, and opted for a package to finance several other projects as well. Later in the process, it was also decided to earmark 20 per cent of the revenue for public transport infrastructure investments. The main reason for the scheme was the general lack of public funds for road investments. A new initiative was required to raise money for investments. Future use as a means for traffic restraint was also an open possibility, but congestion relief was not an objective per se. Political agreement for the Oslo package 1 was obtained, despite the lack of support from a majority of the population in the Oslo region.

A few years after the toll ring was introduced, and following extensive road investments in the region, there was a growing concern about car traffic increasing more rapidly than expected, as well as a lack of investments in public transport (PT) infrastructure. In 1996, the Norwegian parliament invited the local authorities in the Oslo region to develop an enforced PT plan based on national and local co-financing, to meet this challenge. This plan (Oslo package 2) was launched in 1998 and approved by Parliament and the local authorities in 2001.

Oslo package 2 is a supplement to the existing Oslo package 1 and consists of an increase in the toll of approximately €0.25 per trip, earmarked for PT infrastructure investments. In addition, the package includes an increase in the public transport fare of approximately €0.10 per trip, earmarked for rolling-stock investments. The planning of Oslo package 2 involved two counties and several authorities and organisations. Investment in PT was expected to double as a result. The main elements in the first four-year period (2002–2005) were railway investments (60 per cent of expenditures), a new metro ring (20 per cent), terminals/stations (11 per cent), and priority measures (9 per cent). The co-financing plan for Oslo package 2 also involved extraordinary national funding and public–private partnership funds from the redevelopment of the old Oslo airport. In Table 7.1, the status quo (ordinary funds) is compared with the extraordinary funds raised by Oslo package 2. The new funding scheme in Oslo has raised several new questions regarding the use of revenue. These questions concern cross-subsidisation from car to PT, regional distribution effects and inter-modal effects.

The Oslo toll ring was due to end in 2007. As the end of the toll ring came closer, two alternatives were examined. Either the toll ring could be

Table 7.1. Financing Plan for Oslo Package 2, 2002–2011.

	Billion Euro		
	Ordinary	Extraordinary	Total
Ordinary public funding (national and local)	0.95	–	0.95
Toll ring revenue	–	0.19	0.19
Increased public transport revenue	–	0.20	0.20
Public–Private-partnership-funding (private entrepreneurs)	–	0.02	0.02
Extra national funding	–	0.41	0.41
Additional financing requirement		0.12	0.12
Total	0.95	0.95	1.90

removed, as happened in Trondheim at the end of 2005, or a new toll scheme "Oslo package 3" could be introduced. The politicians opted for the latter. Because the planning started late, it was decided to continue the original toll ring ("Oslo package 1") until a new scheme was in place. The plans for Oslo package 3 were presented by a working group in May 2006. Most political parties have accepted the general concept of the package, but some of the contents are still under discussion. The final scheme is planned to be presented to Parliament in 2007. If passed by Parliament, the new scheme will result in 20 more years of urban tolling in Oslo.

The Norwegian toll-financing system is regulated through the Road Act of 1963 (§27) and subsequent regulations. The law clearly states that the purpose of tolls is to finance infrastructure investments, and that the toll period should be limited to a maximum of 15 years (20 years in certain cases). In 2001, a law on road pricing was passed by the Norwegian parliament (an amendment to the Road Traffic Act of 1965, §7.a), making this a possibility for Oslo package 3. The Norwegian law on road pricing states that road pricing is different from toll roads in the sense that the purpose of tolls is solely to finance investment, whereas road pricing is intended to internalise externalities (primarily congestion). Furthermore, the law also states that road-pricing revenue must be split between local authorities and the State, whereas revenue from tolls is intended for the local project/package only. The law also states that road tolls and road pricing cannot exist in the same area at the same time, and indirectly defines these as distinct tools. The usual definition of road pricing is that it is a generic term for charging for the use of roads. In that sense, congestion charging is a better word for the Norwegian law of road pricing. Throughout this chapter, we use road pricing as a generic term covering both tolls and congestion charging.

The overall goal of this study was to consider efficiency and acceptability issues within the current and planned Oslo packages, and how efficiency and acceptability considerations influenced the revenue use. Our expectation was that revenue use was a result of a political trade-off between these two desiderata. Section 2 outlines the model used to consider efficiency issues. Section 3 describes three simplified scenarios for the Oslo packages. The scenarios focus on different pricing levels for the toll ring and for PT, and different restrictions on revenue use for PT operations and investments. Section 4 discusses acceptability. Results of interviews and a Stated Preference (SP) survey are considered, together with the political process behind the packages. Section 5 draws conclusions.

2. THE MODEL

The FINMOD model is used to analyse the social costs and benefits of alternative scenarios. The model uses the number of vehicle kilometres in operation (revenue kilometres) as an indicator of the level of public transport service and it is implicitly assumed that any given number of revenue kilometres is optimally distributed between routes. Any change in the number of revenue kilometres supplied in an area will necessarily affect waiting times and/or walking distances for PT users, and may also affect on-board travel times and number of transfers. Thus, while passengers are not interested in aggregated revenue kilometres per se, this indicator is strongly correlated with variables that are more direct indicators of the level of service.

FINMOD is designed to identify optimal policies for PT at an aggregate level for regions or smaller local areas. As an aggregate model it does not feature individual routes, but rather uses aggregate measures of PT supply. FINMOD was first developed by Larsen (1993) to analyse optimal subsidies under various constraints. The structure of the model was inspired by Jansson (1979, 1984). Larsen and Østmoe (2001) used some aspects of FINMOD when they considered lessons learned from Oslo package 1. The model was further developed to investigate optimal incentives for PT contracts in different urban and regional areas (Johansen & Norheim, 1998; Carlquist, Larsen, Norheim, Hoelsæter, & Hagen, 1999; Johansen & Norheim, 2000; Fearnley, Bekken, & Norheim, 2002; Longva, Bekken, & Norheim, 2003; Fearnley, Bekken, & Norheim, 2004). Larsen (2004) and Larsen, Johansen, and Norheim (2001) provide a thorough mathematical description of FINMOD.

For the REVENUE project, the model was developed further (Tricker et al., 2006) to analyse the efficiency effects of different road-pricing schemes and budget constraints. In Norheim (2005) and Bekken and Norheim (2006), the same model is used to consider incentive-based subsidies between different modes and different regions.

2.1. Calibration and Segmentation

In the Oslo region there are five distinct PT modes: urban bus, metro, tram, regional bus and train. Each mode has its own cost and demand structures. Furthermore, the modes are controlled by different authorities with different budgets; the city of Oslo (tram, metro and urban bus), the county of Akershus (regional bus) and the Ministry of Transport and Communications (train). As a result, we have developed the model to consider these modes as distinct market segments serving different parts of the urban region. By separating the modes, the model is able to handle different restrictions on the allocation of funds between the modes and thus between the authorities.

To calibrate and develop the FINMOD model, extensive data sets from each PT operator were collected and cost models for each mode were developed (Bekken, 2004). Crowding on buses or trains can be included under the umbrella of "quality of service" in PT. Crowding is not treated explicitly in FINMOD. Instead, we assume that the level of crowding is kept constant, presumably at an optimum level, when demand and/or the supply of transport services change. To implement this, we fix the load factor at the reference level.

2.2. Two-step Optimisation

The optimisation model employs two steps. The first step estimates the mode choice between car and PT depending on car user cost, parking fees, urban density, PT service frequency and fares. The second step optimises the PT service under different financial constraints.

The main purpose of the model evaluation is to investigate the costs and benefits for the five PT modes in the different scenarios, and the benefit from a more flexible revenue use at the regional level. The use of revenue is studied under three different constraints on the use of revenues.

- Alternative I: Fixed subsidy level for each PT mode. The subsidy for each mode is fixed and cannot be transferred between modes.

- Alternative II: The total subsidy level for all modes is fixed, but the allocation between modes is unrestricted and chosen to maximise welfare.[1]
- Alternative III: Welfare optimal subsidy level without financial constraints. No restrictions on the level of subsidy for PT operation.

3. MODEL SCENARIOS

Based on the contents of the Oslo packages, we have developed three simplified scenarios. The scenarios focus on different pricing levels, both for the toll ring and for PT, as well as different restrictions on revenue use for PT operations and investments.

1. The Oslo package 1 has low tolls throughout the day, and no additional funding to cover PT operating costs. The model scenario for Oslo package 1 is optimisation with the financial constraint of a fixed subsidy level either for each mode or for the modes in aggregate. The subsidy level is fixed at 2004 values.
2. The Oslo package 2 has an additional toll earmarked for PT improvements, including operating costs. The model scenario for Oslo package 2 is optimisation with the financial constraint of a fixed subsidy level either for each mode or for the modes in aggregate, including the additional earmarked toll revenue.
3. The Oslo package 3 has time-differentiated marginal social cost pricing (MSCP) road tolls and no financial constraints. The tolls are applied on the same links as for the existing toll cordon.

Table 7.2 provides a brief overview of the scenarios. The first scenario is an approximation to Oslo package 1. The average toll is €1.00, with the

Table 7.2. Overview of Scenarios.

Scenario	Alternative	Toll Level	PT Fare Level	PT Operation Subsidies
Oslo package 1	I	€1 all day	€1.44	Fixed subsidy per mode
	II		€1.44	Fixed total subsidy
	III		MSCP	No budget restrictions
Oslo package 2	I	€1.25 all day	€1.44 (+€0.10)	Fixed subsidy per mode
	II		€1.44 (+€0.10)	Fixed total subsidy
	III		MSCP	No budget restrictions
Oslo package 3	III	MSCP	MSCP	No budget restrictions

highest price for a single ticket and the lowest for a season ticket. All revenue from the toll ring is used for road investment. In our model, as in real life, the revenue has already been used for road investments, so that the existing toll is actually used to repay loans.

The second scenario is the newly established Oslo package 2. The main element of Oslo package 2 is a PT investment plan financed by a €0.10 PT fare increase and a €0.25 toll increase. The PT fare increase is earmarked for PT rolling stock, and the toll increase (estimated to yield €30 million per year) is earmarked for infrastructure investment. In the model, the investment in PT infrastructure is treated by removing the capacity constraint on PT in the peak period. All the alternatives assume a marginal cost of public funds (MCPF) of 1.25. This is 5 per cent higher than the recommended value for cost-benefit analyses in Norway, because the budget constraints for local authorities are assumed to be tighter. The marginal external cost of a car trip is assumed to be €4.27 in the peak period before implementation of tolls, based on a calculation of congestion costs in Oslo made by Grue, Larsen, Tretvik, and Rekdal (1997). In the Oslo package 3 scenario we have used this estimate as an approximation of the toll to be imposed for MSCP on roads. MSCP calls for a 219 per cent increase in tolls, and a 55 per cent increase in average car user cost compared to the Oslo package 1 level. The actual marginal external cost of a car trip depends on the level of peak demand, and will fall when tolls are increased in Oslo packages 2 and 3.

The actual Oslo package 1 is the reference level for all scenarios, and the model estimates the changes in PT service provision and welfare level for different financial constraints described by Alternatives I–III.

3.1. Optimised PT Service Provision under Different Budget Constraints

Both the PT sector in general and PT in the Oslo region face budget restrictions. We consider two types of budget restrictions for Oslo packages 1 and 2. The first (Alternative I) allows no redistribution between the modes so that marginal social benefits are not equalised across PT modes, whereas the second (Alternative II) allows redistribution between the modes. In each case the total subsidy is held fixed at the reference level. In Oslo package 2, a 7 per cent fare increase is earmarked for PT service provision. The 25 per cent toll increase in Oslo package 2 will shift the demand for PT.

The results are reported in Table 7.3. The Oslo package 1 columns present the benefit of introducing MSCP for PT without increased tolls and under different budget constraints on revenue use. With fixed total subsidies

Table 7.3. Optimised PT Service Provision for Oslo Packages 1 and 2 with Fixed Subsidies Either Per Mode or in Total.

	Oslo Package 1 Reference Level	Oslo Package 1		Oslo Package 2	
		Fixed subsidies per mode (Alternative I)	Fixed total subsidies, reallocation between modes (Alternative II)	Fixed subsidies per mode (Alternative I)	Fixed total subsidies, reallocation between modes (Alternative II)
Fare level (€/trip)		Changes from reference level			
Peak	1.44	21%	21%	7%	7%
Off-peak	1.44	44%	43%	7%	7%
Network km (1000/h)					
Peak	1.94	145%	146%	73%	68%
Off-peak	1.64	13%	13%	−14%	−15%
Passenger capacity/vehicle					
Additional peak	144	−59%	−59%	−45%	−47%
Off-peak	144	−27%	−27%	2%	0%
Optimised number of PT trips					
Capacity peak trips	90.5	18%	18%	15%	15%
Non-capacity peak trips	34.1	4%	4%	15%	15%
Off-peak trips	75.3	−4%	−3%	2%	2%
Total number of trips	200.0	7%	8%	10%	10%
Cost and benefit (€ million/year)					
Cost of public funds (increase in subsidy)		0	0	0	0
Passenger benefit		85	85	107	108
External benefit		44	47	40	39
Total social benefit (€ million/year)		129	131	146	148

and no reallocation between modes, Oslo package 1 calls for increased fares of 21 per cent in the peak period and 44 per cent in the off peak. The increase is proportionally lower in the peak period because the benefits of reducing car trips are greater in the peak. The increased fare box revenue is used to increase service frequency, primarily in the peak (by 145 per cent). The increased frequency is achieved by running smaller vehicles. These changes result in an 18 per cent increase in capacity peak trips, a 7 per cent increase in overall trips, and a total social benefit of €129 million. This benefit can be regarded as the welfare loss from not implementing MSCP for PT.

Alternative II of Oslo package 1 allows reallocation of subsidies between PT modes. Surprisingly, this greater flexibility yields only a minor rise in the welfare gain from €129 million to €131 million.

The budget restrictions have similar effects in Oslo package 2, and the benefits from redistribution of funds between PT modes are again minor. But the optimised service provision is different because of the fixed fare increase earmarked for PT provision and the increased tolls. The increased fare box revenue is funding 73 per cent increased capacity peak frequency with smaller vehicles. This is significantly lower than the optimised frequency in Oslo package 1 scenario, but the demand for PT is higher due to the increased tolls. The total welfare gain of Oslo package 2 is €146 million with no reallocation of subsidies between modes, and €148 million with reallocation.

3.2. Optimised PT Service Provision and Marginal Social Cost Pricing

The second part of the analysis is to impose MSCP of PT without any constraints on budget reallocation or service level, and compare the outcome for each scenario with the actual Oslo package 1. The differences between the optimum and Oslo package 1 reflect the welfare loss from the initial funding model and organisational and financial constraints on PT service in the Oslo region.

The results are presented in Table 7.4. For Oslo package 1 the peak-period fare in the peak-capacity direction is reduced 21 per cent. The fare is decreased to encourage a reduction in car use, which is priced well below marginal social cost. By contrast, off-peak fares are increased marginally because the external costs of auto travel are lower during the off-peak, and the benefits of reducing auto trips are outweighed by the benefits of funding increased off-peak frequency.

Table 7.4. Optimised PT Service Provision with no Budget Restrictions for Oslo Packages 1, 2 and 3.

	Oslo Package 1 Reference Level	Oslo Package 1	Oslo Package 2	Oslo Package 3
Fare level (€/trip)		Changes from reference level		
Peak	1.44	−21%	−14%	62%
Off-peak	1.44	2%	3%	−3%
Network km (1000/h)				
Peak	1.94	165%	169%	200%
Off-peak	1.64	30%	34%	30%
Passenger capacity/vehicle				
Additional peak	144	−60%	−61%	−54%
Off-peak	144	−25%	−24%	−24%
Optimised number of PT trips				
Capacity peak trips	90.5	28%	30%	52%
Non-capacity peak trips	34.1	20%	24%	23%
Off-peak trips	75.3	14%	19%	16%
Total number of trips	200.0	21%	25%	33%
Cost and benefit (€ million/year)				
Cost of public funds		−29	−26	0
(increase in subsidy)		(115)	(105)	(1)
Passenger benefit		226	261	255
External benefit		62	64	68
Total benefit (€ million/year)		259	299	323

Service frequency is increased by a whopping 165 per cent during the peak, and a much smaller – but still significant – 30 per cent off-peak.[2] But the large increases in frequency are largely offset by running smaller vehicles, and the number of PT trips increases only moderately. This result is consistent with those of other studies that suggest bus size is often too large (Carlquist et al., 1999; Larsen, 2004).

As a consequence of the increase in service provision, and the reduction in peak-period fares, the PT subsidy required rises by €115 million. Nevertheless, a large welfare gain of €259 million results. This can be regarded as the welfare loss from not introducing MSCP of PT under Oslo package 1.

In the Oslo package 2 scenario, tolls are slightly increased and in Oslo package 3 they are raised to MSCP levels. The toll increases influence the optimised PT service provision in two ways: first by increasing PT demand, and second by reducing the external benefits of reducing auto trips. PT is adjusted by increasing frequency to meet the demand, and increasing fares to finance peak capacity trips.

In the Oslo package 2 scenario, the peak fare reduction of 14 per cent is smaller than the reduction in Oslo package 1, while the service frequency is slightly larger. The number of PT trips increases by 25 per cent compared to 21 per cent, and the annual welfare gain is €299 million compared to €259 million. All these differences are attributable to the higher tolls in Oslo package 2.

As Oslo package 3 is still under political discussion, and not yet clearly defined, our scenario may deviate from the actual Oslo package 3 in a number of respects. The benefit from this scheme is largely due to the funds made available to the PT sector for service improvements. The congestion-charging scheme also removes the problem of under-priced car traffic. As a consequence of these two factors peak-period fares are raised by a large 62 per cent and peak-period frequency is raised (200 per cent) by more than in Oslo packages 1 and 2. A similar outcome for PT service under MSCP was obtained in the Brussels case study for the MC-ICAM project (Proost & Sen, 2006). Both the increased toll and the increased fare provide new funds to the PT sector, making the optimisation viable without increasing subsidies. However, the annual welfare gain of €323 million is only moderately higher than the gain from Oslo package 2.

4. ACCEPTABILITY OF THE OSLO PACKAGES

There has been no referendum for the Oslo packages. Thus, it is important to understand how the politicians and the administrative bodies actually consider and weight the different aspects of the schemes against each other. For this purpose, both interviews (Bekken & Osland, 2004) and a SP survey (Nossum & Norheim, 2004) were conducted. The SP survey was carried out among politicians and transport planners with regard to different transport-funding schemes and use of revenue. As discussed below the results show that it has been considered more important to find packages that are acceptable to all parties than to select an optimal package. Yearly attitudinal surveys of the general population were also carried out (Prosam, 2004). These surveys provide useful insights into aspects of general acceptability of the packages and how attitudes have evolved over time.

4.1. The Political and Administrative Process Behind the Oslo Packages

Bekken and Osland (2004) investigated the political and administrative processes leading up to the Oslo packages. The study was carried out as part of a research project for the Ministry of Transportation and Communications and ran parallel with the REVENUE project. The main purpose was to understand how the Oslo packages were made possible and how compromises were fashioned. The study shows that this was accomplished through negotiations between stakeholders. Three important elements in that respect were earmarking some of the revenue for "high-profile" investments, low fare levels with large discounts for heavy users and no time variation in the toll levels.

4.2. General Attitudinal Surveys of Citizens

The toll ring in Oslo commenced operation in February 1990. Each year since 1989 a survey of attitudes towards the toll ring has been carried out among the citizens in Oslo and Akershus. The sample is randomly selected among the population, with roughly 1,000 interviews carried out each time by telephone. The aim has been to track changes in attitudes over time. The result is a time series of attitudes covering a period of 16 years. Fig. 7.1 summarizes the general results from the survey both for the entire sample and for the part of the sample passing through the toll ring on the way to work. Respondents were asked whether they were positive, indifferent or negative to this way of collecting revenue.

Fig. 7.1 shows that there is no overwhelming public support for the packages. Even though this survey cannot be compared with the result from a potential referendum, it is fair to say that the schemes would have a hard time being accepted in a general referendum. Acceptance has, however, increased over time since each scheme was introduced. This was also apparent in the Stockholm congestion charging trial scheme, where the public turned more positive after it was introduced (Gustavsson, 2006). The introduction of Oslo package 2 in 2001, and the corresponding fare hike, reduced acceptability. However, after few years acceptability was back to the pre-Oslo package 2 levels.

Since 2001, the survey has included a question on attitudes towards Oslo package 2. About two-thirds express a positive attitude towards Oslo package 2 after being informed about the contents of the package. Close to one-third agree with a proposal that half the revenue be used for PT investments

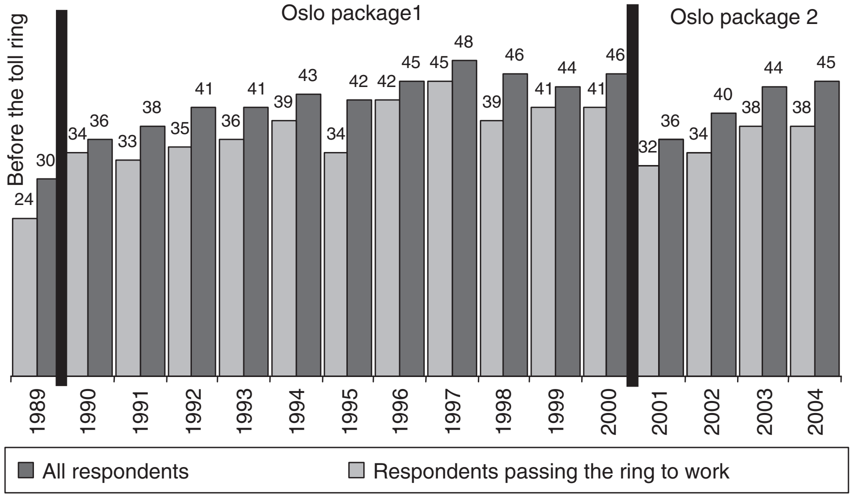

Fig. 7.1. Positive Attitudes Towards the Oslo Toll Ring (Percentage of Sample) *Source:* Prosam (2004).

(the PT share of Oslo package 1 is 20 per cent). Respondents are split equally on whether tolling should be continued regardless of how the funds are spent. However, the increased use of revenues for PT and road investments is the main reason why the public accepts a prolongation of the toll ring (Prosam, 2004).

4.3. Survey of Preferences among Decision Makers

A preference survey of local decision makers in Oslo and some other large urban areas in Norway (Bergen, Trondheim, Stavanger and Kristiansand) was carried out in the autumn of 2004 (Nossum & Norheim, 2004). The survey was part of the project "Alternative financing schemes for urban transport" financed by the Ministry of Transportation and Communications. The survey was also a part of the Oslo case study of the REVENUE project. In this study, decision makers are defined as those who take part in the decision-making process in a wider context, i.e. county and municipal politicians, administrative staff in counties and municipalities and road authorities.

The SP method was used to assess the decision makers' preferences for alternative packages of restrictive measures, pricing policy and revenue-use options. Tailor-made questionnaires were developed for each decision maker and region with the aid of the computer software SAWTOOTH. The survey was carried out as a self-administered Internet survey to enable customisation. It is not possible to combine this SP data set with Revealed Preference data, primarily because the choice situation is much more complex in the real world than this SP-survey can investigate.

The survey revealed strong support for a combined funding approach with contributions from toll revenues, transit fare revenues and local and national authorities. Close to 80 per cent of decision makers favoured the idea of a joint contribution between authorities and passengers to finance better PT service. Toll ring revenue and congestion charging are both preferred to local taxes as funding sources.

One goal of the SP survey was to reveal the different stakeholders' inclination to recommend positive measures (increased PT frequency, reduced PT fares and car-free city centres), restrictive measures (increased road tolls, congestion charging, reduced number of parking spaces and increased parking fees) and combinations of these measures. There was also interest in determining the stakeholders' assessment of the effects of the measures.

The surveys identified a strong positive correlation between the expected efficiency of the measures and acceptability, except for car-free cities. The politicians consider car-free city centres as a very efficient measure for reducing car traffic, but they do not recommend it. All respondents expect the levels of restrictive measures to influence positively the impact of the measures, but the probability that the respondents would recommend different measures was independent of this level. This may explain why efficient measures have increasingly been implemented in the Oslo packages, but less intensively than would be optimal.

The attitudinal and preference surveys help to explain some of the characteristics of the Oslo packages. Currently, the toll ring scheme provides a lion's share of funding, with an increasing share of the funds being earmarked for PT. This is one step in the right direction from an economic point of view, but it is also rational, based on the attitudes of the politicians and the professionals in the field. There is little support for funding with other local taxation.

5. CONCLUSIONS

For a long time, economists have argued in favour of congestion charging in cities with heavy peak traffic. It should come as no surprise that this is also one of the main conclusions of this study. Our model also shows that because car traffic is under-priced during peak periods, PT also has to be under-priced during peak periods as a second-best solution. The result is a choice between an unnecessarily high level of subsidies for PT, and a level of PT provision which is too low. It would be better to price car traffic at marginal social costs. This would also make it possible to improve PT services substantially and/or to reduce the subsidies because there would be no reason to under-price PT. With regard to revenue use, we find that a more efficient allocation of funds between different modes would also improve efficiency, but only to a small extent. All these results support conventional economic perspectives. So why is it that they have not been implemented?

In light of the failure to implement road pricing in Edinburgh, and the general lack of success at implementing true congestion charging schemes in urban areas, it is clear that pure economic arguments are not enough. We live in a democracy, where economists and car drivers have the same right to be heard in general elections. The Norwegian urban toll schemes have developed with this as a clear prerequisite. This makes it necessary, in relation to model simulations of optimal policies, also to put the Oslo toll schemes in a political and organisational context, and consider the acceptability of the schemes discussed.

The main objective of the model scenarios has been to investigate the interaction between different toll-ring regimes, budget constraints and PT optimisation. The toll regimes are called Oslo packages 1, 2 and 3, and they define the financial framework for these packages. It is important to view these in light of the acceptability of the Oslo packages.

First, our analysis shows that the tolls and PT fares of Oslo packages 1 and 2 are far from optimal. The tolls for the toll ring are set at a politically "acceptable" level that only generates enough funds for some particular investments. The increase in fare from Oslo package 1 to Oslo package 2 is enough to finance a certain amount of rolling stock investments for each PT mode. Our analysis also shows that there are positive marginal benefits from increased subsidies. The benefits depend on whether funds are reallocated between PT modes. Earmarking reduces social surplus, but not by much since the existing revenue allocation is close to optimal.

Second, we showed that Oslo package 2 is a small step in the right direction from an economic perspective. Due to additional funds from the increased toll it is possible to improve the level of PT service with a lower fare increase than in Oslo package 1.

Third, the model evaluations for Oslo package 3 indicate that congestion charging yields appreciable benefits by internalising the external costs of car traffic and by raising enough funds to eliminate the need for more PT subsidies. However, the benefits from shifting the toll ring featured in Oslo packages 1 and 2 to congestion charging in Oslo package 3 are not as large as the benefits from optimising PT. This is attributable to the relatively low congestion levels in Oslo and a PT system that is far from optimal.

Finally, the role of earmarking in enhancing acceptability of transport pricing reform must be taken into account. Earmarking is regarded as a sub-optimal approach from an economic perspective, but it is an integral part of most urban road-pricing schemes. The evaluation of the process behind the packages (Bekken & Osland, 2004) showed that a degree of earmarking up front has been important to making the Oslo packages politically viable. In particular, having funds earmarked for PT and for use within the different regions, seems to be important. The main focus of the policy process in Oslo has been to find a compromise that is agreeable to all stakeholders, rather than to select policy packages that are economically efficient. The result has been that all stakeholders have been better off compared to a situation without the packages. Furthermore, some high-profile investments (such as the metro ring) have been included to sweeten the pill. Yet, surveys of the general public indicate that the Oslo packages would be turned down in a referendum. The preference survey of decision makers also supports this view. These findings highlight the challenges of implementing urban road-pricing schemes in democracies.

NOTES

1. Redistribution of budgets between Oslo and Akershus is not as realistic as reallocations within the regions; i.e. between tram, metro and local buses in Oslo, and between train and regional buses in Akershus.
2. The increase in frequency is not surprising given the relatively low overall frequency in the region compared to cities in other countries. The metro, for instance, runs with 15-min headways frequency during peak periods, and 30-min headways during evenings. The metro is also the mode of transport with the highest increase in frequency in our analysis.

ACKNOWLEDGEMENTS

Financial support from the European Commission (EU Fifth Framework research project REVENUE (DG TREN, Contract No. GMA2-2002-52011)) is gratefully acknowledged. The Norwegian Ministry of

Transportation and Communications has supported several parallel projects, providing important input for the REVENUE project.

REFERENCES

Bekken, J.-T. (2004). *Finmod—an aggregated cost model for Norwegian public transport*. Oslo: Institute of Transport Economics.

Bekken, J.-T., & Norheim, B. (2006). *Optimal subsidies for Norwegian urban public transport*. Oslo: Institute of Transport Economics.

Bekken, J.-T., & Osland, O. (2004). *"An offer you can't refuse…" Introduction to the Norwegian urban toll cordon tradition. Alternative financing schemes for urban transport – Part 1*. Oslo: Institute of Transport Economics.

Carlquist, E., Larsen, O. I., Norheim, B., Hoelsæter, A., & Hagen, T. (1999). *Quality contracts in Hordaland County – Discussion of alternative contract forms*. Oslo: Institute of Transport Economics.

Fearnley, N., Bekken, J.-T., & Norheim, B. (2002). *Developing quality contracts for the Norwegian state railways' intercity market*. Oslo: Institute of Transport Economics.

Fearnley, N., Bekken, J.-T., & Norheim, B. (2004). Optimal performance-based subsidies in Norwegian intercity rail transport. *International Journal of Transport Management*, *2*(1), 29–38.

Gustavsson, H.-Å. (2006). *Kunskapen om och attityder till Stockholmsförsöket*. Utrednings- och statistikkontoret. Stockholm: Stockholm city.

Grue, B., Larsen, O. I., Tretvik, T., & Rekdal, J. (1997). *Køkostnader og køprising i bytrafikk*. Oslo: Institute of Transport Economics.

Jansson, J. O. (1979). Marginal cost prising of scheduled transport services: A development and generalisation of turvey and mohring's theory of optimal bus fares. *Journal of Transport Economics and Policy*, *13*(3), 268–294.

Jansson, J. O. (1984). *Transport system optimization and pricing*. Chichester: Wiley.

Johansen, K. W., & Norheim, B. (1998). *Quality contracts – alternative to competitive tendering for public transport in Norway*. Oslo: Institute of Transport Economics.

Johansen, K. W., & Norheim, B. (2000). *Output-based funding of urban public transport. Model simulation of alternative incentives and financial framework*. Oslo: Institute of Transport Economics.

Larsen, O. I. (1993). *Samfunnsnytte av tilskudd til kollektivtrafikk*. Oslo: Institute of Transport Economics.

Larsen, O. I. (2004). *The social optimum public transport model (SOPTRAM)*. Oslo: Institute of Transport Economics.

Larsen, O. I., Johansen, K. W., & Norheim, B. (2001). Towards achievement of both allocative efficiency and X-efficiency in public transport. *Journal of Transport Economics and Policy*, *35*(3), 491–511.

Larsen, O. I., & Østmoe, K. (2001). The experience of urban toll cordons in Norway – lessons for the future. *Journal of Transport Economics and Policy*, *35*(3), 457–471.

Longva, F., Bekken, J.-T., & Norheim, B. (2003). *New contractual arrangements for the purchase of public transport in Telemark*. Oslo: Institute of Transport Economics.

Norheim, B. (2005). *Cost-benefit analysis of alternative public transport funding in four Norwegian cities. Alternative financing schemes for urban transport – Part 4*. Oslo: Institute of Transport Economics.

Nossum, Å., & Norheim, B. (2004). *Stated preference study among local decision makers in the transport and communications sector. Alternative financing schemes for urban transport – Part 2*. Oslo: Institute of Transport Economics.

Proost, S., & Sen, A. (2006). Urban transport pricing reform with two levels of government: A case study of Brussels. *Transport Policy*, *13*(2), 127–139.

Prosam. (2004). *Holdningsundersøkelse for bomringen og Oslopakke 2 1989-2003*. Prosam report 109. Oslo: Prosam.

Tricker, R., Fereday, D., Pickup, L., Norheim, B., Bekken, J.-T., Shepherd, S. P., Laird, J. J., Nash, C. A., & Suchorzewski, W. (2006). *Report on the implementation of urban case studies*. REVENUE Project Deliverable 5. Funded by 5th Framework RTD Programme. ISIS, Rome, 16th January 2006.

CHAPTER 8

CORDON CHARGES AND THE USE OF REVENUE: A CASE STUDY OF EDINBURGH

James Laird, Chris Nash and Simon Shepherd

ABSTRACT

This case study examines the proposed Edinburgh cordon charge scheme, which – despite earmarking revenues for use in the transport sector – was still defeated at a referendum. This research suggests that whilst earmarking revenues reduced the efficiency of the scheme compared to a first best scenario, with lower toll charges than optimal, the scheme still offered substantial net benefits. Also whilst the City of Edinburgh Council had an incentive to charge residents of other authorities to favour its own residents, the proposed revenue distribution was equitable. However, the scheme still proved controversial; its defeat shows the importance of carrying such schemes forward on a regional basis with consensus between the local authorities involved.

1. INTRODUCTION

The widely perceived success of the London congestion pricing scheme led many commentators to expect other British cities to follow suit. Legislation

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 161–187

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19008-9

permits other cities to introduce road pricing schemes, although they require approval of a higher authority (the Department for Transport in the case of English or Welsh cities and the Scottish Executive in the case of Scottish) to implement their proposals. However, in a local referendum in February 2005, the citizens of Edinburgh rejected the congestion charging scheme proposed by the City of Edinburgh Council (CEC). The scheme would have consisted of an inner and outer cordon in and around the city of Edinburgh plus an associated transport investment package. The revenues derived from the congestion charge were to be earmarked for re-investment in the local transport system.

Two major issues concerning urban road pricing arise in this case study. The first is the issue of earmarking of revenues. On the one hand, there is considerable evidence that the acceptability of urban road pricing is enhanced if the revenue is earmarked for local transport projects (Ison, 2000). On the other hand, there is concern that this might not be the most efficient use of the revenue raised (Bös, 2000). The second concerns the jurisdiction of the authority raising the charge. A key issue in the Edinburgh scheme was the fact that the proposed outer cordon would largely charge residents of local authorities outside Edinburgh to enter Edinburgh. It has previously been argued (Proost & Sen, 2006) that in such a situation the city authority has an incentive to charge too high a price, in order to benefit its own residents at the expense of those of the surrounding area. This chapter examines the Edinburgh experience, using a strategic modelling approach to review the impact of the level of and proposed use of revenue from the cordon charges and compare against alternatives that would maximise welfare for all residents in the region and for Edinburgh residents alone. The results are used to discuss the effect of alternative levels of charge and revenue uses together with the proposed transport investment packages on the level and distribution of benefits.

As explained above, despite the fact that the charge was earmarked for local transport projects and that a substantial proportion of the revenue raised would come from residents of the surrounding area, the citizens of Edinburgh overwhelmingly rejected the charge in the referendum. A parallel piece of work was undertaken by Transport and Travel Research (Tricker et al., 2006) in which a number of interviews with key players and stakeholders were undertaken to try to understand the lack of acceptability of the proposal.

The following section provides the background to the Edinburgh proposals. We then explain the modelling approach and outline the results of the modelling work. Finally, we consider the issue of acceptability, before drawing our conclusions.

2. THE EDINBURGH CONGESTION CHARGE

Edinburgh is the capital of Scotland, with a population of 453,400 (in the year 2000), which increases to around one million when including the surrounding "travel to work" area. The city's economy is based on a modern service industry. There is a large knowledge economy, with three universities and a number of further education institutions. A background of high economic growth, rising car ownership and increased traffic and congestion combined with the need to maintain accessibility and protect the historic heart of Edinburgh led to the development of an integrated transport initiative. At the centre of the initiative was a congestion charge which would reduce private road traffic and help fund major public transport infrastructure projects.

CEC proposed a double cordon charging structure with a £2 charge on vehicles in-bound to Edinburgh (see Box 1 for further details). A cordon charge system was preferred by the CEC to an area charge as implemented in London for a mixture of administrative and political reasons:

> The introduction of [an area licensing scheme or zone scheme] would mean that those residents between the inner and outer cordons would have to pay the congestion charge when making even a relatively short journey from home. This would give rise to an array of social problems including severance of communities, technical and environmental difficulties consequent on the proliferation of charging points, and the substantial additional cost of handling the exemptions which would inevitably be claimed by residents in the affected areas.…A cordon system is better suited to Edinburgh than the area based scheme favoured in London. That scheme has been relatively costly to introduce, administer and enforce; and there would have been inevitable difficulties in securing public support without the granting of numerous resident exemptions. (City of Edinburgh Council, 2004, p. 24)

The inner cordon is located around the city centre (locations of the cordons are detailed in Fig. 8.1). The Road User Charging (RUC) legislation in Scotland requires all revenue to be earmarked for the transport sector. Additionally, the responsibility for setting the charge, collecting the revenue and overseeing revenue use has to lie with a single local authority – which in this case was proposed to be the CEC. In 2003, CEC formed a wholly owned arms-length company, Transport Initiatives Edinburgh (TIE), to deliver major transport projects. TIE would have been responsible for administering the congestion charge and overseeing the investment of the revenue. CEC also proposed a revenue sharing scheme in which the net revenue[1] from the congestion charge would be used for projects that would benefit residents of local authorities in proportion to the trip origins of those paying

Box 1. Edinburgh Congestion Charge Structure..

THE PROPOSED SCHEME

- Charging from Monday to Friday only (No charge at weekends or public holidays).
- Two charging cordons:
  - City centre cordon operating from 7am to 6.30pm.
  - Outer cordon inside city bypass operating from 7.00 to 10.00 am.
- £2 charge, only one charge each day, no matter how many times you cross either cordon.
- After introduction in 2006, the charge would be linked to inflation.
- Charge would only apply to vehicles entering the city. No charge would be made for crossing either cordon on trips heading out of the city.
- Drivers would be able to choose from a wide range of methods to pay the charge: ticket machines, internet, mobile and payment at shops. Payments could be made on a daily, weekly, monthly or annual basis.

EXEMPTIONS:

- Emergency vehicles, motorcycles, all taxis licensed under the Civic Government (Scotland) Act 1982, all buses and coaches including taxibuses and vehicles used for the transport of disabled people, blue badge holders, purpose built breakdown vehicles operated by accredited breakdown and recovery organisations and registered car club vehicles will be exempt.
- Residents of Edinburgh, living outside the outer cordon (including Currie, Balerno, Juniper Green, Ratho, South Queensferry, Kirkliston) would be exempt from paying the charge at the outer cordon.

The scheme will operate for 20 years from the actual date on which charging starts.

On any charging day, details of vehicle registration numbers for which a charge has been paid will be held on a database. This will be compared with vehicles identified at the charging cordons as liable for the charge. A penalty charge would be payable if the standard charge had not been paid by midnight on the day cordons are crossed. The penalty charge is proposed to be the same as a parking penalty charge – currently £60.00 with a 50 per cent reduction for payment within 14 days, rising to £90.00 if the penalty is not paid after 28 days.

Source: TIE website http://iti.tiedinburgh.co.uk/ (2 September 2004).

Fig. 8.1. Inner and Outer Cordon Boundaries. *Source:* TIE (2004).

the congestion charge. This would have seen 46 per cent of the net revenue being invested in transport projects to benefit South East Scotland Transport (SESTRAN) authorities[2] other than CEC.

CEC also outlined a proposed investment package that would be associated with the congestion charge. To differentiate transport investments arising from the congestion charge from those that would occur using funding from existing sources, CEC set out two transport investment packages for use in the public consultation, the public inquiry and the referendum. A *Base* scenario that would occur irrespective of whether or not a congestion charge went ahead and a *Base+Additional* scenario that was contingent on funding from the congestion charge. In March 2003, four months after the Scottish Executive had given approval in principle to the congestion charge and its associated *Additional* investment package, the Scottish Executive decided it would provide up to £375 million of public

Table 8.1. Revenue Allocation by Package and Area.

Package	Area		
	Edinburgh (million)	SESTRAN authorities other than Edinburgh (million)	Total (million)
Base investment package (public sector funding unconditional on the congestion charge)	£422	£56	£478
Additional investment package (funded by surplus revenue from the congestion charge)	£410	£351	£761
Total	£832	£407	£1239

Note: 2002 prices.
Source: TIE (2004) "Statement of case".

funding for a tram scheme in Edinburgh. This funding was not contingent on the congestion charge. The effect of this decision was to move the two highest profile transport infrastructure projects from the *Additional* investment package to the *Base* investment package. Table 8.1 sets out the resultant funding totals for the different packages by area and funding source.

The congestion charging proposals were developed by CEC over a 12-year period that followed the statutory planning requirements in Scotland. The proposals first entered the local plans in 1993 and the referendum represented the sixth phase of public consultation. The first phase of consultation occurred in 1999 when a questionnaire was distributed throughout Edinburgh from which 19,000 responses were obtained with a net support[3] for congestion charging of almost 40 per cent. By February 2005, the local referendum of residents of Edinburgh demonstrated that, support had completely eroded to a net opposition of almost 50 per cent.

3. THE MODELLING APPROACH

The primary research question addressed by the modelling part of our study of Edinburgh is: *What would a system of charging and revenue sharing that maximised social welfare within the local political and institutional constraints look like for Edinburgh?* This question is bound up with the related issues regarding: who should set the cordon charges; who should collect the revenue; and who should decide how the revenue should be spent. In this study we therefore consider two situations: (i) where a "high" authority, such as some sort of partnership between local authorities or the Scottish Executive, is the decision maker and acts so as to maximise the welfare of society as a whole; and (ii) where the CEC is the unconstrained decision maker and is free to do as it wishes so as to maximise the welfare of Edinburgh residents. Within these investigations we implicitly tackle the issue of revenue use and look at how the solutions differ if revenue use is earmarked for the transport sector (proportional to residency of those who pay the congestion charge) or revenue is used to reduce general taxation. We also consider earmarking for the transport sector as a whole and geographically.

3.1. The MARS Model

To tackle these issues, we apply the strategic land use[4] transport interaction model MARS (Metropolitan Activity Relocation Simulator) (Pfaffenbichler, 2003; Pfaffenbichler & Shepherd, 2003) to the city of Edinburgh and the surrounding region. Fig. 8.2 illustrates the MARS study area and zoning

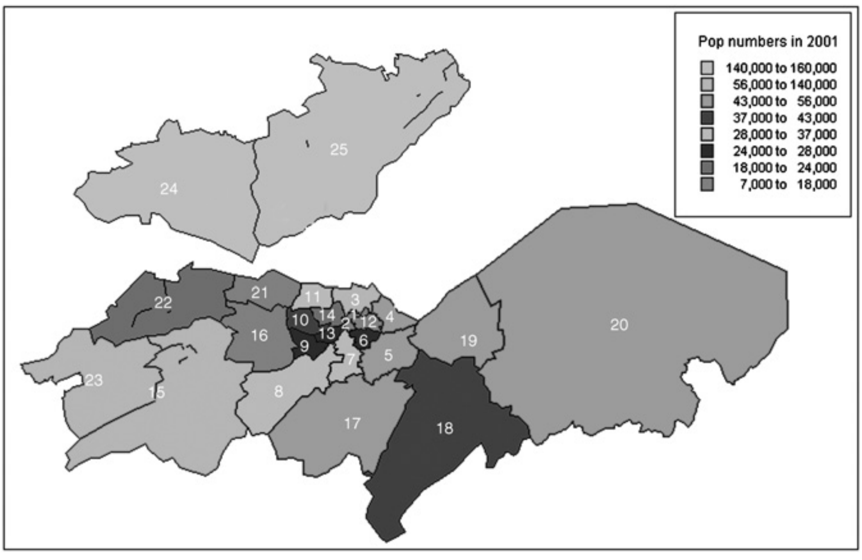

Fig. 8.2. MARS Model Study Area, Zoning System and Population Distribution in Year 2001.

system. Zones 1–14 represent the CEC area (as also shown in Fig. 8.1), whilst the zones to the east represent the East Lothian local authority area, to the south the Mid-Lothian authority, to the west the West Lothian authority and to the north the Fife authority.

MARS is a strategic, interactive land-use and transport interaction (LUTI) model. It was developed as a time-saving alternative to traditional four-step transport models.

The first stage of the MARS development was a qualitative analysis using causal loop diagramming (CLD).[5] Fig. 8.3 shows the result of this initial process.[6] MARS can be divided into two main sub-models: the land-use and the transport model. These two model parts are linked together with time lags. Input to MARS comes from external scenarios and policy instruments. Outputs can be in the form of indicators or these can be adapted to form an objective function (OF).

MARS can model the transport and behavioural responses to several demand and supply-side instruments. These impacts can then be used to carry out a standard cost-benefit analysis. MARS assumes that land-use is not a constant but is rather part of a dynamic system that is influenced by

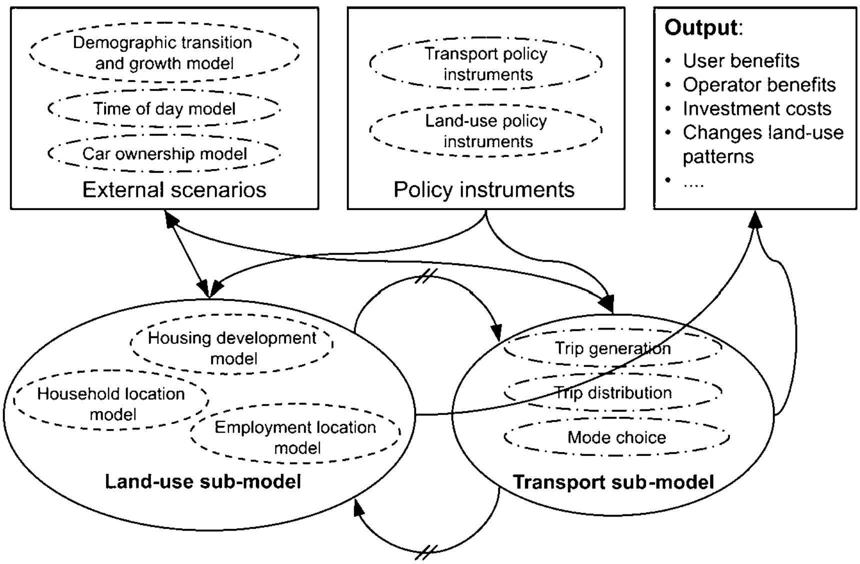

Fig. 8.3. Development of the Qualitative Structure of MARS using Causal Loop Diagramming.

transport infrastructure. The interaction process is modelled using time-lagged feedback loops between the transport and land-use sub-models over a period of 30 years. Two person groups, with and without access to a car, are considered in the transport model. The transport model is broken down by commuting and non-commuting trips, including travel by non-motorized modes. For a given level of employment, the number of peak trips is assumed constant, so that changes in demand reflect changes in land use, trip distribution, modal split and assignment. The modal is calibrated to reproduce observed mode splits and trip generation rates for commuting trips. Comparisons with CEC's model indicate that the MARS model produces similar forecasts, which gives confidence in the use of MARS as a predictive tool. In the version used in this study, the inter-peak is assumed to be un-congested. The land-use model considers residential and workplace location preferences based on accessibility, available land, average rents and amount of green space available. Between 20 and 30 zones are used to represent a city and its immediate environs. The outputs of the transport model include accessibility measures for each zone while the land-use model yields workplace and residential location preferences per zone. The MARS model calculates changes in transport-related impacts, including changes in travel demands by mode, travel time, accidents and pollution. This data in combination with the unit valuations form an input to the welfare function (OF). For a more technical description of the MARS model see Pfaffenbichler and Shepherd (2003).

3.2. Model Scenarios

Three types of regulation schemes are analysed: the existing (*base*) situation, the congestion charging proposals and accompanying uses of revenue set out by the CEC (*Base+Additional* scenario) and variants of those. Whilst the *Base* and *Base+Additional* scenarios have already been defined above, modelling limitations[7] meant that these had to be simplified to the scenarios as defined in Table 8.2. This table also contains the definitions for the eight variant scenarios whose results are presented in this chapter. Variants V0–V3 model different levels of charge on the assumption that revenue is used to fund tramline 3 but not increases in bus frequency. Variants V4 and V5 assume no tramlines, but increases in bus frequencies with and without road pricing at the proposed level. V6 and V7 combine very high cordon charges with no improvement in buses, with and without the three tramlines, respectively. Scenarios V1 and V6 were developed, through the use of multiple model runs, to identify the conditions under which welfare is maximised under alternative OFs. Thus these scenarios, which find optimal cordon charges for the region as a whole and for Edinburgh residents alone with all 3 tramlines and no increase in bus frequencies, were identified after undertaking 12 model runs with varying levels of cordon charge. Optimum bus frequency increases inside Edinburgh (a 20 per cent increase) were identified in a similar manner in scenarios V4 and V5.

3.3. The Welfare Function

All the scenarios were evaluated using the same welfare function. Welfare for the different geographic areas (CEC and rest of study area) was analysed separately. The welfare or OF used is based on previous research work carried out in PROSPECTS (May et al., 2003). The OF consists of a transport user benefit term (consumers' surplus), a transport supplier cost term (producers' surplus), a government revenue term, a CO_2 costs term and a term for monetised values for local pollution and accidents (Minken et al., 2003). All these costs are discounted at the same rate (the official UK government test discount rate of 3.5 per cent) over a 30-year evaluation period. The government revenue term includes a calculation for changes in indirect taxation – including fuel tax. The welfare calculation does not consider the effect of the marginal cost of public funds (MCPF).[8] The user benefits and indirect tax calculations follow UK Department for Transport (DfT) advice (DfT, 2004a). The valuations for the impacts, which are not traded in markets (i.e. time, accidents, carbon emissions, pollution) are those used for the appraisal of major schemes within the UK. Infrastructure investment costs, operating costs and maintenance costs for all modes of transport (including administration of the congestion charge) were derived for each of the scenarios from sources made available to the public inquiry (e.g. TIE, 2004). Each of the scenarios set out in Table 8.2 is therefore fully costed. Inclusion of the costs of administering the scheme is very important, as it has been argued (Prud'homme & Bocarejo, 2005) that in the case of London – using similar technology – the costs are so high that they outweigh the substantial benefits of the scheme. This is not however the finding of the official monitoring study (TfL, 2005), which finds the scheme to be worthwhile.

An adapted version of the welfare function was utilised to analyse the scenario if Edinburgh residents (i.e. CEC) were responsible for setting the cordon charges and determining the investment strategy. The OF for

Table 8.2. Scenario Definition.

Scenario	Tramlines 1 and 2	Tramline 3	Increase in PT Frequency Outside Edinburgh (%)	Increase in PT Frequency Inside Edinburgh (%)	Parking Charges[1]	Cordon Charges Peak (Off-Peak) in Euros	Revenue Sharing
Base scenario	Yes	No			Yes	None	N/A
Base + additional scenario	Yes	Yes	3	3	Yes	3.20 (1.60)	Shared
V0 scenario	Yes	Yes			Yes	3.20 (1.60)	Shared
V1 scenario	Yes	Yes			Yes	10.00 (5.00)	Shared
V2 scenario	Yes	Yes			Yes	13.00 (0)	Shared
V3 scenario	Yes	Yes			Yes	3.20 (0) rising to 15.00 (0)	Shared
V4 scenario	No	No		20	Yes	None	N/A
V5 scenario	No	No		20	Yes	3.20 (1.60)	Shared
V6 scenario	Yes	Yes			Yes	40.00 (20)	None – CEC only
V7 scenario	No	No			Yes	40.00 (20)	None – CEC only

Note: Parking charges remain the same in all scenarios. There are no restrictions on parking availability in any scenario.

Edinburgh residents' comprised the user benefits that accrue to them, the change in external costs that are felt by them, the change in revenues that accrue to transport providers in Edinburgh and the total increase in transport provider operating costs plus the costs of any investment associated with cordon charging (except that already committed by the Scottish Executive).

The OF and its components are used as a first means of comparing the relative impacts of the transport instruments. In addition we discuss the cost implications of each instrument in terms of the change in present value of finance (PVF). The PVF of an instrument or set of instruments is defined as the net discounted financial benefit to government and other providers of transport facilities, both public and private, over a 30-year time horizon, relative to the Do-Minimum. The Do Minimum in this study is the Base scenario as set out in Table 8.2.

PVF is defined as:

$$PVF = \sum_{t=1}^{30} \frac{1}{(1+r)^t}(f_t - I_t)$$

where I_t is the investment cost in year t, r the discount rate used (3.5 per cent) and f_t the net financial benefit to transport suppliers in the modelled target year, compared to the Do-Minimum scenario, taking into account both revenue and operating costs. Road maintenance costs are assumed to be equal to those in the Do Minimum. Any changes in operating costs and running costs for road pricing and tram systems are taken into account.

Aside from the exclusion of the MCPF, the welfare function (or OF) used in the Edinburgh case study has the same structural form as that used in the MOLINO model.

3.4. Assumptions Regarding Population, Employment and Car Ownership

The Do Nothing simply projects the current trends in population growth and car ownership. The study area consists of the CEC authority area and the adjacent authorities (called the Lothians) and consists of 25 zones populated as shown in Fig. 8.2. The total population in the study area in 2001 is 1.07 million. Population growth assumptions have been based on the Edinburgh and Lothians structure plan developed in August 2000:

- The Lothians are forecast to have the biggest population growth in Scotland;
- The Lothians' population increases by 50,000, in 2001–2015;

- The number of households increases by 52,000, in 2001–2015;
- Thirty seven per cent of all households are expected to be single person by 2015;
- The population aged 85+ is expected to increase by 25 per cent to 16,000; and
- The population will be older with 43 per cent aged over 45 compared to 37 per cent in 2000.

Jobs and the economy are growing rapidly; in particular:

- The Lothians economy is expanding rapidly;
- The current unemployment rate of 3.5 per cent is the lowest on record;
- Growth of 30,000 jobs is predicted in Lothian to 2015;
- Seven per cent growth in manufacturing in West Lothian is expected; and
- Most new jobs are being created in the retail, finance and other service sectors;

Transport and in particular the use of car is growing as a result:

- In 1991, 46 per cent of Lothian households did not own a car;
- Car ownership has risen significantly in Edinburgh (by 50 per cent) and the Lothians since 1991;
- In-commuters to Edinburgh are forecast to rise from 88,000 to 100,000, in 2015; and
- By 2015, 42,000 cars will enter the city daily on work trips;

The population is assumed to grow at 0.606 per cent p.a. for the next 25 years followed by 0.2 per cent p.a. until year 30 giving an increase of 17.5 per cent over the 30-year period or an extra 187,000 residents. These assumptions are based on the Edinburgh and Lothians Structure plan (2000) detailed above and extrapolated over the evaluation period. During the same period the number of jobs is expected to grow by 24 per cent. The average level of car ownership is 370 per 1000 population, and this is assumed to grow by 1.2 per cent p.a. based on UK national data for the evaluation period. The vehicle fleet is made up of "average cars" and their fuel consumption is based on UK DfT advice (DfT, 2004b), i.e. the fuel consumption is dependent on a quadratic function of average speed with assumptions about the proportion of diesel cars in the overall fleet. In the base year (2001), the mode split is as shown in Table 8.3.

Table 8.3. Base Year Mode Splits.

	Slow Mode (%)	Public Transport (%)	Private Car (%)
Commute	26	22	52
Other trips	29	14	57

Source: 2001 Journey to Work Census and Scottish Household Survey.

4. MODELLING RESULTS

The following sections present the results of the model runs in terms of effects on welfare and equity. The results presented are comparisons between the existing situation (Do Nothing) and each of the different scenarios analysed (*Base* scenario, *Base+Additional* scenario and the variant scenarios, V0–V7).

Table 8.4 shows the impact on mode split of the various scenarios. Analysis of the *Base scenario* indicates that implementation of tramlines 1 and 2 in 2006 increases the number of PT trips by approximately 51,000 per day (in 2006). This is in line with results from the study commissioned by TIE using the much more detailed TRAM-DELTA modelling system (Leitham, 2003), and presented to the public inquiry. The TRAM-DELTA model suggested an increase in trips of 78,000 – but for all three tramlines.

The *Base+Additional scenario* includes all three tramlines, the double cordon with charges of €3.2 and €1.6 in the peak and off-peak, respectively, and the additional investment in bus services equivalent to a 3 per cent increase in area-wide bus service levels. In this scenario, car use is reduced in the peak by 16,000 trips. In particular, the introduction of the cordon charge reduced the number of outer cordon crossings by 12.9 per cent in the peak period. The combined effect of cordon charges and public transport improvements increased public transport trips by 16,000 in the peak and 75,000 in the off-peak. Implementing the double charging cordon brings in €89 million (£55 million) in toll revenue in the opening year, which is a little lower than the £75 million obtained in the TRAM-DELTA study. This difference is however ascribed to the fact that the MARS model cannot exactly replicate the actual charging regime proposed for Edinburgh – which the TRAM-DELTA model can. The relative consistency between the MARS model forecasts and the TRAM-DELTA model provides confidence that the MARS model is behaving in line with expectation and that its output can be used to form the basis for research conclusions.

Table 8.4. Percentage Change in Mode Split by Scenario Compared to the Do Nothing (Commute trips in 2006).

Scenario	Percentage Change in Slow Mode Use (%)	Percentage Change in Public Transport Use (%)	Percentage Change in Private Car Use (%)
Base scenario	−2.2	4.2	−0.6
Base + additional scenario	−0.2	10.5	−4.1
V0 scenario	−0.1	10.1	−4.0
V1 scenario	3.4	15.8	−7.8
V2 scenario	4.5	17.5	−9.0
V3 scenario	−0.1	10.1	−4.0
V4 scenario	−0.2	0.5	−0.1
V5 scenario	2.7	4.7	−3.1
V6 scenario	9.2	25.1	−14.2
V7 scenario	12.6	18.8	−13.2

Notes:

(i) The peak in the MARS model only comprises of commute trips.

(ii) The total number of commute trips remains constant in each scenario. In response to the transport interventions land use can change, trip origins and destinations can change (re-distribution) and mode can change.

4.1. Maximisation of Welfare Effects for the Region

Table 8.5 and Table 8.6 shows the overall welfare impacts for the region and for Edinburgh, the PVF and an analysis of the time and money impacts for each of the scenarios analysed.

The *Base* scenario, which consists of investment in tramlines 1 and 2, increases overall welfare, but requires significant funding from Government, hence the highly negative PVF. As shown by the *Base+Additional* scenario, the introduction of cordon charging is very beneficial to society, increasing welfare (the OF) by a further €800 million (see Table 8.5). By comparing variant V0 with the *Base+Additional* scenario, it can also be seen that using the revenue from the cordon charge to fund an area-wide increase in bus frequencies is not beneficial to society as a whole.

Total welfare for the whole study area is maximised when the cordon charges are set to €10 (peak) and €5 (off-peak) if the ratio between peak and off-peak has to be maintained (V1). This doubles welfare compared to when cordon charges are set at the level proposed by the CEC. These optimal tolls

Table 8.5. Summary of Overall Impacts on Welfare and PVF (2001 Market Prices, €m, 30 Year Present Value).

Scenario	Travellers' User Benefit[a,d] (Consumer Surplus)		External Costs (Accidents, CO_2, NO_x)	Transport Providers' Producer Surplus[b]			Government Surplus		Economic Indicator		
	Edinburgh residents	Residents of other authorities		Public transport operators	Parking operators	Road toll operators[c,d]	Capital	Indirect taxation	Objective function	Edinburgh residents' objective Function	PVF
Base scenario	475.4	236.8	−11.4	−31.7	−0.8	2.4	−605.6	−33.4	31.7	−232	−669.1
Base + additional scenario	322.0	281.2	6.9	84.9	−8.0	1183.9	−605.6	−434.6	830.8	525	220.6
V0 scenario	295.1	248.8	2.5	250.6	−8.0	1184.6	−605.6	−428.1	940.0	683	393.6
V1 scenario	−708.5	−692.4	69.3	543.7	−15.7	4117.8	−605.6	−1057.9	1650.8	2456	2982.3
V2 scenario	460.5	308.7	11.4	499.9	−14.7	1709.6	−605.6	−550.1	1819.8	1423	1039.2
V3 scenario	538.1	456.3	−7.1	438.0	−13.0	1350.4	−605.6	−458.7	1698.5	1148	711.3
V4 scenario	72.3	22.5	0.0	−74.3	−0.1	0.2	0.0	−4.7	16.0	−8.8	−78.9
V5 scenario	−404.2	2.9	11.1	162.2	−7.3	1468.7	0.0	−365.8	867.7	882	1257.8
V6 scenario	−4631.9	−6351.4	281.8	1006.4	−26.2	10581.5	−605.6	−2506.2	−2251.6	4630	8449.9
V7 scenario	−5420.7	−6469.1	277.6	973.7	−25.6	10903.5	0.0	−2425.4	−2186.1	4824	9426.2

[a]User benefit comprises of time, money and vehicle operating costs.

[b]Transport providers' producer surplus comprises of the difference in revenues and operating costs.

[c]There are two road toll operators: Forth Estuary Transport Authority (Forth Road Bridge) and the authority responsible for congestion charging. Road toll operator surplus is net of congestion charging operating costs and investment costs of the associated investment package.

[d]Excludes the user benefits and capital costs associated with the environmental and social inclusion orientated transport projects in the *Additional* investment package that could not be modelled in MARS. The capital cost of the projects plus the administrative costs of operating the scheme balances the revenue generated in the *Base+Additional* scenario (this is by design).

generate embarrassingly high revenues, which it appears likely, given the results from the analysis of bus frequency increases, could not be spent in an economically efficient manner within the transport sector alone.

As the off-peak cordon charge reduces to zero[9] whilst the peak charge rises to €13, there is a further increase of welfare, though the PVF surplus reduces significantly (V2). The large decrease in the PVF occurs as a consequence of the fact that the majority of trips occur in the off-peak, though as the model assumes that there is no congestion in the off-peak there is no loss of welfare due to the off-peak toll being zero. Variant V3 shows that it is possible to reduce the initial charge in year 5 to the pre-defined level of €3.2 (rising to €15 by year 20) whilst reducing the increase in welfare by only 7 per cent – this could be a means of increasing acceptability without reducing the efficiency of the strategy too much.

Finally, variants V4 and V5 show the effect of increasing urban bus frequencies by 20 per cent without and with the cordon charges proposed by the CEC. This 20 per cent increase in urban bus frequencies is beneficial to existing bus users, new bus users, and by encouraging some mode switching also improves the situation for car travellers. As the benefits to all these users are greater than the costs to society of operating the additional buses, the bus frequency improvements in Edinburgh alone are beneficial to society as a whole. This suggests that targeted (and possibly limited) bus frequency improvements are beneficial. It should be noted that due to modelling restrictions it was not possible to simulate the addition of tramlines along with increased urban bus frequencies.

4.2. Maximisation of Welfare for Edinburgh Residents Only

The welfare of Edinburgh residents was maximised assuming that all toll revenue could be used to benefit these residents. Table 8.6 shows the tests conducted and the change in welfare to Edinburgh residents, to all residents of the region and the PVF. Of particular interest are variants V6 and V7.

As the Variant V6–V7 scenarios assume that all toll revenue can be retained by the residents of Edinburgh to benefit them in some way (through reduced local taxation, for example), it is not surprising that the cordon charges that maximise Edinburgh residents' welfare, €40 (peak) and €20 (off-peak), far exceed the cordon charges that maximise the welfare of all residents in the region. We can also see that such high charges result in highly negative welfare change for the region and are therefore socially undesirable from a regional perspective. From a comparison of V7 and V6 it

Table 8.6. Tolls Paid and Time Savings by Area (2001 Market Prices, €m, 30 Year Present Value).

Package	Edinburgh		Rest of Study Area	
	Tolls paid (€m)	Time savings (€m)	Tolls paid (€m)	Time savings (€m)
Base	0	476	0	244
B+A	719	931	1119	1259
V0	720	904	1120	1229
V1	1876	1041	2897	2077
V2	513	1056	1852	2265
V3	437	1029	1569	2090
V4	0	72	0	23
V5	723	206	1116	978
V6	4461	1225	6776	2679
V7	4498	456	6775	2563

can be seen that from the point of view of Edinburgh residents, investment in trams is not desirable. Tramlines are only justified due to the inclusion of congestion benefits for car users travelling to/from outside Edinburgh. It is more desirable for Edinburgh residents to collect the toll revenue and use it for purposes that benefit them directly (e.g. reduced local taxation) than it is for them to invest it into projects that benefit residents of other local authorities. The main message from these tests is that allowing the CEC to set the charges and collect all revenue to benefit its citizens with no form of control or regulation by a higher authority could lead to excessive charges and a failure to invest the revenue in the local transport system. A similar result has been found for Brussels by Proost and Sen (2006).

Given that Edinburgh stood to gain from much higher cordon charges, it may be questioned why in fact the proposal was for cordon charges below, rather than above, the socially optimal level. Firstly, it should of course be noted that there will be a number of effects from implementing such high cordon charges which the model does not capture, the most significant being economic competition from neighbouring cities such as Glasgow. For example, with such high cordon charges, Edinburgh's retail sector would be put under severe pressure and may experience a significant downturn. Secondly, as we see below, in order to implement the charges there was a need to show that Edinburgh had consulted widely and obtained public support for the proposals. It could only implement them with the agreement of the higher level of authority, the Scottish Executive, who might be expected to pay attention to the interests of the neighbouring authorities.

Lastly but certainly not least, even the modest charges proposed failed to obtain the support of the majority of the population of Edinburgh. Introducing higher charges, at least initially, was clearly not to be thought of, although a process of introducing low charges and then gradually raising them appears to make good political and economic sense.

4.3. Equity

This case study was concerned with spatial equity – in the sense of whether or not the distribution of costs and benefits throughout space was broadly equal. Of particular interest is whether costs were borne by those who live outside Edinburgh whilst the benefits were principally enjoyed by those who live in Edinburgh. Table 8.7 shows the distribution of time benefits, toll revenue, PVF and the Edinburgh residents' share of these for the *Base* scenario, *Base + Additional* scenario and the variant scenarios. Table 8.6 shows the absolute tolls paid and time savings by area.

Firstly, it can be seen that the Edinburgh residents benefit from improved public transport services as they receive the majority share of the time benefits from the tramlines.[10] Conversely, the time benefits accrue to the car users from outside Edinburgh who benefit from reduced demand within the Edinburgh area. In terms of the overall share of time benefits, the Edinburgh

Table 8.7. Distribution of Benefits and Revenues (2001 Market Prices, €m, 30 Year Present Value).

Package	Time Benefits					Money	
	Car		Public transport		Total	Toll revenue	
	€m	Edinburgh share (%)	€m	Edinburgh share (%)	Edinburgh share (%)	€m	Edinburgh share (%)
Base	218	10.1	502	90.5	66.1	0	N/A
B + A	1127	10.7	1063	76.2	42.5	1839	39.1
V0	1112	10.7	1021	76.9	42.4	1840	39.1
V1	1863	11.6	1259	65.8	33.5	4773	39.3
V2	2017	11.6	1304	63.0	31.8	2365	21.7
V3	1869	11.5	1250	65.2	33.0	2006	21.8
V4	21	11.7	74	94.0	75.6	0	N/A
V5	891	10.8	293	37.5	17.4	1839	39.3
V6	2340	14.2	1564	57.1	31.4	11237	39.7
V7	2245	14.3	774	17.7	15.1	11273	39.9

share is reduced as the charges are increased – as car time benefits increase with increased charges at a greater rate than the public transport time benefits.

In terms of tolls paid, the proportion of tolls paid by residents of Edinburgh remains fairly constant at around 39 per cent. The share of any surplus finance is assumed to be pro-rata with the population in all cases except for V6 and V7 where the welfare of Edinburgh residents was maximised with the assumption that they could retain all toll revenues for their own benefit.

In the *Base* scenario, benefits are shared but with the largest share going to Edinburgh residents. With the *Base+Additional* scenario, both sets of residents receive benefits greater than the tolls they pay, although Edinburgh residents achieve a greater surplus of benefits over tolls than do the rest. With the optimal tolls, tolls exceed benefits for both sets of residents; the sharing of revenue between authorities would permit both sets of residents to be fully compensated for the tolls paid. Only option V6 and V7, where all toll revenue accrues to Edinburgh, leads to a situation in which residents from outside Edinburgh are made substantially worse off, and which they would undoubtedly see as highly inequitable.

These results suggest that the distribution of revenue, as proposed by CEC, was broadly equitable in a spatial sense and compensated the populations of the areas, which incurred additional costs as a consequence of the congestion charge. Whether or not the cordon charging structure is regarded as equitable and therefore acceptable is a different issue to which we turn in the next section.

5. ACCEPTABILITY

Ultimately, it was the unacceptability of the congestion charge to the residents of Edinburgh that resulted in the CEC dropping it as a policy. Despite the finding that overall no group of residents would lose, and those of Edinburgh itself would gain substantially from the CEC proposals, almost 75 per cent of Edinburgh residents who voted rejected the congestion charging proposals – with a turnout of over 61 per cent. Additionally, three of the four adjacent local authorities, whose residents had no say in the local referendum, were opposed to the proposals. The reasons for their opposition differ from that of the Edinburgh residents.

The adjacent local authorities opposed the congestion charging proposals due to a perceived unfairness in the charging regime. It was felt that the double cordon nature of the charge affected their residents more than it did Edinburgh residents. For example, half of the working population of mid-Lothian travel into Edinburgh every day for work and would be directly affected by the charge, whilst Edinburgh residents who live and work between the cordons would not have to pay the charge. This issue was exacerbated by the fact that CEC residents who lived outside the outer cordon were to be exempted from paying the charge to cross the outer cordon.

The legal position regarding road user charging meant that a single statutory authority would be responsible for the collection and distribution of revenue from the congestion charging scheme. As the regional transport authority (SESTRAN) was only a voluntary organisation it could not hold that responsibility. Instead, the responsibility for the setting of charges and distributing the revenue would have fallen to the CEC. Whilst CEC had given an undertaking to distribute revenue in proportion to trip origins this was not underpinned by any legal obligation on their part. The neighbouring local authorities therefore had no legal grounds in which to secure funding for any public transport improvements they needed if congestion charging was to go ahead. This created a degree of uncertainty regarding their ability to deliver the public transport improvements that their residents would wish to see.

Residents of adjacent local authorities also felt that the contents of the proposed investment package (the *Base + Additional* scenario) were more suited to the needs of Edinburgh residents than to their needs. For example, residents of mid-Lothian would have preferred a larger proportion of the congestion charging revenue to be directed towards heavy rail projects – which are more suited to the longer distance commutes that they would be making – than to bus and tram.

A common concern amongst both residents of Edinburgh, and residents of neighbouring local authorities was that the congestion charging scheme would be in place before the alternative public transport facilities were available. For example, the congestion charge was proposed to begin in 2006, and yet the first tram scheme would not open until 2011. Without alternatives to the car the congestion charge was therefore perceived to be a tax on car users. Whilst this claim was rejected by the public inquiry the perception of the charge as a tax remained. The concern regarding lack of public transport alternatives in the early years of the congestion charge was compounded by the suspicion that CEC would not be able to deliver the promised public transport improvements, particularly the major infrastructure schemes – where cost over-runs are common place – but also the bus improvements, as the bus market in Edinburgh is de-regulated and CEC's influence over it is limited. A negative media campaign and the decision by Transport for London to raise the London congestion charge from £5 to £8 relatively soon after its introduction in London may also have undermined residents' confidence in CEC and congestion charging schemes in general.

A major cause of opposition to the scheme was thought to be related to the public's limited understanding of the scheme's operational aspects and how it would affect them. This arose for several reasons. Firstly, local transport officers and local academics have argued that a bigger and better information campaign would have considerably helped user acceptance (Saunders, 2005). A budget of £600,000 was allocated to consultation and promotion activities by CEC, which is low compared to a reported £12 million budget for the London scheme. Secondly, the marketing campaign for the Edinburgh proposals was hindered by the fact that there were no headline grabbing projects to be facilitated by the congestion charge – the Scottish Executive had decided to centrally fund the tram schemes up to £375 million. CEC had initially proposed to the Scottish Executive that the Executive provide match funding of £375 million in recognition of the importance of the congestion charging policy as a means of achieving national transport objectives. Parallels were drawn with Norway, where income from the toll rings was match funded by additional government funding. The effect of the Scottish Executive's decision to independently fund two of the three highest profile projects not only significantly changed the scheme from a presentational point of view, but also altered the appraisal outcome (Saunders, 2005). Finally, the marketing campaign lacked a high-profile public champion for the proposals – a role fulfilled very well for the London proposals by Ken Livingston, the London mayor.

There is some debate regarding whether or not a referendum on the congestion charge was required. Norwegian evidence suggests that opposition to road user charging proposals is higher before implementation than after it, thus it might be better to undertake a referendum after the scheme has been in place – as is proposed for Stockholm. However, the ministerial requirement to show "clear public support" that was placed on CEC, although coming after the decision to hold a referendum, reinforced the Council in its view that a referendum was needed. The referendum however ended up being influenced by party politics and wider political issues such as a backlash against the Labour party. The strong "no" vote therefore, whilst not necessarily reflecting a complete rejection of the concept of congestion charging, has had that outcome.

All of this is in marked contrast to the situation in London. The London congestion charge only applies in Central London, which is a small part of

Greater London, and only a very small proportion of commuters to Central London use the car. It was proposed by the mayor of London, in his role as leader of the body responsible for planning and transport throughout Greater London and who at the time of its introduction was independent of all political parties. He devoted substantial resources to information and promotion of the scheme, and was able to take it through without requiring the agreement of any other authority. Arguably, the Edinburgh scheme would have made much more progress if it too had been promoted initially just as a city centre cordon, with the option of adding an outer cordon later.

The main lesson of the Edinburgh experience is the need to build up consensus on a regional basis with an agreed and clearly committed use of revenue that is seen as both efficient and fair, before the process of implementing a reform of road prices begins. For the administration of the scheme to be seen as fair it may also be necessary to review the legal basis for implementing road user charging to ensure that the legal framework permits the adoption of a scheme that can be administered at, for example, a regional level. Additionally, the Edinburgh experience indicates that it is important for the public transport measures, associated with the road user charge package, to be implemented simultaneously with, or before, the road user charge, otherwise the proposals can be perceived as a tax.

6. CONCLUSIONS

Our research has looked at cordon charges and revenue use for Edinburgh in terms of who should set charges and how any revenue raised should be used. To some extent the answers to these questions are inter-linked. Firstly, we can say that leaving the CEC to set the cordon charges and decide how the revenue raised is to be used, without any form of external constraint or regulation by a higher authority, may give rise to a sub-optimal situation from the regional perspective (and country perspective). In the worst case, such a scenario may lead to a more inefficient situation for the region compared to the existing situation (where congestion charging does not exist). This is because the CEC will act to maximise the welfare of its own residents and therefore has a significant incentive to charge residents of other authorities high prices (i.e. a high cordon charge). Surplus revenue would be invested in schemes that benefit only Edinburgh residents, which leads to low investment in public transport services (compared to, for example, alternatives such as reducing local taxation). Residents of Edinburgh would benefit at the expense of those in the surrounding area.

In practice, the CEC actually proposed charges below the level that would maximise the welfare of its residents as well as below the economically efficient level for the region. It also proposed a revenue sharing agreement with its neighbouring authorities. There are various possible reasons for this. Firstly, as we noted it may have feared the consequences of excessive charges for the retailing and business sector. This will be a restraining influence on cities particularly where they have closely competing neighbours; it will be less of a constraint on London than on Edinburgh. Secondly, in Edinburgh the proposal faced a public inquiry and was subject to the approval of the Scottish Executive and therefore subject to a degree of higher control. Thus, there was a need for the CEC to try to establish consensus with its neighbouring authorities. Finally, current legislation in the UK requires that revenue raised has to be used to finance local transport schemes and there are only a limited number of economically efficient public transport projects.

If we were to maximise welfare from the viewpoint of the region, then the optimal cordon charges produce a significant financial surplus, which could not be spent in the transport sector without investing in inefficient projects. A relaxation of current legislation to permit any financial surplus to be spent in other sectors would therefore be necessary. In theory, then, control of price setting and revenue distribution should reside with either a joint partnership of local authorities or the Scottish Executive as realistic scheme managers. Charges would be set at an optimal level for the region and investments made in transport projects, which return a reasonable benefit (positive welfare function or benefit/cost ratio above a certain amount). Surplus revenues would either be invested in other sectors or shared amongst all residents. In the short term such a change in legislation may be unacceptable and cordon charges may need to be set lower than optimal. In the longer term, and if the concept of tolling becomes acceptable, the Scottish Executive might support further increases in toll levels to an optimal level, with surplus revenues allocated outside the transport sector. Since earmarking appears a pre-requisite for public acceptability, some form of earmarking might need to be maintained, but revenues could be earmarked for pre-determined efficient purposes associated with strong social benefits.

As we have also shown, cordon charges at the level that the CEC had proposed are highly beneficial and earmarking revenue use to investment in light rail or urban bus services does not change this. The proposed package with charges set at €3.20 in the peak (€1.60 off-peak) generates enough net revenue to cover the costs of the total earmarked investment package. The proposed revenue sharing arrangements appear to work well in leading to an efficient and equitable solution.

Given that as proposed, the scheme would have had major benefits, and no area would have been disadvantaged overall, the question naturally arises of why it was ultimately rejected. It appears that the public consultation exercise was not as effective as it might have been – too few resources were devoted to explaining the scheme and its benefits. There were also concerns that alternatives such as improved public transport would not be in place soon enough, and the use of revenue was not made sufficiently explicit. Undoubtedly, the opposition of surrounding authorities and the view that the scheme unfairly favoured Edinburgh residents were significant factors. The fact that the scheme was being carried forward by CEC and that there was no legally binding mechanism to ensure that the revenue was spent in the way promised will undoubtedly have added to this hostility.

Ultimately, the lesson of the Edinburgh experience is the need to build up consensus on a regional basis with an agreed and clearly committed use of revenue that is seen as both efficient and fair. The new statutory Regional Transport Partnerships in Scotland may make it easier to take forward such schemes in the future.

NOTES

1. Net revenue is total revenue minus the costs of operating the congestion charging scheme.
2. SESTRAN is a voluntary regional transport body for South East Scotland whose members comprise the local authorities within the region.
3. Net support is the difference between the proportion of people that support the congestion charge and the proportion that oppose it.
4. It should be noted that in our Edinburgh case study the land use responses to the transport policy interventions are small and are not therefore reported.
5. For a description of CLD see, e.g. Emberger and Fischer (2000).
6. Route choice in MARS is not modelled explicitly. MARS is a sketch model in which each origin–destination movement is represented by a congestible notional link (by mode and time period).
7. It should be noted that due to modelling restrictions a number of simplifications to the proposed cordon charging scheme had to be made.

(i) The modelled scheme includes a €3.20 (£2) charge in the peak for both cordons with exemptions from the outer cordon charge for city of Edinburgh residents residing outside that cordon (e.g. in Currie, Balerno and Ratho), which is consistent with the scheme the CEC had proposed. The proposed scheme does not include a charge for the outer cordon in the off-peak period. As it is not possible to change the structure of the cordon charges between periods using MARS, this was approximated by charging €1.6 (£1) for both the inner and outer cordon in the off-peak period.

(ii) The MARS model is also a strategic model therefore only significant public transport investment schemes such as trams or wide-area bus improvements can be represented within it. Changes in quality of road surface (e.g. through increased maintenance) cannot be represented in the model nor can improvements in the city centre streetscape (e.g. better lighting, pavements, etc.).
(iii) The £174 million in the *Additional* package for bus service improvements was assumed to be an undiscounted sum available for 25 years from 2006 to coincide with the introduction of road user charging and the tramlines. This is then equivalent to £6.96 million p.a., which can purchase a 3 per cent increase in area-wide public transport frequencies.

8. Current UK appraisal practice does not include the MCPF. It is felt that the inclusion of the MCPF would give a set of results in which the value of the OF function would alter, but would not change the conclusions for the use of revenue.
9. The MARS model assumes there is no congestion in the off-peak.
10. Though as mentioned earlier such benefits are less than the costs of investing in and operating the tram system. The tram system only gives a net social benefit if the benefits to car users outside of Edinburgh are included in the analysis.

ACKNOWLEDGEMENTS

The authors are grateful to Jo Baker and Russell Tricker from TTR Ltd who undertook the acceptability research, and to Pauli Pfaffenbichler and Guenter Emberger from the Technical University of Vienna for help with the MARS model. Any errors or shortcomings are, of course, the sole responsibility of the authors.

REFERENCES

Bös, D. (2000). Earmarked taxation. Welfare versus policy support. *Journal of Public Economics*, *75*(3), 439–462.

City of Edinburgh Council (2004). *Report on the inquiry into the draft Edinburgh congestion charging order*. City of Edinburgh Council, October 2004. http://download.edinburgh.gov.uk/TransportEdinburgh/Report_on_the_Public_Inquir2004.doc

DfT (2004a). *Transport user benefit calculation: Tag unit 3.5.3*. Departmentt for Transport, Transport Analysis Guidance TAG. www.webtag.org.uk

DfT (2004b). *Value of time and operating costs: Tag unit 3.5.6*. Department for Transport, Transport Analysis Guidance TAG. www.webtag.org.uk

Emberger, G., & Fischer. P. (2000). *Easy understanding of causal loop inherent dynamics – EUCLID method.* Report, Institute of Transport Planning and Traffic Engineering, University of Technology Vienna, Vienna, Austria, pp. 4, www.ivv.tuwien.ac/institute/members/GuenterEmberger/euclid.pdf

Ison, S. (2000). Local authority and academic attitudes to urban road pricing. *Transport Policy*, *7*(4), 269–278.

Leitham, S. (2003). *Use of CEC/TIE LUTI model. Information Note No 11*. Report to TIE. MVA Ltd, Edinburgh.

May, A. D., Karlstrom, A., Marler, N., Matthews, B., Minken, H., Monzon, A., Page, M., Pfaffenbichler, P. C., & Shepherd, S. P. (2003). *Developing sustainable urban land use and transport strategies – A decision makers' guidebook*. Leeds: Institute for Transport Studies, University of Leeds.

Minken, H., Jonsson, D., Shepherd, S. P., Jarvi, T., May, A. D., Page, M., Pearman, A., Pfaffenbichler, P. C., Timms, P., & Vold, A. (2003). *Developing sustainable urban land use and transport strategies – A methodological guidebook*. Oslo, Norway: Institute of Transport Economics.

Pfaffenbichler, P. (2003). *The strategic, dynamic and integrated urban land use and transport model MARS (metropolitan activity relocation simulator)*. Thesis available from Technical University of Vienna.

Pfaffenbichler, P., & Shepherd, S. P. (2003). A dynamic model to appraise strategic land-use and transport policies. *European Journal of Transport and Infrastructure Research. Special Issue: Sustainable Transport*, *2*(3/4), 255–283.

Proost, S., & Sen, A. (2006). Urban transport pricing reform with two levels of government: A case study of Brussels. *Transport Policy*, *13*(2), 127–139.

Prud'homme, R., & Bocarejo, J. P. (2005). The London congestion charge: A tentative economic appraisal. *Transport Policy*, *12*(3), 279–287.

Saunders, J. (2005) Congestion charging in Edinburgh: Some lessons on acceptability. *Proceedings of planning for sustainable land use and transport, European transport conference*, Strasbourg, AET Transport, UK.

TIE (2004). *Statement of case*. Edinburgh: Transport Initiatives Edinburgh.

Transport for London (2005). *Central London congestion charging scheme impact monitoring summary review*. London: TfL.

Tricker, R., Fereday, D., Pickup, L., Norheim, B., Bekken, J.-T., Shepherd, S. P., Laird, J. J., Nash, C. A., & Suchorzewski, W. (2006). *Report on the implementation of urban case studies*. REVENUE Project Deliverable 5, Funded by 5th Framework RTD Programme, ISIS, Rome, 16 January.

This page intentionally left blank

CHAPTER 9

THE SWISS RAILWAY INVESTMENT FUND

Laurent Cretegny, Urs Springer and Stefan Suter

ABSTRACT

The new alpine railway tunnels and the distance-dependent heavy vehicle fee (HVF) are key elements of Swiss transport policy. The tunnels are financed by the railway investment fund, which is fed mainly by a two-thirds share of the revenues from the HVF. We analyse the railway investment fund and earmarking policy using a modified version of the MOLINO model. We find that with given investment, welfare increases with the proportion of HVF revenues that are allocated to the rail fund. The current practice of using two-thirds of HVF revenues to finance new railway infrastructure is therefore welfare enhancing although full earmarking would be better still. Alternative pricing regimes with higher transport prices are also found to be superior to the status quo.

1. INTRODUCTION

1.1. Swiss Transport Policy

Between 1981 and 2001, the number of heavy goods vehicles crossing the Swiss Alps increased from 312,000 to 1,371,000.[1] Transit traffic accounts for

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 189–215

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19009-0

the majority of heavy goods vehicles on the transalpine corridors: In 2003, 77.8 per cent of the vehicles recorded on the Swiss Alpine Corridors had neither origin nor destination in Switzerland.[2]

To cope with the growing transalpine freight traffic on the road, Switzerland introduced the distance-dependent Heavy Vehicle Fee (HVF) in January 2001. The HVF is levied on vehicles with maximum permissible weight above 3.5 tons, and must be paid on all roads in the country.[3] Two-thirds of the revenues from the HVF are allocated to the railway investment fund. The objective of this fund is to finance the construction of new railway infrastructure, particularly the New Alpine Rail Tunnels (NART). This novel way to use revenues from road transport was important for the political acceptance of the HVF. The NART will contribute to shift freight transport from road to rail and to improve passenger transport connections to Italy.

1.2. The Railway Investment Fund

The railway investment fund, FINÖV,[4] was created in 1998 as a legally independent fund. Its objective is to finance a number of new railway infrastructure projects, namely improvements of the railway network, the NART, noise reduction programs, and new interurban high-speed railway connections. The fund has an overall value of €20 billion (CHF 30.5 billion, 1995 prices). We limit our analysis to the NART, which are currently estimated to cost about €10 billion. The income of the FINÖV fund shall come from the following sources:

- Two-thirds of the revenues of the HVF
- Revenues from the federal fuel tax
- Capital market loans
- 0.1 per cent of the value-added tax (VAT)
- Private-public partnerships (PPP)

Up to now, capital market loans and PPP have not been used as financing mechanisms. The federal government has provided substantial loans to the fund. Fuel tax revenues may be used to finance maximum 25 per cent of the total investment.

1.3. Objective and Structure of the Chapter

We examine the efficiency and distributive effects of the FINÖV fund. Specifically, we address the following questions: What are the welfare costs of earmarking and cross-financing from the road to the railway sectors?

What is the optimal scope of earmarking? Should the revenues from road pricing be used to finance investments in the same mode, in another mode (railways) or in both modes? Does economic efficiency demand the extension of both railway and road capacity in the transalpine corridors? Should congestion charges be added to existing transport taxes, or should existing taxes such as the fuel tax be reduced with the introduction of congestion charges?

The chapter is organised as follows: Section 2 describes the modelling approach, in particular our adaptations of the MOLINO model, and the data used to calibrate the model. In Section 3, the policy scenarios and their impacts on the welfare of different user groups are described. The final section contains a summary of the results and conclusions.

2. MODEL

We analyse the railway investment fund using a modified version of the MOLINO model. The standard model has been developed within the REVENUE project.[5] The adaptations and amendments of the original model are described in Section 2.3 below.

2.1. Scope

The NART will expand transport capacity on the Gotthard and Lötschberg transalpine corridors. Therefore, we limit our analysis to these corridors and do not model transport in the rest of Switzerland. The Lötschberg base tunnel is located between Frutigen and Steg, the Gotthard tunnel between Erstfeld and Biasca (see Fig. 9.1 below).

We consider two modes, railway and road transport. Railway transport takes place on both the Gotthard and Lötschberg corridors, road transport is limited to the Gotthard. Specifically, we analyse the Gotthard road from Erstfeld to Biasca (80 km) and the competing railway links at Gotthard (Erstfeld to Biasca) and Lötschberg (Thun to Brig) (88 km).[6] Within road transport, we distinguish between peak and off-peak traffic.

In both modes, there are four types of transport users:

- Passengers from low-income households
- Passengers from high-income households
- Domestic freight traffic
- Transit freight traffic

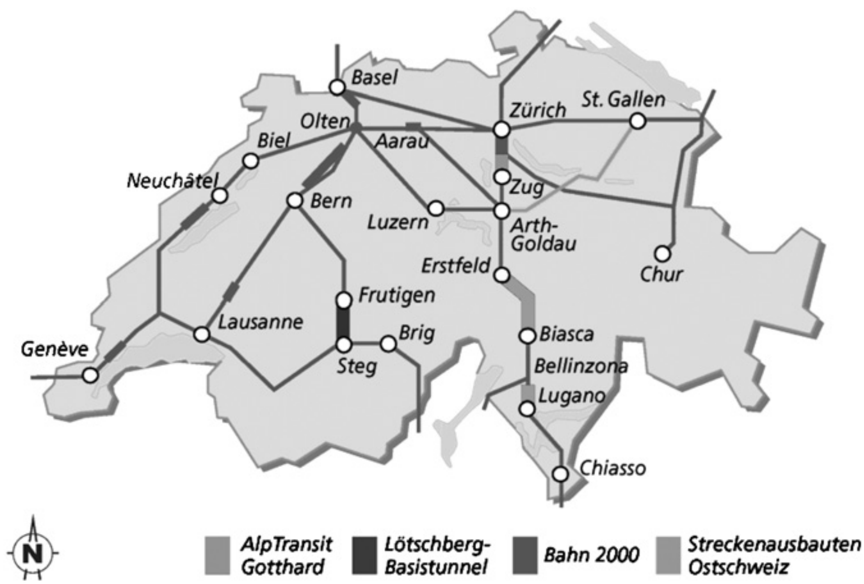

Fig. 9.1. New Railway Investment Projects in Switzerland. *Source:* http://www.alptransit.ch

Domestic freight transport includes all traffic with origin or destination in Switzerland (local, export, and import). Transit freight transport only passes through the country.

There are two funds in the model. The railway investment fund finances the construction of the two new railway tunnels through the Alps. The road investment fund finances the extension of the existing Gotthard road tunnel from two to four lanes.

2.2. Regulation Schemes

We distinguish three regulation schemes:

- *Scheme A: Existing transport pricing*: Transport pricing is based on fuel and vehicle taxes, the passenger car vignette for the usage of motorways, the heavy vehicle fee, and railway track charges.
- *Scheme B: Existing taxation with internalisation*: The existing fuel and vehicle taxes as well as track charges remain in place. Instead of the HVF, charges for marginal environmental and congestion costs are levied on

passenger and freight transport. This scheme is considered as a politically feasible road pricing scenario (the complete abolishment of fuel taxation is not considered as politically viable). It contains some elements of social marginal cost pricing (SMCP) (environmental and congestion costs are taken into account), but does not entirely follow a marginal cost approach.[7]

- *Scheme C: Congestion and infrastructure charging*: All existing taxes and charges are replaced by transport prices which equal the sum of marginal costs of infrastructure operation and maintenance, marginal external environmental costs, and marginal external congestion costs. Only congestion costs are determined endogenously, the other cost components are exogenous inputs into the model. Scheme C partly resembles SMCP and would appear to be the most efficient pricing scheme. However, the subsequent analysis will show that this is not the case.

Within each regulation scheme, there are different *scenarios* with varying assumptions about revenue use and investment. Two different *investment paths* are considered within the regulation schemes (MOLINO requires the specification of an exogenous investment path):

- Two new railway tunnels opening in 2007 (Lötschberg) and 2015 (Gotthard) are built. Road capacity at Gotthard is not expanded. This scenario corresponds to the current Swiss transport policy strategy for transalpine traffic which aims at a modal shift from road to rail transport and therefore abstains from expanding the capacity of the transalpine road network.
- Two new railway tunnels are built, and road capacity is expanded from two to four lanes in the Gotthard tunnel.

The investment in the extension of the road tunnel would alleviate congestion. Even though we have assumed no congestion for rail in the benchmark case, the investment in the new railway tunnels increases the attractiveness of the railway links, because they make them significantly shorter than the existing ones. Consequently, less time and other resources are needed for a trip through the Alps.

Four different levels of cross financing are examined:

1. *No cross financing*: None of the revenues from the HVF are used to finance the construction of the new railway tunnels. Two-thirds of HVF revenues are allocated to the road investment fund, one-third goes to the local government. The new railway tunnels are financed by labour tax and fuel tax revenues from the national government.

2. *Equal distribution*: The revenues from the HVF are distributed equally between the railway investment fund, the road investment fund, and the local government (one-third of revenues each).
3. *Enhanced cross-financing*: Two-thirds of the revenues from the HVF are used to finance the construction of the new railway tunnels, one-third is allocated to the local government. This corresponds to the existing policy.
4. *Full cross financing*: All revenues from the HVF are used to finance the construction of the new railway tunnels.

2.3. MOLINOinGAMS

We have extended MOLINO in several ways. As the extended model has been coded in GAMS (Brooke, Kendrick, Meeraus, & Raman, 1998), it is called MOLINOinGAMS.

2.3.1. Pricing

In the *congestion and infrastructure charging* scenario (scheme C), we have *set existing taxes to zero*. They are replaced by transport prices, which equal the sum of marginal infrastructure operation costs (MOC), marginal infrastructure maintenance costs (MMC), marginal external environmental costs (MEC), and marginal external congestion costs (MCC).

Note that the resulting transport prices are not optimal transport prices. This is due to the way the toll is implemented in MOLINO: the introduction of the toll leads to a change of transport demand. In order to determine the level of the optimal toll, the model would have to take into account the change of demand resulting from the implementation of the toll. This feedback, however, is only implemented for the congestion charge (MCC), but not for the other three components.[8]

2.3.2. Structure of Preferences and Production

The structure of passenger preferences and freight production is depicted in Fig. 9.2. σ_s and σ_f represent the elasticities of substitution for passenger and freight transport, respectively.[9] Railway and road passengers are assumed to maximise their utility under a constant budget constraint, whereas producers (freight trains, trucks) are assumed to minimise their cost subject to a constant production technology. We assume no congestion in railway transport. The utility function of passengers, U_s, where $s = \{$low-income passengers, high-income passengers$\}$, is identical to the freight production function U_f where $f = \{$local freight, transit freight$\}$. Consumers choose first

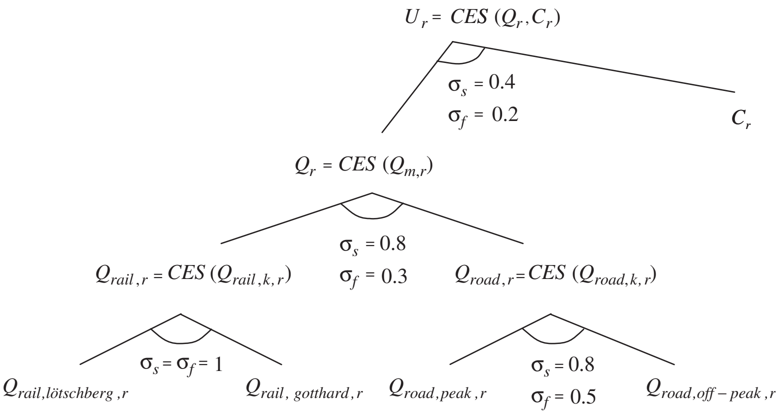

Fig. 9.2. Structure of Preferences and Production.

between a transport composite good (Q_r) and a general consumption composite good (C_r). The upper level branch from the transport composite good presents the choice of transport mode (Q_m), i.e. rail or road. The lower level of the rail mode is a mixture of Lötschberg and Gotthard links whereas the lower level of the road mode is a combination of driving during peak and off-peak periods.

2.3.3. Social Welfare Function

The social welfare function for each period in MOLINOinGAMS is similar to the one in MOLINO:

$$U = \alpha_1(U_{P1} - \beta_1 \times U_{F1}) + \alpha_2(U_{P2} - (1 - \beta_1)U_{F1}) - \beta_2 \times U_{F2} \\ - C_{Ex} + MCF_1 \times NR_{F1} + MCF_2 \times NR_{F2} \quad (1)$$

The first component of social welfare is the monetised utility of consumers. The welfare of rich and poor households is given equal weight ($\alpha_1 = \alpha_2 = 1$). The cost of freight transport is taken as a measure of the consumers' surplus of the demand for freight transport services (U_F),[10] and is subtracted from social welfare.[11] Since the cost of freight transport is larger than the consumers' surplus of passenger transport, total welfare reported by MOLINOinGAMS is negative.[12] Social welfare is diminished by external costs (C_{Ex}) from accidents, air pollution, and noise (external congestion costs do not enter the welfare function directly, but indirectly through the welfare of consumers). Finally, social welfare includes net revenues (NR) – the difference between annual income and expenditure – of both investment funds. The marginal cost of public funds (*MCF*) of the net revenue is 1 for both funds.

To obtain a single measure of welfare, we compute the *present value* of all future welfare using a discount rate of 2 per cent.

2.3.4. Accounting

The managers and operators of infrastructure for both modes are *public*. They are assumed to balance their budget within each period. This means that the subsidy to each manager and the subsidy to each operator from the funds are set to equal the annual differences between charges and revenues. The rail operator pays infrastructure charges (track charge revenues and subsidies from the road fund) to the railway infrastructure manager. The road operator does not pay any infrastructure charge. To cover its operating costs, 7 per cent of the gross HVF revenues are used. The net HVF revenue goes to the railway investment fund and to the local government (see Fig. 9.3). Both the railway fund and the road fund are assumed to balance their budget annually with respect to operating expenses of managers and operators, but not with respect to investment flows. Surpluses or losses are passed on to the federal government.

Investment in roads is financed only by the road fund. Investment in the new railway infrastructure is financed by two-thirds of net HVF revenues, part of the revenues from the fuel tax, and 0.1 per cent of the VAT revenue in reality. In our model, we have implemented the investment as a constant flow over the first 15 years during which both tunnels are constructed.

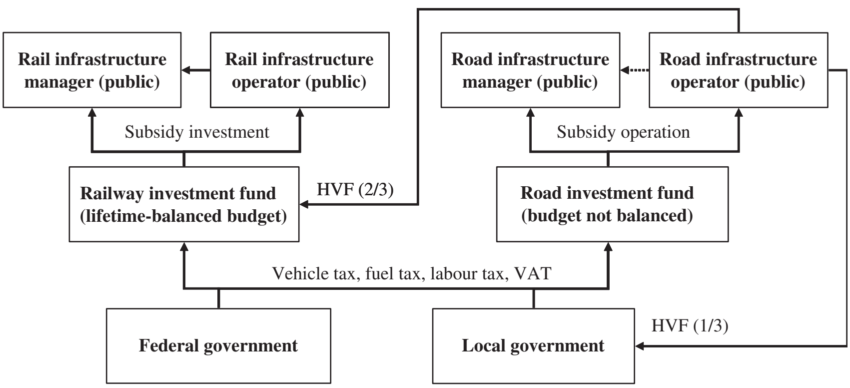

Fig. 9.3. Flow of Funds in MOLINOinGAMS.

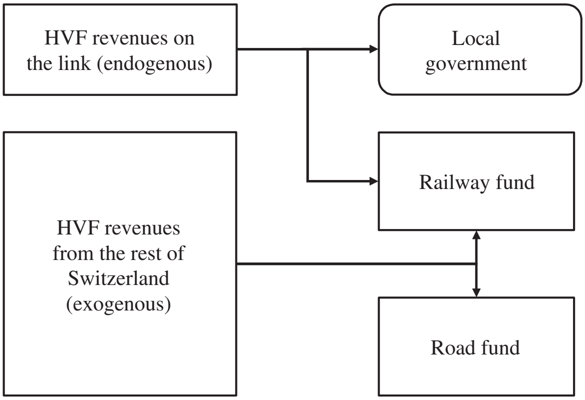

Fig. 9.4. Allocation of HVF Revenues.

The HVF accounts for 50 per cent of the annuity entering the railway fund, the rest comes from the VAT and the fuel tax (25 per cent each).

Investment expenditures exceed the income of the railway fund, which means that it runs a deficit during the first 15 years. When the tunnels are completed, the net revenue becomes positive and is used to repay the debt of the fund. At the end of the time period considered (40 years), the entire debt plus interest are repaid and the fund ends up with a *zero net asset value*. This repayment constraint has been introduced because the FINÖV fund has been created solely for the purpose of financing the four infrastructure projects mentioned above.

The HVF collected on the link is endogenous and varies between pricing scenarios and over time. The HVF from the rest of Switzerland, in contrast, is a constant exogenous variable.[13] In contrast to the current policy, the two components of HVF revenues are treated differently under the earmarking rules: HVF revenues collected *on the link* are either allocated to the local government or to the railway fund, whereas HVF revenues from the rest of Switzerland are allocated either to the railway or to the road fund (see Fig. 9.4 above). If all HVF revenues are allocated to the railway fund, its income exceeds total expenditures, so that some of the money can be recycled through a reduction of the labour tax. If only a small share of the HVF revenues is allocated to the railway fund, the resulting deficit of the fund has to be financed by labour tax money.

2.3.5. Capacity Expansion

Unlike in MOLINO, investment costs do not occur at the same time as the increase in capacity in MOLINOinGAMS. This change was necessary because of the long construction time of the railway tunnels. Furthermore, investment levels within the same mode are allowed to differ, since the Lötschberg and Gotthard tunnels will not open at the same time and have a different capacity expansion.

2.4. Data

The model runs from the year 2000 until 2040. It has been calibrated to the year 2000. The input data for transport volumes used in the model are summarised in Table 9.1. A trip is a passage through the transalpine corridor in one direction of a vehicle of the respective mode, i.e. cars for passenger road transport, trucks for road freight transport, and trains for railway transport.

We set the threshold between high-income and low-income households at € 2,666 gross monthly income. In the year 2000, about 18 per cent of households fell below this threshold.

MOLINO requires a uniform growth rate of transport demand. We assume that demand increases at a rate of 1.35 per cent per year in all modes.

We use a *maximum speed* of 120 km/h on the road link from Erstfeld to Biasca, which corresponds to the speed limit on the motorway (in the tunnel, the limit is 80 km/h). During peak periods, we assume a speed of 56 km/h. This yields a benchmark capacity for the Gotthard road tunnel of approximately 1,000 vehicles per hour and direction, which is consistent with the results of several studies of road congestion at Gotthard. The velocity of trains is set at 60 km/h (on the whole rail link through Switzerland).

Congestion at Gotthard has been recorded during 760 h in the year 2000 (in both directions), which corresponds to 9 per cent of total time.[14] Since

Table 9.1. Number of Trips (Rail: Trains, Road: Cars/Trucks).

	Passenger Transport		Freight Transport	
	Low income	High income	Domestic	Transit
Rail Lötschberg	1,539	8,807	8,792	3,768
Rail Gotthard	2,957	16,925	22,342	9,575
Road peak	215,691	1,234,392	73,937	118,702
Road off-peak	726,540	4,157,952	311,313	499,798

the traffic intensity is much higher during peak times, the share of vehicles in the peak period is higher than the share of time. Based on automatic traffic recording data,[15] we estimate a share of vehicles during the peak period of 22.9 per cent (passenger cars) and 19.2 per cent (trucks).

2.4.1. Cost Data

The value of time for the different user categories is given in Table 9.2 below. The figures for passenger transport are for leisure trips between 80 and 85 km. According to Axhausen et al. (2004), the value of time is lower for public transport (rail) than for private transport (road). For freight transport, the same value for road and rail transport has been used.

Resource costs for passenger cars are based on data from Touring Club Suisse (TCS, 2004). For low-income households, we use a resource cost rate of €0.39/km. For high income households, the rate is €0.42/km. Resource costs for transalpine railway transport are based on Ecoplan (1997).

Maintenance costs for the road and railway links are based on Ecoplan (2004). Total maintenance costs for the road from Erstfeld to Biasca are estimated at €7 million annually, 27 per cent of which are considered fixed maintenance costs. For the railway link, total maintenance costs are approximately €12 million annually. The share of fixed costs is estimated to be 40 per cent.

The information on *operation costs* is from the operators of the Gotthard road tunnel and the new Lötschberg railway tunnel. We expect total operation costs for the Gotthard road tunnel to be about €18 million/year.[16] Operation costs of the Lötschberg railway tunnel are confidential.

Marginal *external costs* of transport include the costs of air pollution and climate change, noise exposure, and accidents. Noise costs were excluded for the railway sector, since most of the traffic on the link passes through the new tunnels.

Table 9.2. Value of Time (€/Passenger-Hour, €/Net Ton-Hour).

	Passenger Transport		Freight Transport	
	Poor	Rich	Domestic	Transit
Rail	9.27	13.76	1.15	0.98
Road	11.76	17.46	1.15	0.98

Sources: Axhausen, König, Abay, and Rapp Trans (2004, pp. 8, 12), and Maggi et al. (1999, p. 13).

According to the most recent estimates, total *investment costs* of the Lötschberg tunnel are estimated to be €2,836 million.[17] Due to its length and difficult geological conditions, the construction of the Gotthard tunnel will be more expensive. Total final costs are currently estimated at €6,697 million.

2.4.2. Taxes and Charges

Fuel taxes are levied by the federal government. In Switzerland, they are lower than in most other European countries for gasoline, but higher for diesel.[18] Using standard consumption rates and a share of 30 per cent diesel cars, we obtain a value of €0.05/vehicle-km. The tax rate per vehicle-km is assumed to be 25 per cent higher for high-income households. For heavy goods vehicles, the fuel tax is €0.142/vehicle-km.

Foreign and domestic passenger cars have to purchase a *vignette* for driving on national highways. The vignette costs €27 and is valid one year. To obtain a tax rate per kilometre, we have divided the price of the vignette by the annual distance travelled by low-income and high-income households.

Owners of motor vehicles pay an annual *vehicle tax* to the canton (regional government). We assume that high-income households drive vehicles with more engine power, which is charged higher. We use the national average for middle- and upper-class passenger cars (year 2002). The annual tax for passenger cars is €243 for low-income households and €278 for high-income households. For heavy goods vehicles, we use the vehicle tax of the canton of Berne, which was €1,705 in 2004.

As of January 2005, the average HVF rate has been raised to €0.016/ton-km. With an average vehicle weight of 25 tons for domestic heavy goods vehicles and 34.5 tons for transit heavy goods vehicles, this amounts to approximately €36.5 (domestic) and €50 (transit), respectively, for the route from Erstfeld to Biasca.

For an average freight train with a gross weight of 1,360 tons, *track charges* are €5.75/km (including subsidies for combined transport). For a passenger train, track charges are €5.30/km.

2.4.3. Marginal Cost of Public Funds

We use a marginal cost of public funds (MCF) for the labour tax of 1.35 for both the local and the federal government. This value corresponds to the MCF for Switzerland in Kleven and Kreiner (2003) for a proportional tax reform (including benefits).[19]

There is no empirical estimate of the cost of public funds for the HVF in Switzerland. Evidence from other European countries suggests that the MCF of *Pigouvian taxes* is significantly lower than the MCF of labour taxes. In Belgium, for example, the MCF of fuel taxes and passenger peak road pricing is below 0.8, if externalities are taken into account. Without externalities, the MCF is between 1 and 1.2.[20] Brendemoen and Vennemo (1996) obtain a negative value for the MCF including externalities in Norway, and conclude that taxes on gasoline and mineral oil are "by far the cheapest ways of financing a public project."[21]

The level of the Swiss HVF is based on average external costs of transport. It is close to the level of social marginal costs and can thus be considered as an approximation to a Pigouvian tax. Therefore, we use a MCF of 1 for the HVF – a plausible estimate considering the values given in the studies cited above.

3. RESULTS

3.1. Price and Demand Reaction

The scenario with the existing pricing regime and the current earmarking rule is used as the *benchmark scenario*. In this scenario, about 75,000 passenger and freight trains, 6.4 million passenger cars, and about 1 million trucks cross the Gotthard tunnel per year. Trains pay track charges of €462 (passenger trains) and €526 (freight trains) on average for using the infrastructure. Road passengers only have to purchase the vignette once a year; trucks pay the HVF, which is between €37 and €50 for domestic and transit vehicles.[22]

3.1.1. Transport Taxes and Tolls

The transport price on the transalpine corridors comprises transport taxes and tolls (road: HVF, road tolls; rail: track access charges). Table 9.3 shows transport prices and their components for two alternative pricing scenarios with the current earmarking rule as well as the deviations from existing transport prices.[23] Prices in the congestion and infrastructure charging regime are for peak periods.

Under the *existing pricing regime* (scheme A), road pricing is mostly based on taxes, whereas in the railway sector, the only price instrument is the track charge (which we consider as a toll in the model). The higher toll for trucks in transit is due to the fact that vehicles in transit are heavier than domestic

Table 9.3. Taxes and Tolls, Peak Period (Average 2000–2040).

Scenario	User	Taxes	Toll	Total Transport Price	Difference to Existing Pricing	
		€/Vehicle-trip (train, car/truck)			€/v-trip	%
B3: existing taxation with internalisation	Rail: passenger poor	0	496	496	34	7
	Rail: passenger rich	0	496	496	34	7
	Rail: freight domestic	0	701	701	175	33
	Rail: freight transit	0	701	701	175	33
	Road: passenger poor	6	21	27	21	348
	Road: passenger rich	7	21	27	20	300
	Road: freight domestic	12	81	93	44	88
	Road: freight transit	14	83	97	34	53
C3: congestion and infrastructure charging	Rail: passenger poor	0	115	115	−346	−75
	Rail: passenger rich	0	115	115	−346	−75
	Rail: freight domestic	0	287	287	−239	−45
	Rail: freight transit	0	287	287	−239	−45
	Road: passenger poor	0	29	29	24	396
	Road: passenger rich	0	29	29	23	332
	Road: freight domestic	0	108	108	59	119
	Road: freight transit	0	112	112	49	77

trucks on average. If *charges for external costs* are put on top of existing taxes and charges as in scenario B3, transport prices are above their current levels for all users. For railway transport, they increase only modestly (7 per cent and 33 per cent for passenger and freight transport, respectively). For road passenger transport, in contrast, the sum of taxes and charges quadruples (An increase of 300+ per cent is equivalent to a quadrupling; i.e. increase to four times as big.) compared to the benchmark scenario (see Table 9.3). Road freight transport prices increase between 53 and 88 per cent.

Under the congestion and infrastructure charging regime (scheme C), all taxes are replaced by tolls oriented at social marginal costs. In this case, transport prices are lower than today for both passenger trains (−75 per cent) and freight trains (−45 per cent). In the road sector, prices are up to four times higher than today for passenger transport. For road freight transport, prices increase by 119 per cent and 77 per cent for domestic and transit trucks, respectively.

During *off-peak* periods, road tolls are substantially lower than those shown in Table 9.3: With congestion and infrastructure charging, for example, they are €14 instead of €29 for passengers, and €61 instead of €108 for domestic trucks. Track charges are not differentiated between peak and off-peak periods.

Unlike in the other pricing schemes, tolls in the congestion and infrastructure charging regime are computed endogenously for every period. Consequently, they *increase over time* as transport demand grows, whereas under the other pricing regimes, the level of the toll remains constant. If road capacity is expanded, the level of the toll drops in 2015 because the extension of the Gotthard road tunnel reduces congestion.

3.1.2. Number of Trips

The investment in additional railway capacity leads to a stronger growth of transalpine traffic in the railway sector than in the road sector (see Table 9.4). Under the existing pricing regime and with current earmarking (scenario A3), the total number of passengers travelling across the Alps by railway, for example, increases by 66 per cent from 2000 to 2040, compared to a growth of 43 per cent on the road.

Under the second pricing regime (“Existing taxation with internalisation”, scenario B3), the market share of the railways in 2000 is slightly higher than under the current pricing regime for passenger transport, and virtually the same for freight transport (65.6 per cent). As in scenario A3, the modal split improves in the period 2000–2040. With congestion pricing

Table 9.4. Growth and Modal Split of Transalpine Traffic.

Scenario	Mode	Passengers/Net-Tons		Growth	Modal Split (%)	
		2000	2040	(%)	2000	2040
A3	Rail: pass	5,169,000	8,602,400	66	27.1	30.2
	Rail: fret	19,480,900	33,482,400	72	65.7	66.5
	Road: pass	13,936,100	19,871,000	43	72.9	69.8
	Road: fret	10,154,100	16,880,300	66	34.3	33.5
B3	Rail: pass	5,310,300	8,738,600	65	28.7	31.4
	Rail: fret	19,226,600	33,010,500	72	65.6	66.3
	Road: pass	13,214,300	19,114,100	45	71.3	68.6
	Road: fret	10,092,700	16,797,200	66	34.4	33.7
C3	Rail: pass	5,528,000	9,279,600	68	29.0	33.1
	Rail: fret	19,770,100	34,575,000	75	66.1	67.9
	Road: pass	13,501,700	18,786,200	39	71.0	66.9
	Road: fret	10,142,800	16,369,400	61	33.9	32.1

(scenario C3), the growth of railway transport is higher than under the existing pricing, and the market share of the railways reaches 33.1 per cent (passengers) and 67.9 per cent (freight) in 2040.[24]

3.2. Efficiency

3.2.1. Revenue Use: Earmarking and Cross-Subsidisation

Under the current Swiss transport policy, two-thirds of the revenues from the HVF are earmarked for the railway construction fund. Our results indicate that this is an *efficient way* of financing new infrastructure. Under the existing pricing regime, welfare would decrease if a smaller share of the revenues from the HVF were used to finance the new railway tunnels (see Fig. 9.5). Allocating all revenues from the HVF to the railway investment fund would further increase the welfare compared to the current solution.

With given investment, *welfare increases with the magnitude of earmarking*. This result holds for all scenarios. It is due to the fact that the investment in the railway tunnels is financed through an approximation of a Pigouvian tax (the heavy vehicle fee) instead of the labour tax. It is an illustration of the basic principle of (green) tax reforms described by Mayeres and Proost (2001):[25]

> Welfare is increased (reduced) when the tax with the highest MCF is reduced (increased) and when simultaneously the tax with the lowest MCF is raised (reduced).

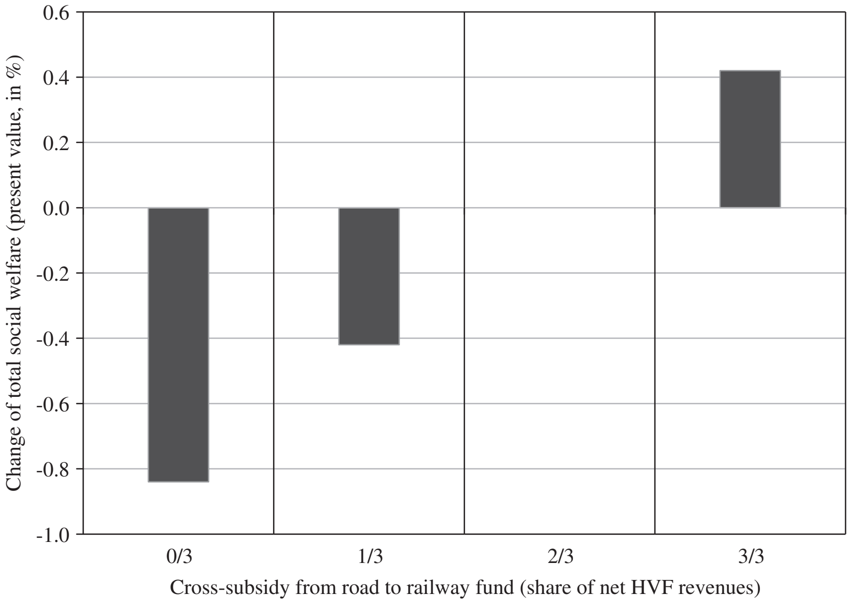

Fig. 9.5. Earmarking of HVF Revenues (with Investment in New Railway Tunnels and Existing Pricing): Change of Total Welfare (Per Cent).[28]

In the present case, the tax with the higher MCF is the labour tax (MCF = 1.35), and the tax with the lower MCF is the HVF (MCF = 1).

Fig. 9.6 shows which *factors* contribute to the change of welfare under different earmarking rules: With a low share of HVF revenues allocated to the railway fund, the federal government needs to raise distortionary labour taxes to finance investments in railway infrastructure.

The higher the share of HVF revenues allocated to the railway fund is, the less "expensive tax money" is needed, and the higher the welfare effects of the federal government's financing of transport infrastructure that are attributed to passengers with the accounting framework. The welfare of the local government, in contrast, decreases with an increasing share of earmarking. If all HVF revenues are allocated to the railway fund (earmarking share of 3/3), none of the HVF revenues are left for the local government, which results in a welfare loss for the local government. Finally, the welfare from transport decreases as a consequence of higher transport prices, regardless of the earmarking rule.

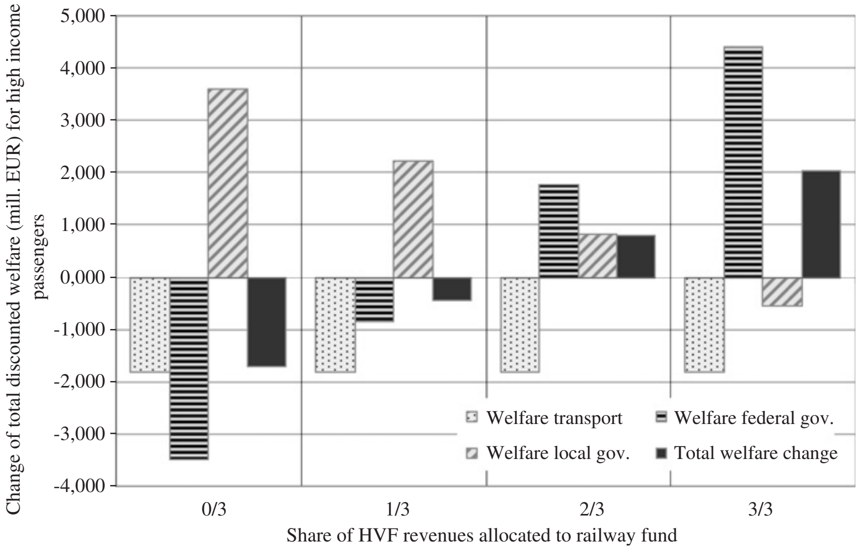

Fig. 9.6. Earmarking of HVF Revenues: Change of Welfare (Million, €) for High-Income Passengers ("Existing Taxation with Internalisation").

3.2.2. Investment: New Railway vs. New Railway and Road Capacity

The results of our analysis suggest that investing in additional capacity in both the road and the railway sectors increases welfare compared to the current strategy of investing in railway capacity only (see Fig. 9.7). The result holds for all other scenarios as well. Fig. 9.7 also shows that *low transport prices increase the pressure to invest*: Under the existing pricing scheme, where no road passenger tolls exist, the welfare gain of extending road capacity is larger than the welfare gain with congestion and infrastructure charging.[26]

Regardless of the earmarking regime, it is better to invest the revenues from transport pricing in additional road capacity than to let them accumulate in the road fund (as it is currently done with fuel tax revenues).[27] This does not imply, however, that the extension of road capacity at Gotthard is the most urgent transport infrastructure project. For a sound judgement about the merits of an extension of road capacity at Gotthard, a comprehensive cost–benefit analysis would have to be carried out.

Sensitivity analysis has shown that the above result is robust to changes of the elasticities of substitution.

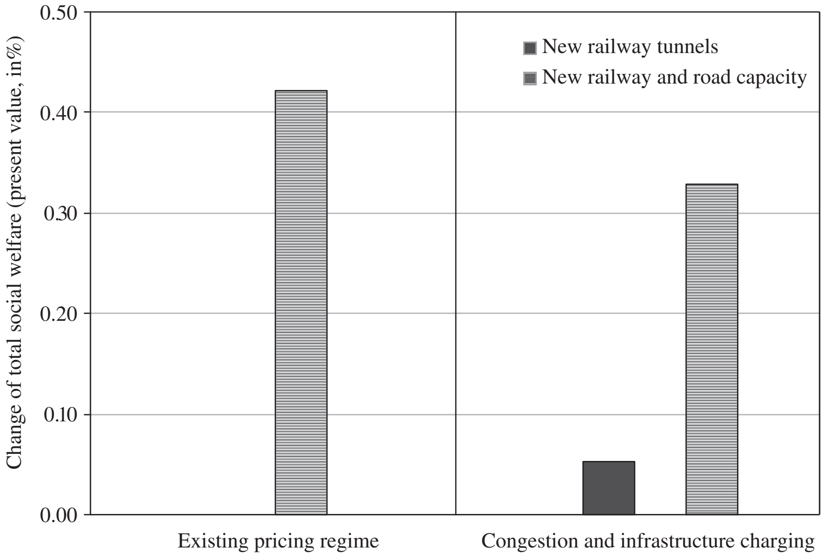

Fig. 9.7. Investment in Railway or Railway and Road (Current Earmarking Regime): Change of Total Welfare (Per Cent).

3.2.3. Pricing Regimes: Revenue and Efficiency Impacts

The pricing regime ''existing taxation with internalisation'' yields the highest social welfare (see Fig. 9.8). In this pricing regime, transport prices in all modes and for all users are above their current level (see Table 9.3). With congestion and infrastructure charging, average road prices are substantially higher than today (and higher than in the scenario ''existing taxation with internalisation''), whereas railway prices are below their current level. Thus, an increase of transport prices in all modes is better than an increase of road transport prices only.

The good performance of the ''existing pricing with internalisation'' pricing regime is due to the *welfare from the use of toll revenues*. Both federal and local governments benefit from the higher revenues associated with alternative pricing regimes, which help them finance transport investments and result in significant welfare gains (see Fig. 9.9). However, tolls are above their optimal levels and produce a loss of welfare from transport. Since the welfare gains for the government (less expensive financing of transport investment) exceed the welfare loss of transport users, total welfare is higher under alternative pricing regimes.

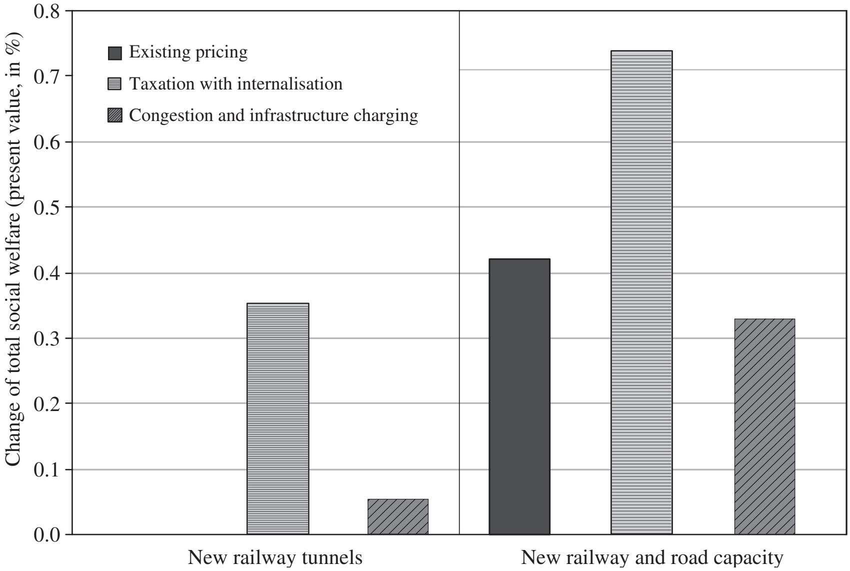

Fig. 9.8. Pricing Regimes (Current Earmarking): Change of Total Welfare (Per Cent).

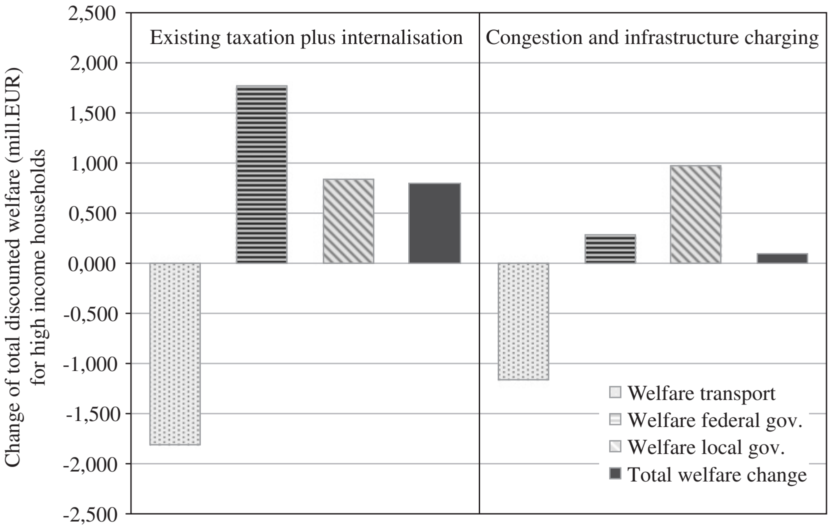

Fig. 9.9. Pricing: Change of Welfare (Million, €) for High-Income Passengers (Current Earmarking).

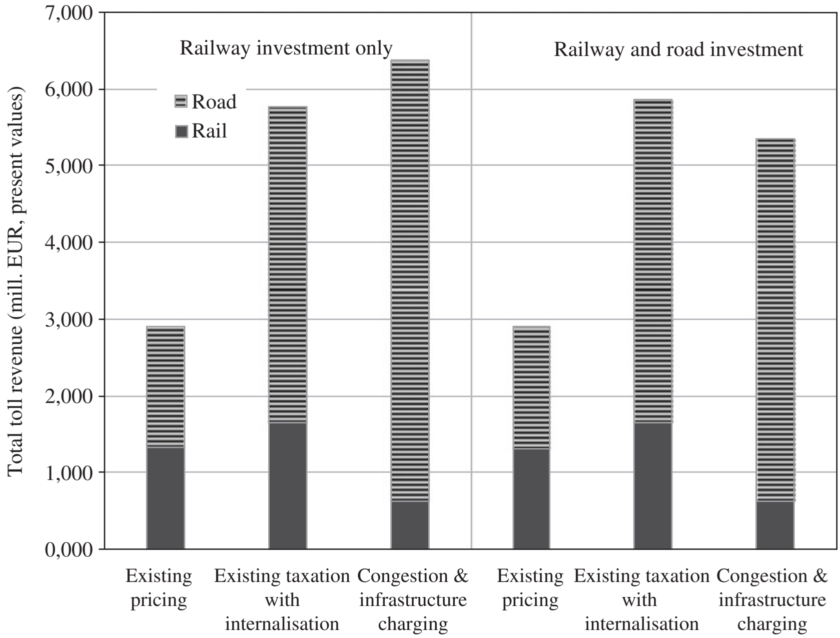

Fig. 9.10. Total Toll Revenue (Million, €, Present Value).

Toll revenues from the road and railway sectors are €2.9 billion under the existing pricing regime with railway investment, €5.7 billion in the "existing pricing with internalisation" scenario, and €6.4 billion in the "congestion and infrastructure charging" scenario (see Fig. 9.10).

In the scenarios where the Gotthard road tunnel is extended from one to two lanes, total toll revenues are equal in the "existing pricing" and "existing taxation with internalisation" scenarios. Since we use the same growth rate for transport volumes in both scenarios, this result could be expected. In the "congestion and infrastructure charging" scenario, total revenues decrease. The reason is that the new road infrastructure leads to a reduction in congestion and therefore to a lower level of the congestion charge. This in turn results in less revenue from the charging scheme.

3.3. Equity

3.3.1. Passenger Transport

Under the current earmarking scheme (two-thirds of HVF to railway fund), the introduction of alternative pricing regimes would benefit both income

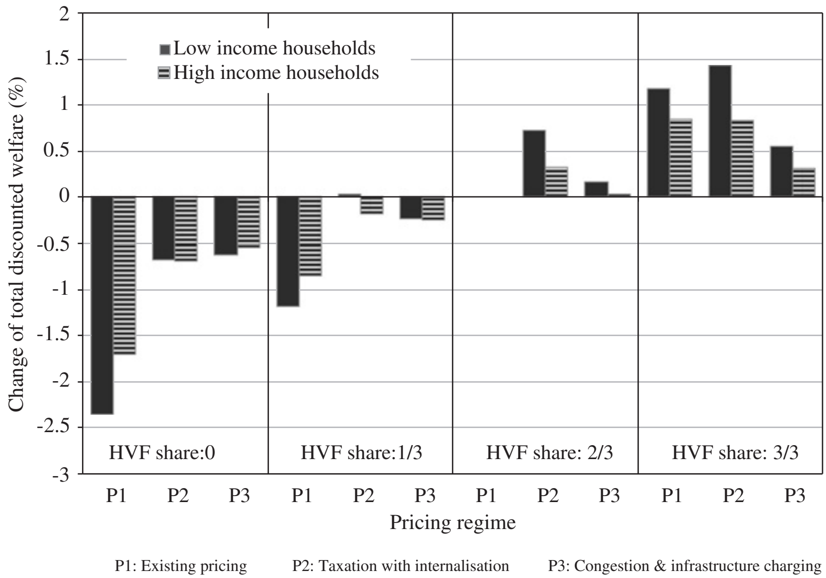

Fig. 9.11. Passenger Transport (Rail and Road): Welfare Change (Per Cent).

groups; low-income households benefit slightly more than high-income households (see Fig. 9.11). A similar pattern can be found in the full earmarking scheme (3/3). If less HVF revenues are earmarked for the railway fund than today (0/3 or 1/3), low-income households are affected more negatively than high-income households under the existing pricing regime. For the other pricing regimes, both income groups are affected about equally.

Increasing cross-subsidisation thus benefits low-income households more than high-income households; reducing cross-subsidisation has the opposite effect. The reason is that the government's welfare gain attributed to households represents a larger share of total welfare for low-income households than for high-income households. But overall, there are only *small differences* between the welfare impacts on low-income households and high-income households.

3.3.2. Freight Transport

While private passengers would gain from the introduction of pricing schemes oriented at social marginal costs, our results indicate that freight transport is likely to lose from higher transport prices. This is because consumers benefit from the toll revenues, whereas in our model, freight

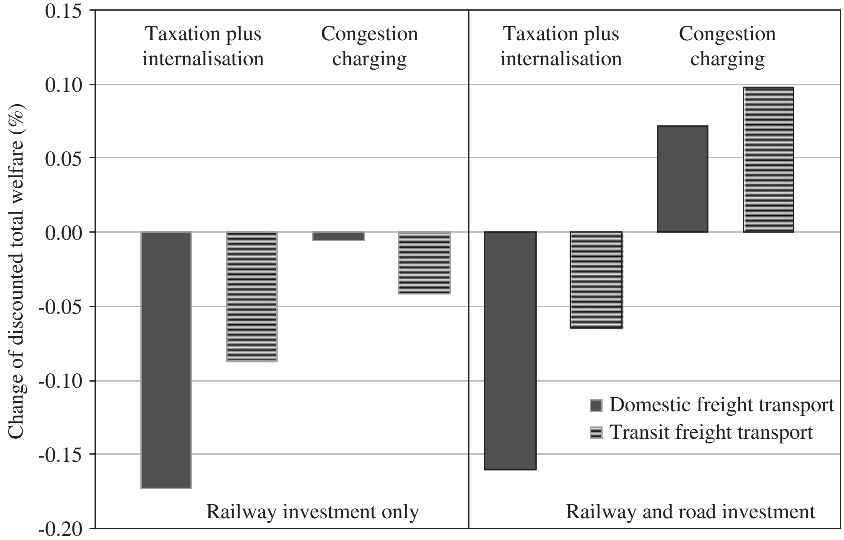

Fig. 9.12. Freight Transport (Rail and Road): Change of Welfare (Per Cent).

transport is not allocated any welfare benefits from the government's reduced financing costs.

The results of MOLINOinGAMS suggest that transit freight transport would be less affected than domestic freight transport by the introduction of alternative pricing regimes (see Fig. 9.12): In three out of four cases, the welfare impact on domestic freight transport is more negative than the impact on transit freight transport. The reason for this is that the total (road) transport price increases more for domestic than for transit transport compared to the benchmark (see Table 9.3). The current HVF is levied according to maximum permissible weight, which is higher in transit traffic on average, whereas the alternative pricing regimes in our model do not distinguish between the weights of trucks. Consequently, the price increase is larger for the lighter trucks in domestic transport.

4. CONCLUSIONS

The results of our analysis can be summarised as follows:

- *Social welfare increases with cross subsidisation from the road to the railway sector*. Thus, the current Swiss policy of using two-thirds of the revenues from the HVF is an efficient way to finance new railway infrastructure. However, welfare would be even higher if all HVF revenues were earmarked for the railway investment fund. The main reason for this result is the low marginal cost of public funds of the HVF.

- *Raising transport prices above marginal social costs may increase welfare*, if the revenues from transport pricing are used to cut existing distortionary taxes. Moreover, the results of MOLINOinGAMS indicate that it is better to introduce congestion charges in addition to existing taxes (such as the fuel tax) than to replace existing taxes.
- Currently, investment in transalpine transport infrastructure is limited to the railway sector. Our results indicate that social welfare could be increased if investment were targeted at *both railway and the road infrastructure* in the transalpine corridors.
- The introduction of a more efficient transport financing system would *not put low income households at a disadvantage* if welfare gains are distributed equally among low-income and high-income households. The differences between the welfare impacts on the two groups of households are small. In the freight sector, transit transport is less affected by higher transport prices than domestic traffic.

For several reasons, the above results should be interpreted with caution. First, the analysis is restricted to the stretches of the transalpine corridors where the investment takes place. Consequently, the effects in the rest of Switzerland are not taken into account. Second, the results of the welfare analysis are sensitive to the choice of the marginal cost of public funds. Since there is no empirical estimate of the marginal cost of the Swiss HVF, we use a plausible estimate based on other European studies. Third, we assume that if HVF revenues allocated to the road fund were not invested in the extension of the Gotthard tunnel, they would be kept in the road fund. While this is happening in fact with the revenues from the fuel tax, it does not necessarily have to be the case for HVF revenues. Finally, economy-wide effects of transport pricing and revenue use cannot be fully captured by partial equilibrium models such as MOLINO. Therefore, the impacts of different transport pricing and revenue use schemes should ideally be analysed using a computable general equilibrium model. Such a model would also make it possible to examine scenarios where all HVF revenues are used to reduce labour or other taxes instead of being invested in new transport infrastructure.

Two main policy recommendations emerge from the preceding analysis:

1. *Increasing transport prices and using revenues from road pricing to finance investments in other modes can be welfare improving*. Naturally, this

depends on the specific situation of a country. In Switzerland, revenues from an efficient road pricing instrument have made it possible to undertake measures (i.e. construct new railway tunnels) to address the severe problems caused by the rapid growth of transit traffic.

2. *Transport pricing, investment, and revenue use should be considered together*. Our case study shows, for example, that for a given investment path, the introduction of congestion charging decreases social welfare in some scenarios. Thus, optimising only one component of transport policy and leaving the others unchanged may in fact worsen the situation.

NOTES

1. See ARE (2001, p. 38).
2. See ARE (2004, p. 2).
3. For a description and analysis, see Borgnolo, Stewart-Ladewig, and Neuenschwander (2005) and ARE (2002).
4. FINÖV is the abbreviation of the name of the corresponding decree "Bundesbeschluss über Bau und Finanzierung von Infrastrukturvorhaben des öffentlichen Verkehrs".
5. See Chapter 5 of this volume.
6. Thun is located 8 km north of Frutigen.
7. A pure social marginal cost pricing approach would exclusively consist of the price-relevant costs (marginal infrastructure costs, marginal user costs (congestion) and marginal external costs). Scheme B, however, also contains charges (fuel tax and railway track charges) contributing to cover total infrastructure costs. The marginal external costs and the marginal congestions costs are added on top of them.
8. Except for the environmental externalities, they would probably not vary much anyway.
9. The values of σ_s are taken from EC, S& P, and K.U. Leuven (1999, p. 7), and the values of σ_f from ECMT (2003, p. 86).
10. See De Borger and Proost (2001, p. 86).
11. β_1 is the share of domestic freight transport allocated to low-income households, $(1-\beta_1)$ the corresponding share allocated to high-income households. β_2 is the weight given to transit freight transport. β_1 equals 0.18, which corresponds to the share of low-income households in the total population. β_2 is set to 1.
12. This does not imply, though, that the regulation schemes produce welfare losses. The only relevant indicator is the change of welfare between scenarios.
13. We assume a constant flow of revenues from the HVF collected on all Swiss motorways to the railway fund, even though total revenues would actually vary with the level of the toll. Since the model is limited to the transalpine corridors, we cannot model the change in travel behaviour and toll revenues in the rest of Switzerland.
14. See Kommission für Verkehr und Fernmeldewesen (2002, p. 13).
15. ASTRA (2003).

16. The operation of the Gotthard road tunnel costs approximately €8 million annually (http://www.gotthard-strassentunnel.ch). To this figure, we have added another €10 million for the operation of the roads in the north and in the south of the tunnel.
17. BAV/Alptransit (2004), p. 7.
18. Source: Erdölvereinigung (2004, p. 76).
19. See Kleven and Kreiner (2003, p. 15).
20. See Mayeres (1998, pp. 26–27).
21. See Brendemoen and Vennemo (1996).
22. The HVF is equal for all heavy vehicles and does not discriminate against transit traffic. The higher HVF in the model is due to the fact that trucks in transit are heavier than domestic trucks on average.
23. For a more detailed exposition of results, see Cretegny, Springer, and Suter (2005).
24. Note that the modal split in 2000 is different from the benchmark case. This is due to the fact that the model determines the optimal division between transport sectors based on relative prices, which vary between pricing scenarios.
25. See Mayeres and Proost (2001, p. 347).
26. The investment in railway tunnels does not produce a change of welfare in the existing pricing scheme because it is part of the reference scenario.
27. All revenues from the HVF that are allocated to the road fund are assumed to be either left in the fund or spent to increase road capacity at the Gotthard.
28. Total social welfare is always compared to the benchmark scenario (A3).

ACKNOWLEDGMENTS

Financial support from the European Commission (EU Fifth Framework research project REVENUE (DG TREN, Contract No. GMA2-2002-52011)) and helpful comments from the editors are gratefully acknowledged.

REFERENCES

ARE Bundesamt für Raumentwicklung (2001). Wege durch die Alpen. Alpenquerender Güterverkehr auf Strasse und Schiene, Bern.

ARE Bundesamt für Raumentwicklung (2002). Fair and efficient. The distance-related heavy vehicle fee (HVF) in Switzerland, Berne.

ARE Bundesamt für Raumentwicklung (2004). Alpinfo 2003. Alpenquerender Güterverkehr auf Strasse und Schiene, Bern.

ASTRA Bundesamt für Strassen (2003). Automatische Strassenverkehrszählung, Januar 2002.

Axhausen K. W., König, A., Abay, G., & Rapp Trans, A. G. (2004). Zeitkostenansätze im Personenverkehr. Forschungsauftrag Nr. 2001/534 auf Antrag der Vereinigung Schweizerischer Verkehrsingenieure (SVI), Zürich.

BAV Bundesamt für Verkehr/Alptransit (2004). NEAT Neue Eisenbahn-Alpentransversale – Kurzfassung, Bern.

Borgnolo, Ch., Stewart-Ladewig, L., & Neuenschwander, R. (2005). Position and recent trends in European countries. In: J. M. Viegas (Ed.), *Interurban road charging for trucks in Europe* (pp. 109–132). Oxford: Elsevier.

Brendemoen, A., & Vennemo, H. (1996). The marginal cost of public funds in the presence of environmental externalities. *Scandinavian Journal of Economics*, *98*(3), 405–422.

Brooke, A., Kendrick, D. A., Meeraus, A., & Raman, R. (1998). GAMS: A users guide. Available at www.gams.com

Cretegny, L., Springer, U., & Suter, S. (2005). *Case study Switzerland. REVENUE project deliverable 4*. Funded by 5th Framework RTD Programme. ISIS, Rome.

De Borger, B., & Proost, S. (Eds). (2001). *Reforming transport pricing in the European Union*. Cheltenham: Edward Elgar.

EC European Commission, S&P Standard & Poor's, & Leuven, K. U. (1999). The AOP II cost-effectiveness study, Part II: The TREMOVE Model 1.3 [The (deleted) reference in footnote 9, page 9, refers to Section 3.1], Annexes. Draft Final Report presented to Working Group 7, August.

ECMT European Conference of Ministers of Transport (2003). Efficient transport taxes and charges 2003. (CEMT/CS(2003)4), Paris.

Ecoplan (1997). Wirtschaftlichkeitsstudie NEAT. Studie im Auftrag des Bundesamts für Verkehr, Bern.

Ecoplan (2004). Unterhaltskosten von Schienen- und Strassen-Infrastrukturprojekten. Überprüfung der Kostensätze für Zweckmässigkeitsbeurteilungen. Studie im Auftrag des Bundesamtes für Verkehr, Bern.

Erdölvereinigung (2004). Jahresbericht 2003, Zürich.

Kleven, H. J., & Kreiner, C. T. (2003). *The marginal cost of public funds in OECD Countries: Hours of work vs. labor force participation*. CESifo Working Paper no. 935.

Kommission für Verkehr und Fernmeldewesen (2002). Bericht Parlamentarische Initiative Gotthard-Strassentunnel/A2, Bern.

Maggi, R., Bolis, S., Maibach, M., Rossera, F., Rudel, R., & Schreyer, Ch. (1999). Multimodale Potenziale im transalpinen Güterverkehr. Bericht B4. NFP 41 Verkehr und Umwelt, Wechselwirkungen Schweiz-Europa, Bern.

Mayeres, I. (1998). *The distributional impacts of policies for the control of transport externalities: An applied general equilibrium model*. FEEM Working Paper no. 8.99, Milan.

Mayeres, I., & Proost, S. (2001). Marginal tax reform, externalities and income distribution. *Journal of Public Economics*, *79*, 343–363.

TCS Touring Club Suisse (2004). Verbrauchskatalog 2004, Bern.

This page intentionally left blank

CHAPTER 10

THE GERMAN HGV MOTORWAY TOLL

Claus Doll and Heike Link

ABSTRACT

This chapter explores the feasibility and possible benefits of amending the German Heavy Goods Vehicle (HGV) motorway toll system by adopting different pricing and revenue use rules. Economic efficiency and equity effects are assessed using the MOLINO partial equilibrium model and the ASTRA system dynamics model. Acceptability issues are also addressed using the results of a survey of actors in the German haulage business. According to first-best welfare theory user charges should be set equal to social marginal costs and the revenues should accrue entirely to the state. However, acceptability and long-term development considerations militate strongly in favour of earmarking revenues to the transport sector.

1. INTRODUCTION

In September 2000, the Commission on transport infrastructure financing set up by the German government published its final report (Pällmann, 2000). There the Commission strongly recommended making the existing tax-based infrastructure financing systems suitable for the future by promoting the user-pays principle. This recommendation was, among other reasons, motivated by the enormous investment deficit in the federal

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 217–241

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19010-7

transport networks that inexorably grows and will eventually lead to considerable losses in service quality. Especially in periods of tight public budgets, the tax-based financing mechanisms will be unable to rectify these investment deficits and thus will create even higher obligations for future generations.

For the federal road network, the Commission proposed to introduce user charges for heavy goods vehicles (HGV) above 12 tonnes gross vehicle weight (GVW), and to follow later with light goods vehicles and passenger cars. The federal government adopted these recommendations and immediately tendered the electronic toll collection system, and in parallel launched a study calculating the HGV motorway tariffs (Rommerskirchen et al., 2002). In summer 2001, the federal cabinet decided to replace the time-based Eurovignette-System with an electronic distance-based motorway toll by approving the draft Motorway Toll Act (Bundesregierung, 2001). However, the Act had to be ratified not only by the parliament, but also by the council of member states (Bundesrat). There was heated discussion about how to use the revenues from the HGV motorway toll and to what degree hauliers should be compensated for the double burden imposed by the toll and fuel taxes. In March 2002, the government was finally able to assert its position and to put the Motorway Toll Act into effect (Bundesregierung, 2004). The Act guarantees compensation of €300 million to German hauliers and stipulates that the revenues be used as follows:

- The compensation is realised by lowering the originally proposed charge level of €15.00/100 km–€12.40/100 km for vehicles above 12 tonnes GVW. The charges are differentiated according to the number of axles and exhaust emission standards ranging from €9.00 to €14.00/100 km.
- Roughly, 20 per cent of the revenues are granted to the toll operator for operating the charging technology.
- The remaining 80 per cent of revenues are allocated solely to the federal transport networks. From these revenues 50 per cent is devoted to the federal road network (which consists primarily of motorways), 38 per cent is for the federal rail network and 12 per cent is for inland waterways (IWW).
- Allocation of revenues to particular investment projects is carried out by the state-owned Transport Infrastructure Financing Society (Verkehrsinfrastruktur-finanzierungs-Gesellschaft VIFG), which was founded in October 2003.

This schedule was very ambitious and it is not surprising that the toll system could not be launched by the planned starting date of September 2003. Because of technical and managerial problems the introduction was postponed several times, and there were serious doubts that the system would ever successfully go into operation. While the first attempts in September 2003 and in January 2004 were abandoned, the Toll Collect consortium eventually managed to launch the HGV motorway toll system in January 2005 without further hitches. Thanks to the guaranteed earmarking of revenues to the federal transport networks and the compensation measures (which are still being argued by the EC and the German government) the scheme is supported by hauliers' organisations. After one year in operation, in 2005, revenues of €2.86 billion were collected. Furthermore, on the initiative of a number of German federal states (länder) and in the light of the enormous public deficit, after the parliamentary elections in September 2005, the Federal Government began exploring ways to introduce motorway tolls for passenger cars.

One disadvantage of the toll system is that many lorries are avoiding the tolled motorways by using toll-free federal roads. As this is causing serious environmental and safety problems on some sections of the federal road network (Bundestag, 2006), the government has empowered the German states to close affected road sections with heavy traffic. Moreover, in the short run the toll has not induced traffic to shift from road to rail or waterways because the charge level is not high enough to offset the general competitive advantage of road haulage against competing modes in terms of delivery time, costs and flexibility. Thus, the system is not helping to achieve the Kyoto targets as had been envisaged by the Federal Government's climate change programme 2000 (BMU, 2000). Environmental organisations, for example, call for higher charges and for an extension of the toll system to the secondary road network as has been implemented in Switzerland (Balmer, 2003). But the rationale for such measures is questionable on both acceptability and efficiency grounds. First, resistance against substantially higher charges from the haulage business might result in growing transaction costs associated with control and enforcement against toll evasion. Second, poorly developed charge structures might lead to adverse environmental, safety or equity impacts. The ways in which additional revenues are used is an important consideration in addressing both of these concerns.

Against this background, this chapter presents the results of two case studies of the German HGV charging scheme. The first case study used two modelling tools. The MOLINO model (described in Chapter 5 of this volume) was applied to study the welfare impacts of different charging options in combination with different options for using the revenues. And a system dynamics model (ASTRA) was used to analyse the longer-term economic impacts of charging on GDP and employment. The second case study carried out an Internet-based survey of German freight hauliers, shippers and logistic providers to examine attitudes towards various pricing and regulation schemes within the German haulage business. It provides insights into such dimensions of acceptability as problem perception, perception of the effectiveness of charging and outcome beliefs.

The main conclusion from the modelling study is that a system of HGV charges based on social marginal cost with the revenues allocated to the general budget is welfare superior to other schemes. However, the acceptability study indicates that such a policy would not be acceptable to those affected by charging. Indeed, there appears to be a trade-off between the charge level and the use of revenues in the sense that road hauliers would accept higher charges in return for a guarantee that revenues were used in the road sector. And the modelling study indicates that if toll revenues are earmarked to the transport sector, as it is current practice in Germany, it is probably more efficient to use them for the road sector rather than for rail and inland waterways.

The modelling case study is presented in Section 12.2 and the acceptability study in Section 12.3. Section 12.4 draws conclusions.

2. ECONOMIC EFFICIENCY OF HGV CHARGING AND REVENUE-USE SCHEMES

2.1. The Assessment Framework

Two modelling tools were used to assess the economic efficiency of alternative HGV charging and revenue-use schemes in Germany. One is the MOLINO model which was developed within the REVENUE project (Chapter 5 of this volume) and the other is the ASTRA system dynamics model which is described in detail in Schade (2005). The MOLINO model focused on microeconomic aspects of price setting, fund design and revenue use, while the ASTRA model was used to examine macroeconomic interactions between the transport sector and the rest of the economy over a longer period of time. Although the two models were applied to similar sets of scenarios the models are very different in structure, and a comparison of their results is of independent interest from a modelling standpoint. The models investigated four settings of motorway tolls and revenue spending against the situation in Germany prior to the introduction of the Toll Collect system.

2.1.1. Setup of the MOLINO Model

To adequately reflect the decision-making structure, responsibilities and financial flows underlying transport infrastructure pricing in Germany, the actors within the MOLINO model were set up as follows (Fig. 10.1 depicts the flow of funds in the MOLINO setup that is used):

- Roads and a combination of rail and inland waterways are treated as competing modes.[1] The infrastructure of each mode is owned and managed by separate actors.
- Transport users and service operators are treated jointly as final users who pay taxes to the state and charges to the infrastructure operators, and who generate external costs of accidents, noise, air pollution and climate change to society. Depending on the scenarios analysed, these charges are either externally computed as average-cost prices or computed internally by MOLINO as marginal external cost charges. In the latter case, the user prices equal the marginal social costs.
- The infrastructure operators administer the infrastructure and undertake maintenance and renewal, and pass on their costs to final users via average or marginal social cost pricing.
- Revenues from pricing go to a transport infrastructure fund which can provide grants to the infrastructure operators.
- The infrastructure owner is responsible for providing new capacity. The level of investment consists of a minimum level financed by the state and an additional level that is a function of the grants received from the infrastructure fund. The development of the infrastructure capacity in each of the competing modes is thus driven by the money flowing into the fund and by the fund allocation rules.

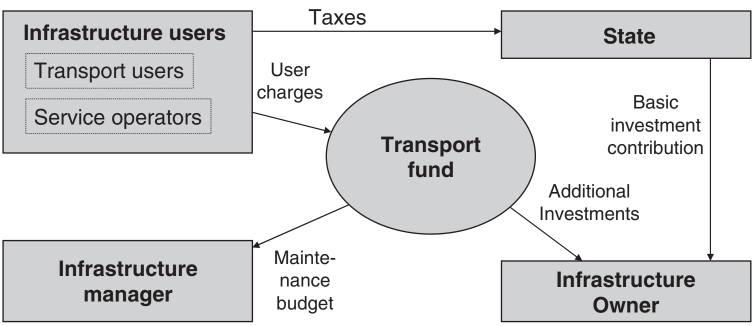

Fig. 10.1. Schematic Flow of Funds in MOLINO. *Source:* Link and Stewart (2005).

The MOLINO model was configured to study alternative revenue allocations between investment and maintenance, between road and rail/IWW and between transport and the public budget. In both modes, passenger and freight transport were considered, each having individual pricing structures.

Each model configuration was calibrated to the transport conditions prevailing in Germany in the year 2000 using data from the general investment plan (BMVBW, 2003), available statistical sources (BMVBW, 2005; KBA, 2004) and business reports of the German rail network operator DB-Netz AG (DB, 2004). Network-wide speed-flow curves were used to compute marginal social congestion charges and the benefits of investment activities. These curves were defined using the network database model developed in NEA, IWW, and COWI et al. (2003) and by analysing the effects of big projects in the rail sector. Since speed-flow functions for railway networks are commonly not available they were derived by examining the impacts of the new high-speed Frankfurt–Cologne and Hanover–Würzburg lines on the network-wide average speeds of passenger and freight services. Elasticities of substitution, which were found to have a major impact on the model results, were chosen very carefully by calibrating MOLINO using transport-pricing impact studies (e.g. Kleist & Doll, 2005; Rothengatter & Doll, 2002; Gresser, Hölsken, & Kienzler, 2001).

Finally, a value for the marginal costs of public funds (MCPF) of 1.58 was derived as an average of 1.70 for income tax, 1.35 for average cost charge and fuel tax and of 1.00 for marginal social cost pricing, weighted by the respective revenues.[2] These figures are taken from an international comparison of different types of tax reforms in Kleven and Kreiner (2003). The users have been classified into low-income and high-income categories according to the survey ''Mobility in Germany'' (Follmer, Kunert, Kloas, & Kuhfeld, 2004). Low-income households are defined to be those with a net income below €2,000 per month.

The MOLINO model was run over the period 2000–2020. The results are reported as the discounted sum of annual social welfare, using a discount rate of 0.025, as well as separately for low-income and high-income users. The social welfare function considers the costs of the users (transport operators, firms and consumers) including time costs and therefore congestion effects, the external costs other than congestion (environmental costs and part of accident costs), the net revenues of the government (weighted by the MCPF), the net surplus of the infrastructure providers and the net surplus of the Fund. For each scenario, payments from the transport fund to the modal networks or public purse were constrained so that the balance in the fund is zero in 2020. This assures that the welfare measures are not influenced by the existence of a surplus or deficit at the end of the planning horizon.

2.1.2. Brief Description of the ASTRA System Dynamics Model

In recognition of the complexity of the interrelationship between transport policy and economic development in a changing environment, the MOLINO results were benchmarked with some outputs of the ASTRA system dynamics model. ASTRA was developed over the course of several EC studies spanning the 4th–6th RTD framework programmes. The system dynamics approach is based on evolutionary economics and is designed to show the development path of complex systems and their response to changing conditions. The version of the ASTRA model used in this study covers the EU-15 member states, of which only the outputs for Germany were used. Each country is divided into four functional zones, of which each embraces all regions with similar settlement structures. The transport sector is disaggregated into several modes, commodities, travel purposes and distance bands and is calibrated at observed transport flows between 250 counties and functional zones. The macro-economic module consists of input–output tables for 25 sectors. Other modules deal with population, the government, trade, regional development, transport demand and distribution, the vehicle fleet, environment and welfare assessment. For the current analyses the model was run until the year 2020. A detailed description of the ASTRA model is given in Schade (2005).

Although the ASTRA model contains far more detail than the MOLINO model, ASTRA cannot model transport funds and it cannot be used to compute neoclassical welfare measures. Thus, the model was only used for a limited number of pricing and revenue-use alternatives. Apart from the business-as-usual case, which describes the member states' transport policies according to current plans, only the pricing scenario of charging all road vehicles on all network plans is analysed. Within this framework three alternatives for revenue spending are considered:

- *Road scenario*: All revenues are earmarked for the road sector, with half for motorways and half for trunk roads. Within each road category 80 per cent of revenues are spent for capacity extension measures, including the construction of new roads and the widening of existing roads. The remaining expenditures are classified as maintenance works, including replacement, renewal, repair and operation activities.
- *Cross scenario*: Revenues are allocated equally to road and to rail transport. Within the railway sector 60 per cent of the revenues are invested in

new network capacity, 20 per cent in facilities (e.g. inter-modal terminals) and 20 per cent in rolling stock.

- *DT-scenario (for direct taxes)*: Revenues are transferred to the general budget and used to decrease direct taxes so that all benefits accrue to consumers.

2.2. Scenarios and Results

The analysis of efficiency and equity issues in the German HGV motorway toll case was organised around four research questions. Out of these, the central concerns were how to allocate revenues between transport modes (Research Question 2) and between the transport sector and the general budget (Research Question 3). These questions, the corresponding model settings and the resulting model outputs are considered or addressed in turn.

2.2.1. Maintenance versus New Construction

2.2.1.1. Research Question 1. If funds are earmarked wholly to the transport sector (50 per cent road, 38 per cent rail and 12 per cent for IWW), what fraction should be allocated to capacity extension, and what fraction to maintenance, replacement and repair?

This question was addressed within the currently applied average cost pricing framework in which a toll of €12.4/100 km is charged on motorways for HGVs. The toll level is somewhat lower than the average infrastructure costs (€15/100 km) because part of the fuel tax revenue is already dedicated to finance the federal road network. Charges for rail and inland navigation are below average costs as well, and therefore these modes receive additional financial contributions and investment aids from the government. According to prevailing rules, half of the toll revenues are distributed to road and the other half to rail and inland waterways.

For each mode the share of revenue allocated to maintenance, replacement and repair activities was increased from 25 to 75 per cent with the remainder used for new construction. Compared to the reference scenario, which describes the situation before 2005, the introduction of the HGV toll scheme led to a loss of social welfare. This loss ranges from €59 million when 75 per cent is devoted to maintenance to €72 million when only 25 per cent goes to maintenance. From the social point of view maintenance is preferred to construction because in the long run new roads create substantial problems with induced traffic, its related environmental burden and with long-term maintenance obligations. These results confirm the position of the Governmental Commission on Transport Infrastructure Financing

(Pällmann, 2000), which explicitly put a priority for repair and renewal of road infrastructure over the provision of new capacity.

The model reveals that, in contrast to the impacts on overall welfare, transport users benefit from introducing the HGV motorway toll scheme and they are better off with more investments in new capacity because this boosts travel speeds. Decreasing the share of revenues devoted to maintenance activities from 75 to 25 per cent, and thus increasing the proportion of new construction activities from 25 to 75 per cent, led to a decrease of total user benefits by €16 million from €10.9 million to €4.9 million. On the other hand, the rest of society, including inhabitants who suffer from local emissions, climate change, accidents and noise pollution, would suffer a larger reduction in welfare of €29 million. Compared to the total annual revenues from the HGV charging system of €2.9 billion, these changes in social welfare are rather small.

2.2.2. The Level of Cross-Subsidisation

2.2.2.1. Research Question 2. If funds are earmarked to the transport sector, how should they be allocated between modes?

This question was addressed using both MOLINO and ASTRA. Within the MOLINO model the share of revenues allocated to roads was varied from 100 per cent down to 25 per cent, with the remainder assigned to rail and inland waterways. Investment levels were set externally without making use of MOLINO's optimal investment function. In all cases, the HGV toll revenues were earmarked to the transport sector and within each mode the revenues were allocated in equal shares to new construction and maintenance. In addition, cross-subsidisation of rail and inland waterways was addressed using the ASTRA model by comparing the *Road* scenario and *Cross* scenario defined earlier.

The MOLINO model results are shown in Fig. 10.2. As the share of revenues devoted to rail and inland waterways increases, total social welfare and the surplus of both income groups steadily decrease. Total social welfare decreases by €9.2 million. User benefits fall by €14.2 million, while other social groups – including the state, network operators and inhabitants – gain by roughly €5 million.

The finding that investments are more productive in the road sector is explained by two facts: (a) expanding the network capacity by a certain proportion is cheaper in the road sector and (b) roads are more heavily used and hence more users benefit from road improvement

The higher benefits of directing grants to the road sector is confirmed by the results of the ASTRA model. But the differences between earmarking

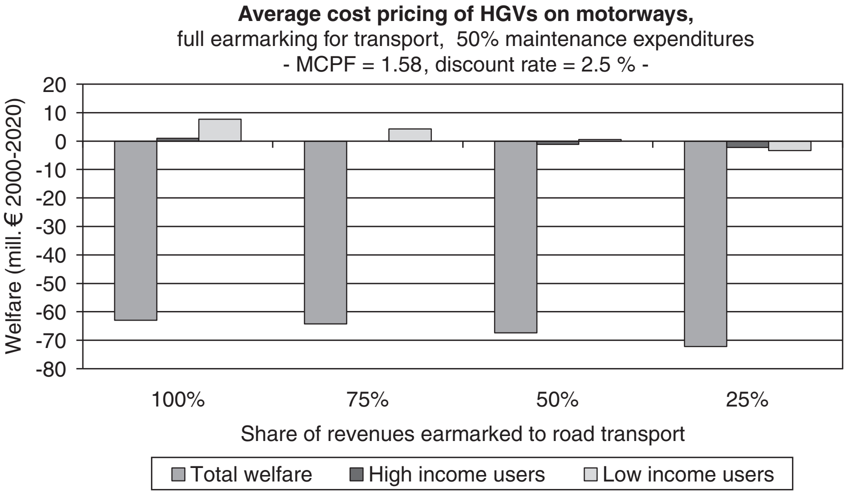

Fig. 10.2. Effect of Cross-Subsidies from Road to Rail and IWW. *Source:* Link and Stewart (2005).

the revenues to the road sector and cross-subsidising rail appear rather small. The simulation runs indicate that the higher cost efficiency in the road sector outweighs the environmental benefits from demand shifts to rail.

2.2.3. The Share of General Budget Contributions

2.2.3.1. Research Question 3. If no earmarking rules are specified, what fraction of funds should be allocated to the transport sector, and what fraction should be assigned to the general budget?

This research question was again analysed within the context of the current charging and regulation scheme using both the MOLINO and ASTRA models. Before discussing the results it should be noted that the revenues from the HGV motorway toll are additional to the government's basic financing levels for road, rail and inland waterways. Both models incorporate this fact by setting basic financing levels for each mode, which remain constant over time and which are not affected by changes in the size or the allocation rules of the transport fund. Thus, tax levels are not affected by the design of the toll scheme.

For the MOLINO simulations, the share of funds allocated to the general budget was varied from 0 to 100 per cent. Half the resources devoted to the transport fund are assumed to be used to subsidise rail (38 per cent of total revenues) and inland waterways (12 per cent of total revenues), and within each mode 50 per cent is used for investment in new capacity and the remainder for maintenance.

The ASTRA model compared the two polar *Cross* and *Direct Taxes* scenarios. For both models it was assumed that the central government uses the toll revenues to lower direct taxes. Alternative forms of spending public money, e.g. for education or health care, are not considered.

The results of the MOLINO and ASTRA models differ sharply on Research question 3. According to the MOLINO model, allocating revenues to the general budget rather than to the transport sector is much preferred for society as a whole, as well as for users. This result is driven by the assumption that the state is both omniscient and benevolent and capable of maximising global social welfare. The transport fund, in contrast, is capable only of allocating funds within the transport sector, and earmarking scenarios thus diminish the state's freedom of action.

A comparison of Figs. 10.2 and 10.3 shows that, according to the MOLINO model results, welfare is much more sensitive to how much revenues are earmarked to the transport sector than to how a given amount of revenues is distributed among modes and types of investment activities. The MOLINO model suggests that using revenues to lower taxes is more important than deciding how to improve transport conditions. This result is driven by the

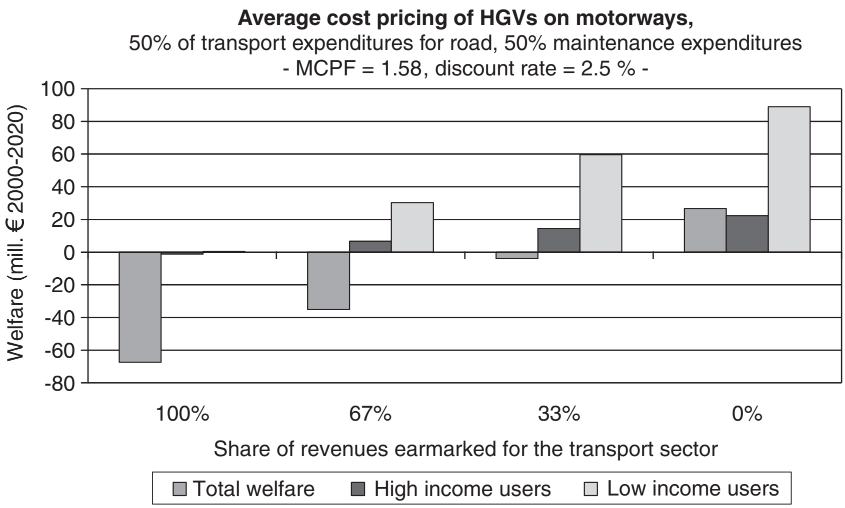

Fig. 10.3. Earmarking to Transport versus Public Sector Support (MOLINO Model Results). *Source:* Link and Stewart (2005).

comparably high value of the MCPF of 1.58: any EURO transferred from the transport sector and or investments generates €0.58 of extra welfare. According to the current tax structure in Germany, high-income users profit most from the redistribution process as they gain a disproportionate share of tax reductions.

To test the robustness of these results a sensitivity test was performed using the estimated marginal costs of public funds of 1.35 for raising fuel taxes. The absolute levels of the welfare changes decrease substantially, but the relative values remain about the same as shown in Fig. 10.3. Earmarking funds to the transport sector is therefore still welfare-inferior to allocating them to the general budget.

As noted above, the results of the ASTRA model differ sharply from those of the MOLINO model. After an initial decline of GDP for five years in the *Road* and *Cross* scenarios, GDP slowly recovers towards its level of the reference case (see Fig. 10.4). By contrast, in the *Direct taxes* scenario, GDP drops steadily below the reference level.

The initial decline of GDP and other macro-economic indicators, such as gross value added by sectors or employment, is due to the shift of transport demand to slower modes, which reduces the productivity of transport-intensive market segments such as the export market. The later positive

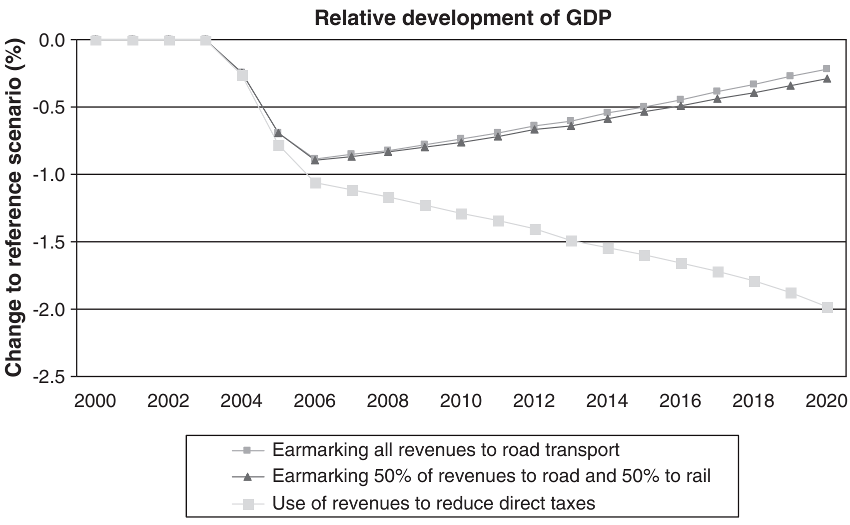

Fig. 10.4. Long-Term Development of GDP for Different Revenue Spending Scenarios (ASTRA Model Results). *Source:* Link and Stewart (2005).

development of the reinvestment scenarios can be explained by productivity gains of the economy and of the trade sector due to improved transport market conditions, and by employment effects entailed by network, facility and rolling stock investments. In contrast, lowering direct taxes in the DT scenario serves only to stimulate consumption. This and the demand-related effects of the HGV Motorway Toll are, however, not enough to offset the additional financial burden imposed on the business sector. In this respect, ASTRA agrees with the less enthusiastic view of the MOLINO model concerning the introduction of average cost-based road pricing schemes.

Similar developments can be observed for the gross value added in different economic sectors, employment, disposable income, exports, transport demand and CO_2 emissions. Only with respect to CO_2 emissions does the *Direct taxes* scenario perform well by curbing economic growth.

The difference between the MOLINO and the ASTRA results can be explained by the existence of an omniscient, benevolent and welfare-maximising state within the MOLINO model. The main difference between earmarking the revenues for the transport sector and fully transferring them to the public budget is the degree of freedom of the state to decide on how to spend the money. In the absence of earmarking rules, the state can always decide to lower labour taxes and bring about a welfare gain from the reduction in excess burden. In contrast to this, the state sector within the ASTRA model follows predefined rules and is thus less affected by changing organisational forms, responsibilities and earmarking rules. In the current model setting the state only has the power to lower direct taxes, and is thus much more limited than the government in the MOLINO model. In ASTRA, the economic effects of investments or tax reductions evolve according to their influence on consumer behaviour, productivity, technical progress and many other intermediate variables, which may change substantially over time.

As the two models incorporate different economic considerations their results complement each other and thus help to draw a wider picture of the various aspects related to the spending of transport pricing revenues. If funding were more productive in some non-transport sector than in transport, cross-sector subsidisation would be warranted. However, a simple re-distribution of revenues by reducing direct taxes should be avoided as this way of revenue use does not stimulate productivity increases in the transport and business sectors in the medium to long run.

In practical terms, the transfer of toll revenues to the public purse is restricted by EU policy. The EC's White Paper on the common transport policy in 2010 (EC, 2001a) promotes the earmarking of transport pricing revenues to the transport sector. Moreover, the call for liberalisation and private capital involvement in the rail and shipping sectors (EC, 2001b; EC, 2005) reduces the capacity of the public sector to absorb large amounts of revenue from transport infrastructure user charges. And public discourse in Germany reveals that serious equity and acceptability problems arise if revenues are not earmarked to transport.

2.2.4. The Pricing Regime

2.2.4.1. Research Question 4. Is the choice of rules for setting prices more important than the decision how to use the revenues as far as affecting welfare? And which pricing regime performs best from a welfare perspective?

To answer this very fundamental question, the current HGV charging regime was compared to two alternative regimes: one in which all road vehicles were charged for average infrastructure costs on all road types, and the other in which marginal social cost pricing was applied to all modes. In the latter case, current taxes were replaced by user charges composed of marginal infrastructure, congestion and environmental costs. In all three cases revenues were fully earmarked to transport, and within the transport sectors revenues were used as for Research question 3.

The MOLINO model results in a welfare increase of €1.0 billion against the reference case when average cost pricing is replaced by a marginal social cost pricing scheme. This is roughly forty times the gain identified for Research question 3 from allocating all revenues to the general fund. This suggests that, if acceptability and distortions in other markets are ignored, it is not the type of revenue spending that matters but the underlying pricing principle.

This conclusion is tempered, however, by the size of the price changes induced by marginal social cost pricing. Under average cost pricing, passengers are charged €1.9/100 km on roads and €2.5/100 km on rail. By contrast, marginal social cost prices in the MOLINO model runs range up to €20/100 km for road and €11/100 km for rail in peak hours. Freight transport shows a similar picture. These high tariffs are due to congestion and external costs. Transport users experience a dis-benefit of €3.8 billion, while under average cost pricing conditions and no earmarking of revenues to the transport sector user benefits increase strongly but overall social welfare decreases.

Another qualification is that the results are rather sensitive to the value for the marginal cost of public funds. If the MCPF value of 1.0 for marginal social cost pricing (MSCP) revenues is replaced by the value assumed for average cost pricing and fuel tax revenues (1.35), the social surplus of marginal social cost pricing increases by 60 per cent.

3. ACCEPTABILITY OF HGV CHARGING AND REVENUE USE SCHEMES

3.1. The Survey Design

Both the German road haulage companies and their interest association (BGL) have on several occasions expressed their support for introducing a distance-related HGV charging scheme on German motorways. One of the main reasons was that this charging scheme enables charges to be imposed on foreign hauliers who had been using the German motorways for free. Nevertheless, specific acceptability problems have occurred such as the calculation and determination of the charge level which had been perceived as too high, the problems experienced in the test phase with the on-board units, the institutional set-up with TollCollect receiving 20 per cent of tolling revenues for operating the scheme, and the use of revenues. In order to gauge the different dimensions of acceptability after the HGV charge was introduced in 2005, an Internet-based acceptability survey with German road hauliers was conducted. The major issues explored in this survey related to:

- The perception of the problems that transport causes (infrastructure damage, congestion, environmental pollution and accidents) and the problems that transport users face including deteriorating road conditions, competition from foreign truckers who do not pay for road use in Germany, and lack of interoperability of charging technologies in Europe;
- The design of charges (levels of charges and variation in charges according to congestion, vehicle characteristics, road type and road condition);
- The effect on acceptability of the charging technology, interoperability of charging technology and the institutional framework;
- The use of revenues in combination with different options for the institutional framework (who collects the charge and who is entitled to decide how revenues are used);
- The perceived capability of road user charging to solve transport-related problems;
- The adaptation strategies of hauliers; and
- The question whether there exists a trade-off between charge level and use of revenues with respect to acceptability.

The survey was a self-administered Internet survey carried out by a professional service for Internet-based surveys.[3] The target sample consisted of road hauliers, shippers, logistics providers and combined transport operators with a registered office in Germany.[4]

The questionnaire was in the form of an attitudinal questionnaire supplemented by a stated-preference (SP) exercise and by a set of questions about company characteristics of respondents. The attitudinal questionnaire included the following sets of questions: problem awareness (six questions), design of charges (eight questions), technology and value-added services (five questions), institutional framework (five questions), use of revenues (four questions) and impacts of road pricing (nine questions). Responses to each attitude scale question were collected using a 5-point semantic agreement scale (from "strongly agree" to "strongly disagree"). In order to gain further insight into the compensatory relationship between different price levels and revenue-spending options, a stated preference exercise was conducted as part of the survey. In this exercise, respondents were asked to rate on a 5-point preference scale relative to current conditions (with "1" indicating "much better than at present" and "5" indicating "much worse than at present"), various packages involving combinations of pricing and spending measures. Two charge levels were presented to the respondents: an average charge of €0.18/km and an average charge of €0.25/km.[5] The charge level of €0.18/km has already been under discussion in Germany as the next potential step in implementing HGV charging. A charge of €0.25/km reflects requests from environmental and rail lobby groups to set the charge at a level high enough to induce shifts from road to rail transport (Rothengatter & Doll, 2001). Both charge levels were linked with the following five revenue-spending options

A. Maintenance of roads;
B. Maintenance of roads and relief of bottlenecks and by-pass roads for heavily utilised urban roads;
C. Roads, rail and inland waterways;
D. Roads and combined transport; and
E. Contribution to the general government budget.

3.2. Hauliers' Responses

The survey yielded in total responses from 315 companies with responses on specific questions varying between 226 and 314. For interpreting the results it should be borne in mind that the distribution of company size, represented by the number of employees, is characterised by a bias towards medium-sized and large companies compared to the official company statistics for 2003 (BAG, 2005).[6] This bias is due to the choice of an Internet-based survey type where small companies are underrepresented. Almost all companies operated business as hauliers, more than two-thirds were shippers, half also offered logistical services, and 14 per cent performed combined transport services (multiple answers). The majority of companies considered their transports as time-sensitive.

The responses of haulage companies revealed the following insights into the different dimensions of charging acceptability.

Problem perception: The free use of German motorways by foreign vehicles (until 2005), the time losses and costs caused by congestion and the worsened condition of motorways are viewed as the major transport-related problems in Germany. A majority of hauliers stated that the government has sufficient tax revenues to keep roads in good condition (46 per cent agreed or strongly agreed[7]).

Design of charges: Although congestion problems are seen as severe, hauliers reject congestion charges, and in particular real-time pricing (with 92 per cent of respondents rejecting such a charging scheme[8]). They also reject an extension of the HGV charging scheme to the whole network. There is strong support for harmonised rules for determining and raising charges in the EU, and for charging vehicles below 12 tonnes GVW. There is no clear picture whether haulage companies would accept a charge design, which varies with road quality (safety standards, etc.) and road condition (maintenance standards), and whether passenger cars should also be subject to charges.

Charging technology: The satellite-based tolling system (GPS/GSM) is viewed as the best technological solution. However, the institutional set-up in which TollCollect receives 20 per cent of toll revenues for operating the system is considered too expensive. A majority of companies has installed or intends to install the TollCollect On-Board Unit (OBU). German hauliers seem to view the microwave-based tolling system, as applied with the Go-Box in Austria, as more user-friendly than GPS/GSM. Lack of interoperability of different charging systems in Europe is perceived as a major disadvantage. Hauliers are undecided regarding the question whether the satellite-based technology will be the long-term binding solution in the EU. This uncertainty might have been increased by the critical discussion on the EC (2003, p. 123).

Institutional framework: The highest agreement was reached for an institutional set-up in which either a motorway operator (independent of the ownership structure) collects the revenues and decides on their use, or a governmental agency collects the revenues while an independent agency (fund operator) decides on the use of revenues (Fig. 10.5).[9] All proposals in which the revenues go to the general budget and the state decides on how they are used were rejected regardless of what party collects the revenues. This reveals a general fear of the haulage business that revenues will be used to cover government deficits. The low level of support amongst hauliers[10] for having a private company collect the charges likely reflects their negative experience with ordering and installing the TollCollect OBUs, and the resulting negative image they developed of TollCollect, the private operator of the German HGV charging scheme.

Use of revenues: German road hauliers prefer that revenues from HGV charging be spent in the road sector (Fig. 10.6); in particular for a combination of road maintenance and new construction (bottlenecks, by-pass roads). This preference is consistent with the perception of deteriorating road conditions in Germany, neglected maintenance and (probably overstated) congestion problems. On the contrary, using road-charging revenues

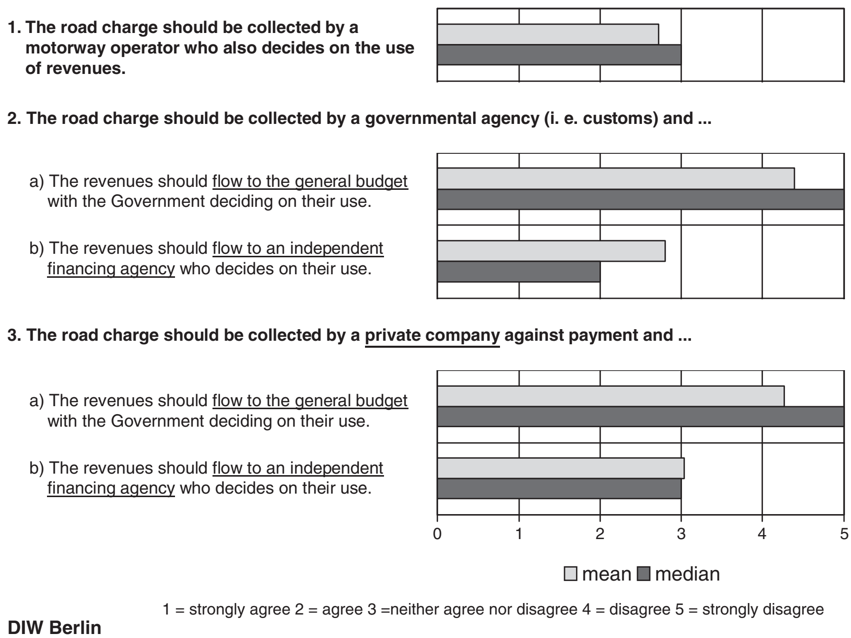

Fig. 10.5. German Haulier Survey: Institutional Framework. *Source:* Link and Stewart (2005).

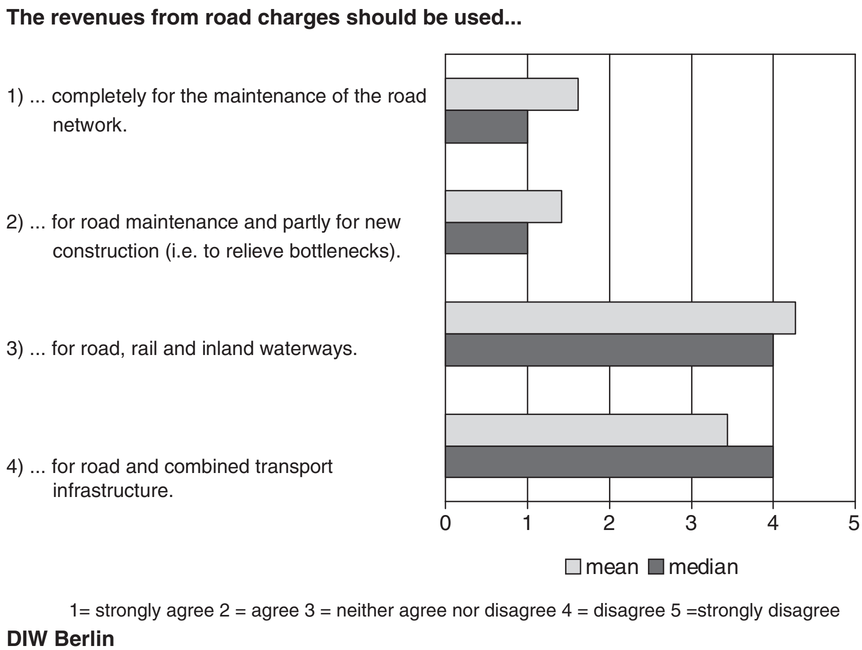

Fig. 10.6. German Haulier Survey: Use of Revenues. *Source:* Link and Stewart (2005).

for alternative modes such as rail and inland waterways is rejected. Half of the respondents also rejected the option of splitting revenues between road and combined transport. The stated preference exercise revealed a trade-off between charge level and use of revenues (Table 10.1). In general, the ranking of spending options is the same for the €0.18 and €0.25 charges, with options A and B being the most accepted ones. However, the acceptability of the €0.25 charge with a guaranteed use of revenues within the road sector is higher than for the €0.18 charge with the spending options C, D, E (intermodal use for road, rail and inland waterways; use for roads and combined transport; contribution to the general state budget). Furthermore, respondents are also more favourable to a higher charge if revenues are used for road and combined transport, than a lower charge with revenues allocated either to road/rail/inland waterways or to the general budget. This trade-off between charge level and use of revenues should be considered when discussing changes to the charging policy, as, for example, the planned increase of charges in Germany.

Table 10.1. Stated Preference Survey: Mean Values of Preference Scores*.

Options of Using Revenues	Charge Level 1: €0.18/km	Charge Level 2: €0.25/km
A. Road maintenance	2.43	2.71
B. Road maintenance, relief of bottlenecks, by-pass roads	2.44	2.66
C. Intermodal – road/rail/inland waterways	4.22	4.28
D. Road and combined transport	3.83	3.91
E. Contribution to general state budget	4.68	4.69

Source: Link and Stewart (2005).
*Lower scores mean higher level of agreement.

Outcome beliefs and adaptation strategies: Finally, German road hauliers remain sceptical that charging HGVs will reduce traffic on motorways. Using non-tolled roads as an alternative is not seen as a promising adaptation strategy. The majority of companies intend to pass cost increases caused by the HGV charge through to their clients. Acquisition of low-polluting vehicles that are charged lower tolls is preferred to acquisition of vehicles below 12 tonnes (for example, purchasing newly developed trucks with a maximum GVW of 11.9 tons) that are not currently charged. This suggests that the environmental incentives in the charging design seem to work. It also indicates that fears and press accounts that companies will avoid charges by acquiring vehicles below 12 tonnes GVW are overblown. Higher load factors are seen as a viable strategy by only one-third of respondents, perhaps because vehicle utilisation may have already approached saturation. Companies are equivocal regarding staff reductions and the risk of shutting down. In both attitudinal questions on these issues about one-third of respondents did not view such a response on the charging scheme as probable, one-third answered with "neither/nor" and the remaining third stated that they consider such options. Respondents viewed rather sceptically potential savings in shifting from own-account transportation[11] to for-hire transport.

4. CONCLUSIONS

This chapter has explored how amending the German HGV charging scheme by adopting different pricing and revenue-use rules would influence efficiency, equity and acceptability of the scheme. The charging scheme currently in operation is based on an average-cost-pricing rule with charges varying according to axle loads and exhaust emission standards. Twenty per cent of the revenues from this scheme (€2.86 billion in 2005) are granted to the private firm TollCollect for operating the scheme. Of the remaining revenues 50 per cent are granted to road, 38 per cent to railway and 12 per cent to inland waterways. The research presented in this chapter covered a set of scenarios in which different pricing rules (including marginal social cost pricing) and different options for spending revenues were analysed. Economic efficiency and equity effects were assessed using two dissimilar models: the MOLINO partial equilibrium model and the ASTRA system dynamics model. In addition, acceptability issues were addressed using the results of a survey of actors in the German haulage business.

In general both models find a decrease in overall social welfare from charging regimes similar to the one currently in operation on German motorways. But transport users (in particular high-income users) are better off with the toll system. Concerning the design of the charging regime the modelling results were as follows.

MOLINO suggests that from a social point of view revenues are better allocated to maintenance since investments in new capacities entail long-term financial burdens. But transport users benefit more from the provision of new capacity. Both the MOLINO and the ASTRA models conclude that, under the existing structure of the German transport system, investments in roads are slightly more beneficial than expenditures for rail and inland waterways. This preference is driven by the greater level of congestion on roads and could change in the future.

The preference for ploughing charge revenues back into the road sector is consistent with the results of the acceptability survey. Haulage companies expressed a clear preference for using HGV charging revenues for road maintenance and new construction (relief of bottlenecks), and a strong preference against cross-subsidisation. This is underscored by the trade-off between charge level and revenue use revealed in the stated preference exercise. This exercise indicated that road haulage companies would accept a higher charge if it were guaranteed that revenues were used in the road sector.

The results of the MOLINO and ASTRA models differ sharply on whether revenues from HGV charges should be allocated to the transport sector. According to the MOLINO model the welfare-superior solution is marginal social cost pricing with revenues allocated to the general budget rather than earmarked to the transport sector.[12] A different conclusion emerges from the system-dynamics ASTRA model. While the use of revenues for tax reduction alone does not have a stimulating effect on the macroeconomy, earmarking revenues to transport induces productivity gains in transport that spill over to other economic sectors. Furthermore, the theoretical superiority of allocating revenues to the state presumes that the state takes optimal pricing and revenue decisions, which is not the case in reality. By comparing the EU transport policy and economic performance to the US, Schade et al. (2006) demonstrate that the impact of transport policies on economic growth is frequently over-estimated. Moreover, the acceptability survey reveals that such a solution does not seem to be acceptable to freight hauliers who are affected by charging.

If, for whatever reasons, it is decided to earmark HGV revenues to the transport system, as it has been done in Germany, the MOLINO model results suggest that it is generally welfare optimal to allocate revenues to the road sector. Using HGV revenues within the road sector rather than cross-subsidising other modes is welfare-superior simply because investments in roads yield larger user cost savings.

According to the acceptability survey, haulage companies have a clear preference for using HGV charging revenues for road maintenance and new construction (relief of bottlenecks), and a strong preference against cross-subsidisation. This is underscored by the trade-off between charge level and revenue use revealed in the stated preference exercise. This exercise indicated that road haulage companies would accept a higher charge if it were guaranteed that revenues were used in the road sector. It appears that those affected by the HGV charging scheme prefer either that a motorway operator (independent of its ownership structure) collect the revenues and decide on how they are used, or that a governmental agency collect tolls while an independent agency (such as the newly founded transport infrastructure financing society, VIFG) decide on revenue allocation. Set-ups in which either the state collects and uses the revenues within the general budget, or a private company such as TollCollect collects the revenues are rejected.

Current practice in Germany with the HGV charging scheme, the institutional framework and the use of revenues, conforms only partly with these case study findings. The HGV charge level is based on average costs rather than the welfare superior marginal social cost pricing principle. Revenues do not go to the general budget as recommended by the MOLINO modelling results. Instead, the current scheme entails an intermodal distribution of revenues that reflects political, practical and acceptability considerations. The 50 per cent share of revenues dedicated to motorways might be considered as a political compromise between efficiency considerations, acceptability and practical financing needs for transport modes.

NOTES

1. Rail and inland waterways are combined because the MOLINO model features a choice between only two transport modes.
2. This value of the MCPF of 1.58 is also applied to estimate the social benefits of tax reductions.
3. See http://www.echopoll.com. Using this kind of service ensured that all legal requirements for data protection were fulfilled. Companies could complete the survey only once, and only companies that were contacted were eligible to complete the survey.
4. The company addresses were taken from the internet database "Wer liefert was" from which 1500 companies were selected that cover the full range of sizes (measured by the number of vehicles and employees, and by financial turnover). In addition, the German Road Haulier's Association (BGL) supported the survey by posting an announcement of the survey including the web address for the questionnaire on the member's-only section of the BGL homepage.
5. The current charge differentiation according to environmental criteria was left unchanged.
6. 23 per cent of responses fall into the company category with less than 10 employees (according to BAG 2005, 74 per cent), one quarter between 10 and 19 and between 20 and 50 employees respectively (BAG, 2005, 15 per cent and 8 per cent, respectively), and another quarter were companies with more than 50 employees (BAG, 2005, 3 per cent).
7. 35 per cent thought that the tax revenues are not sufficient, 18 per cent responded with neither/nor, 46 per cent viewed the state's income as sufficient.
8. Congestion charging was explained as setting higher charges at peak times and lower charges in off-peak periods.
9. Forty-eight percent agreed or strongly agreed with the motorway operator solution while 27 per cent disagreed or strongly disagreed. Fifty percent agreed or strongly agreed with the state being responsible for collecting the charges and an independent agency deciding on the use of revenues, while 32 per cent disagreed or strongly disagreed.
10. Eighty-three percent rejected private toll collection if the state decides on the use of revenues, and 40 per cent reject it even if an independent agency decides on the use of revenues.
11. Transportation that is performed by goods producers with company-owned vehicles. In-house transportation achieves lower load factors than does for-hire transport, and could be put at a competitive disadvantage by the HGV charge.
12. This result is rather sensitive to the value chosen for the marginal cost of public funds. In particular, the advantage of allocating revenues to the general budget is considerably smaller with a MCPF of 1.35 than with the base-case value for Germany of 1.58.

REFERENCES

BAG (2005). Struktur der Unternehmen des gewerblichen Güterkraftverkehrs und des Werkverkehrs. Volume USTAT 9, Cologne.

Balmer, U. (2003). Practice and experience with implementing transport pricing reform in heavy goods transport in Switzerland. IMPRINT seminar, Leuven.

BMU (2000). *Nationales Klimaschutzprogramm 2000. Series "Umwelt" no. 11/2000*. Federal Ministry for Environment, Nature Conversation and Nuclear Safety (BMUI), Berlin.

BMVBW (2003). *Bundesverkehrswegeplan 2003 – Grundlagen für die Mobilität in Deutschland*. Project Group on the Federal Investment Plan at the Federal Ministry for Transport, Building and Housing (BMVBW), Berlin, July.

BMVBW (2005). *Verkehr in Zahlen 2004/05*. Federal Ministry for Transport, Building and Housing, Deutscher Verkehrs-Verlag, Hamburg.

Bundesregierung (2001). *LKW-Maut auf Bundesautobahnen beschlossen*. Federal Ministry of Transportation, Construction and Housing. Press release number 208/01, Berlin, August 15.

Bundesregierung (2004). *Autobahnmautgesetz für schwere Nutzfahrzeuge (ABMG) of 5.4.2002*. BGBl.III/FNA Nr.9290-13, revised edition, December 2.

Bundestag (2006). Bericht der Bundesregierung über die Verlagerungen von schwerem Lkw-Verkehr auf das nachgeordnete Straßennetz infolge der Einführung der Lkw-Maut. Printed matter of the German Parliament no. 16/298.

DB (2004). *Geschäftsbericht 2003 der DB Netz AG*. Frankfurt: Deutsche Bahn AG

EC (2001a). *European transport policy for 2010: Time to decide*. White Paper of the European Commission COM(2001) 370 final. Brussels, September.

EC (2001b). Directive 2001/14 of the European parliament and of the council of 26 February 2001 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification. *Official Journal of the European Communities, L 75/29*(March).

EC (2003). Proposal for a Directive of the European Parliament and of the Council amending Directive 1999/62/EC on the charging of heavy goods vehicles for the use of certain infrastructure. Draft proposal from 23.07.2003.

EC (2005). *Reinforcing quality services in sea ports: A key for European transport*. Communication from the European Commission to the European Parliament and the Council. Brussels, February.

Follmer, R., Kunert, U., Kloas, J., & Kuhfeld, H. (2004). *Mobilität in Deutschland*. Results report to the Federal Ministry for Transport, Building and Housing (BMVBW). Projekt no. 70.0736/2003. Infas GmbH, Bonn and DIW, Berlin, April.

Gresser, K., Hölsken, D., & Kienzler, H.-P. (2001). *Verlagerungswirkungen bei Einführung der Lkw-Maut*. Final report to the Federal Ministry for Transport, Building and Housing (BMVBW). Project no.: 96.695/2001. Working group BVU / IVV / K+P. Berlin, Freiburg, Aachen.

KBA (2004). Kraftfahrt-Bundesamt (German Federal Bureau of Motor Vehicles and Drivers) Grenzüberschreitender Güterkraftverkehr im Jahr 2002 nach Heimatländern der Lastkraftfahrzeug Reihe 9, Sgb. 332, 4. VJ 2002 Flensburg, June 6.

Kleist, L., & Doll, C. (2005). Economic and environmental impacts of road tolls for HGVs in Europe. In: J. M. Viegas (Ed.), *Interurban road charging for trucks in Europe*. Research in Transportation Economics (Vol. 11). Amsterdam: Elsevier Ltd.

Kleven, H. J., & Kreiner, C. T. (2003). *The marginal costs of public funds in OECD countries. Hours of work versus labour force participation*. CESIFO Working Paper no. 935 category 1: public finance, April (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=404582).

Link, H., & Stewart, L. H. (2005). *Inter-urban case studies. REVENUE project deliverable 4*. Funded by the 5th framework RTD programme, ISIS, Rome, July 2000. Draft version.

NEA, COWI, IWW, NESTEAR, PWC, TINA, IVT, Herry, & MKmetric (2003). Scenarios, traffic forecasts, and analysis of corridors on the trans-European networks (TEN-STAC), Deliverable D6: Traffic, bottlenecks, and environmental analysis on 25 corridors. Study funded by the European Commission.

Pällmann, W. (2000). *Kommission Verkehrsinfrastrukturfinanzierung*. Final report of the governmental commission on transport infrastructure financing to the federal government. Berlin, September.

Rommerskirchen, S., Rothengatter, W., Helms, M., Doll, C., Liedtke, G., & Vödisch, M. (2002). *Wegekostenrechnung für das Bundesfernstraßennetz unter Berücksichtigung der Vorbereitung einer streckenbezogenen Autobahnbenutzungsgebühr für schwere Lkw*. Final report to the Federal Ministry for Transport, Building and Housing (BMVBW), Berlin. Basle, Karlsruhe.

Rothengatter, W., & Doll, K. (2001). *Anforderungen an eine umweltorientierte Schwerverkehrsabgabe für den Straßengüterverkehr*. Study 20096130 UBA-FB 00217, commissioned by the German Federal Environmental Agency. Final Report. Texte des Umweltbundesamtes, No. 57/01.

Rothengatter, W., & Doll, C. (2002). Design of a user charge for heavy duty vehicles on German motorways considering the objective of efficiency, fairness and environmental protection – findings from the EU research project DESIRE. *IATSS Research*, *26*(1), 6–16 Tokyo.

Schade, W. (2005). *Strategic sustainability analysis: Concept and application for the assessment of European transport policy*. Unpublished doctoral dissertation, Faculty for Economics, University of Karlsruhe. NOMOS-Verlag, Baden-Baden, April.

Schade, W., Doll, C., Maibach, M., Peter, M., Lilico, A., Conti, M., Crespo, F., & Carvalho, D. (2006). Analysis of the contribution of transport policies to the competitiveness of the EU economy and comparison with the United States (COMPETE). Study submitted to the European Commission co-ordinated by Fraunhofer-ISI, Karlsruhe.

This page intentionally left blank

CHAPTER 11

FRENCH MULTI-MODAL TRANSPORT FUNDS: ISSUES OF CROSS-FINANCING AND PRICING

Charles Raux, Aurélie Mercier and Stéphanie Souche

ABSTRACT

This paper assesses the efficiency of cross-financing new motorway or rail projects from established toll motorways as an alternative to direct public subsidies. For new motorway projects a combination of short-run marginal social cost pricing and cross-financing is the best of the alternative schemes tested for increasing overall welfare. Regarding the Lyon–Turin rail project, an alpine fund supplied by toll mark-ups on the Alpine motorways in combination with the same national transport fund as in the road case study, would eliminate the need for public subsidies and simultaneously improve the financial balances of the rail operator and manager.

1. INTRODUCTION

In December 2003, the French government decided to implement an ambitious transport programme with 35 major road and rail infrastructure projects, representing an overall investment of €20 billion between 2005 and 2012. In order to finance part of this programme a new funding agency with

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 243–268

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19011-9

an independent budget, known as AFITF ("Agence pour le Financement des Infrastructures de Transport de France") was created on 1 January, 2005. Twenty billion euros is required for the programme. AFITF was expected to provide €7.5 billion of this with a borrowing capacity on the capital market based on revenues from paid-off state-owned toll motorway companies. However, this revenue source may be undermined by the ongoing privatisation of motorway companies. The remaining €12.5 billion for the programme was expected to be met by external partners (local governments, European Union, private sector with public–private partnerships).

AFITF is the fourth transport investment fund to be created in France since 1995. The first, the FITTVN ("Fonds d'Investissement pour les Transports Terrestres et les Voies Navigables"), was created in 1995 to relaunch public investment in transport infrastructure but was abolished in 2001. Two other funds were created in 2002 to develop intermodal transport in France and the Alpine region. However, these two funds never became operational.

When it was created, the FITTVN symbolised a new policy trend aimed at promoting regional development and intermodal transport. Funded from specific taxes, it was intended to be independent from the general budget and to circumvent the budget constraint. The fund was financed by the "taxe d'aménagement du territoire" (a disguised additional motorway toll; see Box 11.1 for the history of motorway tolling in France), which was expected to provide a permanent source of finance. The fund has financed several major transport links, including motorways, the Mediterranean TGV line and some stretches of inland waterways.

At the outset, the fund appeared to have some advantages. It was launched by a political leader with the support of elected representatives (the Senate), and was based on a "cash cow" revenue source (motorway users who have no real capacity to organise themselves as a structured interest group). However, the FITTVN had two weaknesses. First, the regional development and intermodality objectives were not clearly defined. Second, as a consequence of this lack of clarity and the existence of an overall public budget constraint, a kind of crowding out occurred by which the fund replaced funding from the central government's general budget instead of providing new financing capacity.

These weaknesses were exploited by several interest groups, and the fund was abolished in 2001 while the motorway tax was maintained. Several conclusions can be drawn from this experience. First, the objectives and scope of this kind of fund must be clearly defined, and a strong legislative and institutional basis is required in order to resist pressures – be they

Box 11.1. The Current Tolling Regime on French Motorways.

The French Revolution of 1789 abolished tolls on paths, bridges and city entrances ("péage d'octroi") and introduced the principle of free usage of all roads in the name of equality between citizens. However, the need to fund transport infrastructure led to a law being passed in 1955 which authorised the creation of toll motorways. The operator is allowed to collect tolls from motorway users in order to pay off capital costs and maintain, or even extend, the motorway. However, a free alternative – such as toll-free parallel road – must be available.

On interurban tolled roads average (flat) distance-based pricing is applied with a distinction between freight and passenger vehicle classes. Tolls may vary from one motorway to another depending on the financial needs of each concession. These tolls are fixed in the concession contracts. Toll revenues accrue to the motorway concessionaire who invests, operates and maintains the infrastructure. Depending on the duration of their concessions and on initial public subsidies, this has enabled concessionaires to cover most of their costs at the network level.

In addition to tolls, motorway operators collect from users a regional development tax ("taxe d'aménagement du territoire") that was €0.007/km in 2000. The revenue from this tax is not earmarked and goes into the central government budget.

Until recently, concessions for new motorway sections were awarded to an existing motorway concessionaire in the same geographical area. The concessionaire used a kind of cross-financing from paid-off sections of its own network to new sections, most of which would not be profitable on their own. As a consequence of the European Council Directive 93/37/EEC concerning the procedures for awarding public works contracts, the government is now required to advertise the contracts for motorways in order to open up these contracts to effective community-wide competition. This means that competitors must be treated equally, so that if subsidies are required they have to be the same for each competitor.

political or financial. Second, sound initial design and functional operation are required in order to control the interest groups and avoid the formation of excessively powerful opposition coalitions. To do this effectively, efficiency and equity as well as legislative aspects must be addressed. Third, to satisfy EU guidelines pricing rules must be defined as well as rules that determine how much of user revenues can be allocated to finance particular modes.

The repeated attempts to establish investment funds in France underscore the need for a permanent (or ''sustainable'') means of financing transport investments for regional development. This chapter will address a number of issues related to cross-financing as an alternative to direct public subsidies, and to the use of user charges as a revenue source and demand-management tool. It will assess the efficiency of cross-financing new motorway or rail projects by established toll motorways via a multi-modal transport fund. The pricing reform aims not only to raise revenue but also to improve the efficiency in the use of infrastructure. The merits of these investment programmes are not questioned; thus the use of money for purposes other than transport infrastructure is not considered.

The evaluation is based on two case studies. The first, described in Section 2, deals with the financing and pricing of a programme of 10 new motorway projects. The second case study (Section 3) concerns the cross-financing of the Lyon–Turin rail link from Alpine motorways. Some general conclusions from the results of these two case studies are drawn in Section 4.

2. THE FINANCING AND PRICING OF NEW MOTORWAY PROJECTS

Ten new motorway projects were selected from the governmental infrastructure programme for our case study. The selection is based on the stage of planning and on the data availability. Table 11.1 provides some general information about the projects including construction costs and the expected level of public subsidies.[1]

The first stage of the analysis was to compare alternative financing schemes involving public subsidies with partial financing by a transport fund. Since public subsidies have an excess burden (the marginal cost of public funds, MCPF), the sensitivity of the welfare computation to the level of the MCPF is briefly evaluated. The next stage was to consider marginal social cost pricing (MSCP) as an alternative to the current pricing regimes. Although MSCP is not optimal in the presence of financing constraints, it is nevertheless a useful benchmark. Finally, the combination of MSCP and cross-financing from a transport fund is evaluated.

Table 11.1. Overview of the Motorway Projects.

Tolled Motorway Project	Length (km)	Alternative Road	Construction Cost[a] (€ million)	Total Public Subsidies[a] (€ million)	Public Subsidies/ Construction Cost (%)	Construction Period
A24 Motorway "Amiens-Lille-Belgium"	120	A1 Tolled motorway	666	375	56	2009–2012
A48 Motorway "Amberieu-Bourgoin Jallieu"	55	N75 Highway	605	365	60	2010–2015
A51 Motorway "Grenoble-Sisteron"	80	N75 Highway	1,200	670	56	2011–2018
A585 Motorway "Les Mees-Digne-les-Bains"	25	Highway 85	209	169	80	2008–2010
A831 Motorway "Fontenay le Comte-Rochefort	64	Highways 11 and 135	468	243	52	2009–2012
A89 Motorway "Lyon-Balbigny"	50	Highways 7, 82 and 89	769	625	81	2006–2011
A19 Motorway "Artenay-Courtenay"	100	N60 Highway	607	165	27	2006–2009
A41 Motorway "Saint-Julien -Villy"	18	N201 Highway	674	277	41	2007–2010
A45 Motorway "Lyon – Saint-Etienne	52	A47 Toll-free motorway	1,300	1,118	86	2008–2011
A65 Motorway "Pau-Langon"	142	N10 and N134 Highways	910	142	15	2008–2011

[a]Excluding VAT.

In this study, the selection of new infrastructure projects is taken as granted since the political decision has already been taken. We only assess alternative ways of financing the given selection of new projects. The financing alternatives will be combinations of public revenues (at a MCPF cost premium) and revenues from transport pricing.

2.1. Methodology

The MOLINO model was used in a "two roads, passenger/freight" configuration. For each of the 10 infrastructure projects, we model a hierarchical decision tree with three levels: the uppermost choice is between transport and other consumption,[2] the second choice is between travelling in a peak or an off-peak period, and the third choice is between the new motorway projects and the alternative highway. As Table 11.1 indicates the alternative roads are not tolled except for the A1 motorway, which is an alternative to the A24.

MOLINO was used in a *static* approach: the context is as if the new projects are already implemented and in operation, and the consequences of variants of financing and pricing the use of infrastructure on welfare are assessed for a representative year in the future. This means that, as implied by this static approach, we take as given the level of traffic initially forecast for each project.[3]

Given the location of the motorway projects, it is assumed that network effects are negligible; i.e. opening any new motorway has no impact on demand for the other new projects. Each project is thus assessed in this MOLINO configuration independently of the others. Moreover, since there is no interaction between projects, the overall welfare result is simply the sum of the individual results.

In the assessment, public subsidies and funds are exogenously allocated to each project. Projects receive an equal allocation from the transport fund, and potential competition between projects for financing is ignored.

Regarding the welfare cost of money, a value of 1.1 is used for the MCPF.[4] All the other parameters used in the modelling exercise are detailed in Appendix A.1.

2.2. Results

The results of the motorway assessment are summarised in four sections. Sections 2.2.1 and 2.2.2 compare the performance of different means of financing and test the sensitivity of the rankings to the value used for the MCPF. Section 2.2.3 assesses the planned pricing and MSCP rules, and the final section deals with MSCP in combination with cross-financing.

2.2.1. Financing

The motorway projects are assumed to be cross-financed through an investment fund that is independent from the general budget and financed by taxes on existing toll motorways. The fund partially finances each project by means of a grant. To determine the size of the grant it is assumed that the fund's financing capacity remains stable over the period 2005–2012, and that the grant for each project is proportional to the fund's annual financing capacity. The proportion is computed on the basis of the ratio between the total subsidy needed by the project and the total expenditures of the fund during the period 2005–2012.[5]

First, the alternative "planned toll + transport fund + public subsidies" scheme is compared with the reference scheme "planned toll + public subsidies''; see Table 11.2. In the reference scheme, the new motorway projects are financed from the forecast toll revenues for these projects supplemented by public subsidies obtained from fiscal revenues. In the alternative scheme, the new motorway projects are financed from the forecast toll revenues from these projects, subsidies from a transport fund (supplied by the tax on existing motorway tolls[6]) and additional public subsidies only when needed.

The difference between these two schemes essentially derives from the welfare cost of money. For public subsidies that are paid out of fiscal revenues, this cost is the MCPF, which is discussed in more detail in Section 2.2.2. For revenues derived from the tax on existing motorway tolls the relevant cost is the welfare loss incurred by traffic that is tolled off the existing motorways. The latter cost can be estimated using information on current demand on tolled motorways, the average toll and the empirical elasticity of motorway demand to tolls which is approximately −0.5.[7] The toll revenues available for cross-financing roughly correspond to the revenues from the current regional development tax on tolls (''taxe d'aménagement du territoire'') of 0.007€/km. This tax has a deadweight loss of €0.023 per euro collected.[8] As shown below this has a marginal impact on overall welfare.

The figures in Tables 11.3 and 11.4 show that the introduction of the transport fund yields a welfare improvement (ratio of annual welfare gain on construction cost[9] of 0.012 per cent). There is an increase in central

Table 11.2. Overview of Financing Schemes.

Alternative Schemes	Pricing Regime	Revenue Use
Planned tolling regime + public subsidies	Tolling as originally planned for the new projects	Revenues from road tolls on each new motorway go to the motorway concessionaire Additional public subsidies to cover new motorway construction costs
Planned tolling regime + transport fund + public subsidies		Revenues from road tolls on each new motorway go to the motorway concessionaire Tax on existing toll motorways goes to the transport fund. The transport fund subsidises new motorways Additional public subsidies when needed

Table 11.3. Assessment of Financing Schemes.

Change in	Planned Tolls + Transport Fund + Public Subsidies (%)
Annual welfare gain[a]/construction cost	0.012
Annual welfare gain[a,b]/construction cost	0.012
Central government net revenue	17
Public subsidies	−95.0
Fuel tax revenues	20.6

[a]Compared to the "planned toll + public subsidies" scheme.
[b]Including welfare loss from payment of the regional development tax by toll motorway users.

Table 11.4. Performance of "Planned Tolling Regime + Transport Fund + Public Subsidies" Scheme.

	Change in Central Government Net Revenues[a] (%)	Change in Need for Public Subsidies[a] (%)	Annual Welfare Gain[a]/Construction Cost (%)
A24	4	−95	0.010
A48	12	−93	0.006
A51	12	−92	0.024
A585	154	−100	0.014
A831	20	−92	0.008
A89	53	−96	0.013
A19	17	−93	0.004
A41	109	−96	0.007
A45	28	−96	0.014
A65	40	−96	0.010
Total	17	−95	0.012
Total[b]	17	−95	0.012

[a]Compared to the "planned toll + public subsidies" scheme.
[b]Including welfare loss from payment of the regional development tax by toll motorway users.

government net revenues (+17 per cent) that derives from the 95 per cent reduction in public subsidies. The transport fund is more advantageous than a public subsidy mainly because the deadweight cost of the tax from reduced traffic is much smaller than the MCPF.

2.2.2. Sensitivity to the MCPFs

The value of 1.1 used for MCPF is a conservative value. It is similar to the value of 1.13 quoted in Lebègue, Hirtzman, and Baumstark (2005). Other studies quoted in the same reference give values of 1.5. To test the sensitivity of the welfare results to the MCPF a value of 1.5 was also considered. This change results in an increase of 36 per cent in the welfare gain from the "planned tolls + transport fund + public subsidies" scheme vis à vis the "planned tolls + public subsidies" scheme.[10] The welfare result therefore depends directly on the MCPF value. This sensitivity test confirms the welfare benefit of financing these projects with a fund financed by external resources – such as contributions from motorway users – which in our study yields a much lower levy cost (1.023 as noted above) than public subsidies from taxation. The benefits of the financing fund increase monotonically with the MCPF.

2.2.3. Marginal Social Cost Pricing

MSCP is not optimal in the presence of financing constraints, but it nevertheless provides a useful benchmark against which to measure other pricing regimes. MSCP is applied to the new motorways and the alternative highways that are currently free. Consistent with pure MSCP, fuel taxes are abolished in this scheme.[11] We do not discuss the reform of the existing tolled motorways since, by assumption, they do not affect the 10 projects that were selected.[12]

The "pure MSCP" scheme is compared with the "planned tolling" scheme (see Table 11.5). For this exercise the costs of constructing the new motorways are ignored in order to focus on pricing effects.[13] Table 11.6 records the results of the "pure MSCP" scheme for each of the 10 planned motorway projects in comparison with the "planned tolling" scheme results.

For the 10 projects in aggregate, "pure MSCP" yields an increase in overall welfare (ratio of annual welfare gain on construction cost of + 0.002

Table 11.5. Overview of Pricing Schemes.

Alternative Schemes	Pricing Regime	Revenue Use
Planned tolling regime	Tolling as originally planned for the new projects	Revenues from road tolls on each new motorway go to the motorway concessionaire No public subsidies
Pure MSCP	MSCP on new motorway projects and their free highway alternatives	No transport fund. No public subsidies

per cent) but total revenues drop by over half (by 56 per cent). The welfare impact of replacing fuel taxes and tolls by MSCP revenues is therefore quite small. To be sure there are some disparities between individual projects. The welfare increase is largest by far (+0.009 per cent) for the A51 motorway (see below for a discussion on this case). For the other projects the welfare change is much smaller, and in the case of the A24 and A45 projects it is negative.

The low overall welfare increase is a net result of three changes that come from MSCP: changes in user charges for motorway users, changes in user charges for highway users, and a change in net government revenues. Each of these changes will be considered in turn.

(a) Changes in user charges for motorway users. The peak-period tolls on some of the motorways are much higher than the planned tolls (see Table 11.7).[14] But because the peak periods are relatively short in duration, the reduction in traffic during the peak is not very big in absolute terms. And the off-peak motorway tolls are mostly lower than the planned tolls. Since off-peak periods are much longer than the peak periods, the net result is that daily traffic for both freight and passengers increases with MSCP (see Table 11.8).
(b) Changes in user charges for highway users. Tolls are introduced on the highways for the first time (except for A1), but fuel taxes are abolished. The net effect is that total user charges increase on all highways except for A1, which is already tolled. The reductions vary widely across projects. This is illustrated by considering three examples:
   - On the A41 motorway project, the decrease in alternative highway demand for freight and passengers (see Table 11.8) can be explained by the introduction of a toll on the highway for freight and for passenger vehicles (see Table 11.7). However, on the motorway, the lower toll level for freight and passenger vehicles in off-peak periods, when compared with the planned flat toll, and small changes in tolls in peak periods generate an increase in motorway demand.
   - The A51 project exhibits the greatest decrease in highway freight demand and the greatest increase in motorway freight demand. The main reason is that the toll level decreases, particularly for freight (division by 10). On the other hand, given the toll implemented on the alternative highway, freight road users are therefore encouraged to use the motorway.
   - The A24 project is unique in facing a tolled alternative i.e. the alternative tolled existing A1 motorway.[15] The A1 motorway already carries

Table 11.6. Performance of the Pure MSCP Scheme.

Projects	Planned Tolling Scheme			Pure MSCP Scheme			
	Fuel tax revenues[a] (€ per day)	Toll revenues (€ per day)	Fuel taxes + toll revenues (€ per day)	MSCP revenues[b] (€ per day)	Change in revenues (%)[c]	Change in central government net revenue[d] (%)	Annual welfare gain[d]/construction cost (%)
A24	1,020,599	703,379[e]	1,723,978	630,207	−63	−65	−0.007
A48	290,383	132,889	423,272	242,486	−43	−39	0.001
A51	457,127	287,517	744,644	223,553	−70	−69	0.009
A585	55,913	21,818	77,730	42,385	−45	−65	0.003
A831	206,443	90,486	296,929	116,228	−61	−64	0.004
A89	263,721	117,982	381,703	201,992	−47	−47	0.001
A19	148,540	86,090	234,629	82,395	−65	−62	0.003
A41	85,085	68,054	153,139	92,773	−39	−49	0.001
A45	851,721	128,153	979,874	609,280	−38	−42	−0.001
A65	276,857	96,152	373,009	125,756	−66	−66	0.003
Total	3,656,389	1,732,519	5,388,908	2,367,055	−56	−58	0.002

[a]Aggregate fuel taxes for new motorways and their highway alternatives.
[b]Including MSCP on new motorways and their highway alternatives.
[c]MSCP revenues (in the "pure MSCP" scheme) – (fuel taxes + toll) (in the "planned tolling" scheme).
[d]Compared to the "planned tolling" scheme.
[e]Including the A1 tolled alternative.

Table 11.7. Toll on Highways and Motorways in the MSCP Scheme Compared with the "Planned Tolling" Scheme.

Projects	Toll[a] on Motorway						Toll[a] on Highway[b]			
	Planned tolling scheme		MSCP scheme				MSCP scheme			
	Freight	Passengers	Freight		Passengers		Freight		Passengers	
			Peak	Off-peak	Peak	Off-peak	Peak	Off-peak	Peak	Off-peak
A24	1.77	4.17	15.92	1.65	34.43	2.16	2.75	1.80	4.38	2.24
A48	1.06	1.97	0.95	1.14	1.58	2.01	3.69	0.93	7.75	1.50
A51	10.95	4.04	1.09	1.14	1.49	1.61	6.84	1.66	13.54	1.82
A585	0.59	1.41	0.45	0.39	0.90	0.77	1.10	0.37	2.40	0.74
A831	1.13	2.66	1.98	0.82	3.75	1.13	3.14	0.94	6.15	1.17
A89	0.87	1.84	0.94	0.97	1.62	1.70	4.02	1.12	8.24	1.67
A19	3.10	4.39	2.95	1.22	5.39	1.47	4.84	1.32	9.43	1.46
A41	0.62	1.24	0.52	0.41	1.17	0.91	1.25	0.39	2.75	0.80
A45	0.85	1.78	0.68	1.11	1.02	1.99	2.79	1.09	5.60	1.75
A65	2.89	5.01	6.22	1.66	12.15	1.83	5.49	1.68	10.35	1.73

[a]Euros per passenger or tonne per trip.
[b]All highways are toll free in the "planned tolling" scheme except for the A1 alternative to the A24.

Table 11.8. Changes in the Traffic from the "Planned Tolling" Scheme to the "Pure MSCP" Scheme.

Projects	Traffic[a] on Highways		Traffic[a] on Motorways		Total Traffic[a]	
	Freight (%)	Passengers (%)	Freight (%)	Passengers (%)	Freight (%)	Passengers (%)
A24	9	14	12	20	9	16
A48	−17	−22	20	33	3	9
A51	−73	−30	285	85	72	31
A585	−19	−40	56	117	5	11
A831	−21	−35	31	83	13	23
A89	−13	−23	18	36	7	16
A19	−31	−40	66	83	19	23
A41	−30	−44	39	54	12	16
A45	−16	−2	10	38	3	7
A65	−19	−27	48	73	11	18
Total	−9	−12	39	53	12	14

[a]Change in passengers or tonnes per day from the "planned tolling" scheme.

a high level of freight traffic. In this case, MSCP increases peak tolls on both alternatives but decreases off-peak tolls for passengers. For freight, the off-peak toll decreases on the A24 and increases on the A1.

The reductions in traffic on the highways are outweighed by the increases in traffic on the motorways. Thus, total traffic increases on all projects (see Table 11.8) and users benefit.

(c) Change in net government revenues. As shown in Table 11.6 there is a sharp decrease in central government net revenues (between 39 per cent and 69 per cent), when one compares the MSCP scheme to the reference "planned tolling" scheme, as a net result of three changes. The first change comes from the abolition of fuel tax in the MSCP scheme: this tax revenue represents more than twice the toll revenues in the "planned tolling" scheme. The second change comes from differences in taxes on profits between the two schemes: in the MSCP scheme the whole highway and motorway operator profits are fed into government revenues (because they are considered as public) whereas in the planned tolling scheme the motorway operators are private and thus pay a 35 per cent tax on profits. The third change comes from a difference in toll revenues: because of the increase in peak-period tolls, the implementation of tolls on previously free highways and the increase in overall demand, these revenues are higher with the MSCP scheme. However, the toll and fiscal revenues from operator profits are not sufficient to compensate for the abolition of fuel tax revenues.

To summarise the findings: the MSCP scheme results in a net reduction in user charges compared to the "planned tolling" scheme. Total freight traffic increases by 12 per cent, and total passenger traffic by 14 per cent (see Table 11.8). But combined user charge revenues (toll + fuel taxes) decline by 56 per cent and the central government revenues decline by 58 per cent.

Moreover, there is a problem as regards the viability of such pricing because MSCP does not produce sufficient revenues to finance new projects. These revenues have to be made up by other financing, as we will see in the next section.

2.2.4. A Combination of MSCP and Cross-Financing

As reported in Section 2, cross-financing the new motorway projects with a transport fund and MSCP are both welfare-improving. This naturally leads to the question whether a combination of the two measures will perform better than either measure independently. To ascertain this we compare a "pure MSCP + transport fund + public subsidies" scheme with the "planned tolling + public subsidies" scheme that was used as a benchmark in Table 11.2 (see Table 11.9). The main difference between the two schemes is the road pricing rule and the absence of fuel taxes in MSCP.

As Table 11.10 shows the combined policy generates a higher aggregate welfare than the cross-financing scheme only (ratio of annual welfare gain on construction cost of +0.027 per cent).

Table 11.9. Overview of Pricing and Cross-Financing Schemes.

Alternative Schemes	Pricing Regime	Revenue Use
Planned tolling regime + public subsidies	Tolling as originally planned for the new projects	Revenues from road tolls on each new motorway go to the motorway concessionaire Public subsidies when needed
Pure MSCP + transport fund + public subsidies	MSCP on the new motorway projects and their free highway alternatives	Tax on tolls from existing motorways go to the transport fund. Transport fund subsidises the new motorways Public subsidies when needed

Table 11.10. Changes from "Planned Toll + Public Subsidies" Scheme to "pure MSCP + Transport Fund + Public Subsidies" Scheme.

Projects	Central Tax Revenues[a] (%)	Change in Central Government Net Revenue[a] (%)	Annual Welfare Gain[a]/ Construction Cost (%)
A24	−44	−27	0.092
A48	−70	−3	0.021
A51	−141	−100	0.019
A585	−61	55	0.023
A831	−99	−55	0.014
A89	−66	14	0.023
A19	−118	−70	0.004
A41	−103	6	0.008
A45	−49	34	0.044
A65	−94	−46	0.015
Total	−141	−26	0.027
Total[b]	−141	−26	0.027

[a]Compared to the "planned toll + public subsidies" scheme.
[b]Including welfare loss from payment of the regional development tax by toll motorway users.

This overall result sums up the situation for most of the projects studied. Indeed, the combination of MSCP and a transport fund improves the welfare result for each of the projects, except for Highway A51. To conclude, the combination of pure MSCP with cross-financing by a transport fund and public subsidies when needed yields a greater overall welfare than the other schemes that were assessed. The combined scheme performs well for two reasons. First, cross-financing from existing motorway users has a lower social cost of funding than do subsidies from general revenues. And second, MSCP brings overall user charges into closer alignment with the social marginal costs of travel.

3. CROSS-FINANCING FROM ROAD TO RAIL: THE LYON–TURIN RAIL LINK

The Lyon–Turin rail link is used as an example for the assessment of rail projects for two reasons. One is that it is part of the Trans-European Network and symbolises European transport policy, which is based on intermodality and the development of alternatives to road transport. The second, pragmatic, reason is that competing parallel road options can be clearly identified: the Mont Blanc and Fréjus Alpine motorway crossings. This facilitates a comparison of the rail and road alternatives and makes it easier to use the MOLINO model.

The Lyon–Turin project is expected to reduce road traffic and to achieve a better balance of traffic between road and rail. The new link is composed of eight segments. The total cost amounts to €9.759 billion (present value 2005) of which €8.631 billion is expected to be publicly financed.

As was done for the motorway projects in Section 2 the investment in the Lyon–Turin link is taken as given. The analysis is limited to an assessment of alternative ways of financing and pricing the use of infrastructure.

3.1. Methodology

The assessment focuses on the cross-financing of new rail infrastructures by road revenues. Five alternative pricing and financing schemes as summarised in Table 11.11 are entertained. Financing is assumed to be channelled via a transport fund that is more specialised than the fund considered in Section 2. It is a kind of ''Alpine fund'' that can draw revenues only from the Mont Blanc and Fréjus crossings that compete with the Lyon–Turin rail link.

The MOLINO is again used in a static approach in which reference traffic volumes are given for a representative year by the initial levels forecast by the ''Lyon–Turin Ferroviaire'' company.

In the first scheme the ''current tolling regime'' is applied (the current tolling scheme on existing Alpine motorways) while public subsidies contribute to the financing of the Lyon–Turin rail link. The second and third schemes include an investment fund (Alpine fund) with existing Alpine motorway companies providing cross-financing to the Lyon–Turin rail project.[16] In Scheme 2 motorway tolls are increased by 25 per cent.[17] It is assumed that the motorway operator transfers the resulting increase in gross operating surplus (GOS) to the fund. Scheme 3 entertains a larger increase in tolls of 80 per cent in order to lower further the level of public subsidies needed for the financing of the rail link.[18] In Scheme 4 the Alpine fund is replaced by the transport fund considered in the motorway study. Thus, revenue from the regional development tax on toll motorways throughout the national network goes to the national transport fund to subsidise the rail link. It is assumed that the fund is able to pay for the entire subsidy required.[19] Scheme 5 is the same as Scheme 4 except that as in Scheme 2 a toll mark-up of 25 per cent is levied on Alpine motorways. The extra

Table 11.11. Overview of Alternative Pricing and Financing Schemes for the Lyon–Turin Rail Link.

	Alternative Pricing and Funding Schemes	Pricing Regime	Revenue Use
1 (Ref.)	Current motorway tolling scheme + public subsidies	Current tolling regime on Alpine motorways	Rail revenues go to rail operator or manager Toll revenues go to motorway operator Public subsidies for the rail link No cross-financing
2	Increase in Alpine motorway tolls by 25% + transport fund + public subsidies	Mark-ups (25%) on current tolling regime on Alpine motorways	Rail revenues go to rail operator or manager "Base" toll revenues go to motorway operator Additional surplus of alpine motorways goes to the Alpine fund which subsidises the rail link Public subsidies when needed
3	Increase in Alpine motorway tolls by 80% + transport fund + public subsidies	Mark-ups (80%) on current tolling regime on Alpine motorways	As above
4	Current motorway tolling scheme + transport fund	Current tolling regime on Alpine motorways	Tax on existing toll motorway (whole national network) goes to the transport fund. The transport fund subsidises the rail link Rail revenues go to rail operator or manager
5	Increase in Alpine motorway tolls by 25% + transport fund	Mark-ups (25%) on current tolling regime on Alpine motorways	Tax on existing toll motorway (whole national network) goes to the transport fund. The transport fund subsidises the rail link Additional surplus of alpine motorways goes to the Alpine fund which subsidises the rail link Rail revenues go to rail operator or manager

revenue is fed to the Alpine fund and the national transport fund is used to top up the amount required to subsidise the rail link. The purpose of considering these last two schemes is to assess the relative welfare performance of financing by the national transport fund and financing with mark-ups on the Alpine motorway tolls.

The MOLINO model is used in a "road/rail, passenger/freight" configuration. In the hierarchical decision tree the top-level choice is between transport and other consumption or other inputs (freight), the second choice is between travelling in peak or off-peak periods and the third is the choice between existing motorways and the future rail link. Other assumptions are detailed in Appendix A.1.

3.2. Results

Table 11.12 reports the welfare effects of Schemes 2–5 relative to the "current toll + public subsidies" scheme as a benchmark.

3.2.1. Schemes 2 and 3: Alpine Fund with Road Toll Mark-ups

The 25 per cent toll increase embodied in Scheme 2 has a positive impact on overall welfare. Central government net revenues rise appreciably (+36 per

Table 11.12. Results for the Lyon–Turin Rail Project for the Different Schemes.

	Scheme 2	Scheme 3	Scheme 4	Scheme 5
	Cross-financing: road toll mark-ups of 25%	Cross-financing: road toll mark-ups of 80%	National transport fund	National transport fund + road toll mark-ups of 25%
Annual welfare gain[a]/ construction cost (%)	0.001	0.003	0.015	0.015
Change in:				
Central government net revenue[a] (%)	36	89	137	155
Public subsidies[a]	−14	−24	−100	−100
Road operator profits[a] (%)	0	0	0	0
Rail operator profits[a] (%)	18	57	0	18.4
Rail infrastructure manager profits[a] (%)	19	58	0	18.8

[a]Compared to the "current toll + public subsidies" Scheme 1.

Table 11.13. Results for Model Shares.

	Current Motorway Tolling Scheme	25% Toll Increase		80% Toll Increase	
		Peak	Off-peak	Peak	Off-peak
Passenger					
Rail (%)	40[a]	43	47	49	61
Road (%)	60	57	53	51	39
Freight					
Rail (%)	50[a]	57	58	69	72
Road (%)	50	43	42	31	28

[a]According to LTF (2003) forecasts.

cent) mainly because of a 14 per cent reduction in public subsidies. However, this has a low impact on overall welfare (ratio of annual welfare gain on construction cost of +0.001 per cent). Profits increase for both rail operators and managers thanks to the modal shift from road to rail. Profits for road operators are not affected because the surplus generated from toll mark-ups is transferred to the fund.

Scheme 3 differs from Scheme 2 in applying a larger toll mark-up. The welfare change is three times higher and the impacts on the various parties are all much larger. The road toll increase generates an increase in central government net revenues of nearly 90 per cent. The considerable increase, of approximately 57 per cent, in the profits of rail operators and infrastructure managers is noteworthy.

The effects of the motorway toll mark-ups on the road/rail modal split are shown in Table 11.13.

For passenger transport rail's share increases with the level of mark-ups on tolls, and is greater during the off-peak than the peak. Starting from an initial share of 40 per cent, it is only when the road toll increases by 80 per cent that rail passenger becomes dominant in off-peak period (61 per cent).

The pattern is similar for freight demand except that the drop in road's share is higher than for passengers for the same toll increases. Starting from an initial share of 50 per cent, rail becomes dominant in both peak and off-peak periods.

3.2.2. Schemes 4 and 5: National Transport Fund Either alone or in Combination with an Alpine Fund

For the last two schemes the rail link is funded from the national transport fund: either alone or as a supplement to the alpine fund. Since the use of the national transport fund eliminates the need for public subsidies the overall welfare increase[20] is higher (ratio of annual welfare gain on construction cost of + 0.015 per cent) when compared with the previous Schemes 2 and 3.

Because public subsidies are eliminated the central government's net revenues increase greatly (+ 137 per cent and + 155 per cent). The only appreciable difference between the welfare effects of Schemes 4 and 5 is that the rail operators and infrastructure managers benefit in Scheme 5 but not 4. From a political economy perspective this might be a reason in favour of financing the rail link partly by a toll mark-up on the Alpine motorways.

4. CONCLUSIONS

Three main conclusions can be drawn from the analysis of the motorway programme financing and pricing (Section 2). The first lesson relates to financing. When compared with public subsidies from the general budget, cross-financing from existing motorways to new motorways increases the level of welfare for all the projects. This overall result is a consequence of the fact that public subsidies are more costly (i.e. the MCPF) than raising revenue from larger tolls on existing motorways via an infrastructure fund. Moreover, the welfare improvement varies directly with the MCPF. The main conclusion is the advantage of additional road tolls for cross-financing (e.g. through a transport fund) when compared with subsidies coming from public money.

Second, with regard to optimal pricing, MSCP yields an increase in overall welfare compared with the current tolling scheme, despite the low level of congestion anticipated for the selected projects. This can mainly be explained by the fact that road traffic will increase since the overall costs borne by road users would fall by 65 per cent with a switch from fuel taxes plus planned tolls to MSCP. However, because of the low level of congestion, pure MSCP cannot solve financing problems. It must be supplemented by subsidies from a transport fund or from central (or local) governments. This is not to say that new investments are not justified. Indeed, motorway projects are typically planned not so much to reduce congestion as to provide better roads with a higher speed limit and improved safety. This improvement induces a socio-economic rate of return high enough to justify these projects.

Third, since pure MSCP is more efficient as regards pricing (without considering financing problems) and cross-financing appears to be more efficient as a means of financing new motorway projects, the combination of the two rules would increase overall welfare. Indeed, for all projects,[21] the scheme that combines pure MSCP with cross-financing provides the greatest increase in overall welfare of the alternative schemes.

Concerning the rail–road case study, the assessment of the Lyon–Turin project also yields three main conclusions. First, cross-financing rail by applying toll mark-ups on alternative Alpine motorways has a limited impact on welfare. However, there would be a higher redistribution towards low-income passengers compared with high-income ones, and the financial balances of the rail operator and manager would be improved while the rail mode share would increase.

Second, it should be stressed that while public subsidies amount to 88 per cent of the construction costs in the first scheme (no cross-financing) the level of public subsidies decreases by 14 per cent and 24 per cent with cross-financing by motorway toll mark-ups of respectively 25 per cent and 80 per cent (the possibility of increasing rail prices has not been considered).

Third, the introduction of the same national transport fund as in the road case study would yield some advantage when combined with the Alpine fund supplied by toll mark-ups, by eliminating the need for public subsidies while simultaneously improving the financial balances of the rail operator and manager.

Moreover, the road toll mark-ups apply only to traffic crossing the Franco–Italian border through the Mont Blanc and Fréjus tunnels. The toll mark-up base could be widened to all traffic using the Alpine motorway network, on the premise that this traffic would benefit from lower road congestion resulting from modal transfer to rail. Widening the base would yield much higher revenues. However, this option raises policy and equity issues which require more thorough analysis.

NOTES

1. Subsidy levels depend on the expected toll revenues, which in turn depend on anticipated traffic forecasts. The forecasts and planned tolls reported in IGF, CGPC (2003) are taken as given.
2. In the case of freight transport this means the choice between transport inputs and other inputs.
3. However, in our modelling exercise this traffic may vary in response to toll variations.
4. This cost is discussed in more detail in Section 2.2.2.
5. Taking the AFITF's initial figures, it is assumed that the fund's financing capacity amounts to €635M/year (in the first years) to reach a cumulative total of €11,482M in the period 2005–2012.

6. See the discussion of the regional development tax in Box 11.1.

7. Since traffic that leaves the motorways can take the alternative free highways this procedure overstates the welfare loss as long as congestion and other external costs on the free highways are not too great.

8. Annual heavy goods vehicle (HGV) traffic is 12 billion vehicle-kilometre, and it pays an average toll of €0.1964 per km. Annual car traffic is 62.4 billion vehicle-kilometre and the average toll paid is €0.0687 per km (DAEI/SES-INSEE, 2004). Using these figures the welfare loss is €11.9 million (€11.1 million for cars and €0.8 million for HGV). Since the tax yields an annual revenue of €520.8 million, the welfare loss is €0.023 per euro collected.

9. This ratio is computed for each project as the annual welfare gain, for the project, of the scheme under assessment compared with the reference scheme, divided by the construction cost of the project, and for the whole of the projects as the sum of annual welfare gains of the scheme under assessment compared with the reference scheme, divided by the total construction cost of all the projects.

10. For all the projects except one (A51) for which the increase of welfare gain is only 8 per cent.

11. MSCP here includes congestion costs and costs that are external to users (see Appendix A.1).

12. The alternative to A24, A1, is tolled but the current tolling on A1 is left unchanged.

13. Construction costs are deemed to have been paid from another source that is not considered in the welfare computation. Consequently, the benchmark scheme is referred to simply as the "planned tolling scheme" in Table 11.5 rather than "planned tolling scheme + public subsidies" as in Table 11.2.

14. However, the off-peak motorway tolls are higher than the peak tolls for both freight and passenger transport services for 4 of the 10 projects (A48, A51, A89 and A45). This shifting peak can be explained due to our assumptions about elasticities of substitution (see Table A.1), i.e. that peak-period demand is more price-elastic than off-peak demand.

15. With a toll of €3.74 for passenger vehicles and €1.72 per tonne for freight.

16. The timing of investment is identical in the different alternatives.

17. This figure is based on the Directive proposal COM (2003) 448 which suggested that toll markups of up to 25 per cent be allowed on motorways in environmentally sensitive areas in order to cross-finance alternative rail routes.

18. This is the toll level that maximises the gross operating surplus of motorway operators, taking into account the toll elasticity of road demand. Such a large increase might be difficult to implement for acceptability reasons.

19. In this scheme the project is so expensive that the entire capacity of the national fund would be required to fund it during its construction period (i.e. about 10 years).

20. Note that here again this welfare variation includes the welfare loss incurred by motorway users whether they pay the regional development tax on the existing national toll motorway network or the 25 per cent mark-up on toll on the Alpine motorways.

21. Except Highway A51.

ACKNOWLEDGEMENTS

Financial support from the European Commission (EU Fifth Framework research project REVENUE (DG TREN, Contract No. GMA2-2002-52011)) is gratefully acknowledged. The authors also wish to thank the editors who provided helpful comments on preliminary drafts.

REFERENCES

ASF (2004). Comptes consolidés 2003, Paris.

ATMB (2004). Rapport annuel 2004, Paris.

Commissariat Général du Plan (CGP) (2001). *Transports: choix des investissements et coût des nuisances.* Rapport du groupe présidé par Marcel Boiteux. Paris: La Documentation Française.

DAEI/SES-INSEE (2004). *Les comptes des transports en 2003*. Paris: INSEE.

DATAR (2002). *La France en Europe: Quelle ambition pour la politique des transports?* Dominique Parthenay et Michel Vermeulen (dir) – Documentation française.

ECMT (2003). *Reforming transport taxes*. Paris: Les éditions de l'OCDE.

Girault, M., Fosse, M., & Jeger, F. (2000). Circulation, Consommation de carburant en France. In: *Notes de synthèse du SES* (pp. 15–22). Paris.

Hivert, L. (2000). Le parc automobile des ménages. *Etude en fin d'année 1998 à partir de la source "Parc Auto" SOFRES*. Rapport sur convention INRETS/ADEME. Arcueil: INRETS.

IGF, CGPC (2003). Inspection Générale des Finances, *Conseil Général des Ponts et Chaussées*. Rapport d'audit sur les grands projets d'infrastructures de transport, Annexe II-F ≪Les projets ferroviaires≫.

Lebègue, D., Hirtzman, P., & Baumstark, L. (2005). *Le prix du temps et de la décision publique*. Révision du taux d'actualisation. Paris : La documentation Française.

LTF (2003). Lyon Turin Ferroviaire. Nouvelle liaison ferroviaire Lyon-Turin, Synthèse des études de trafic, économiques, et socio-économiques. Soumission no 25.

MINEFI (2004). Rapport sur l'Etat actionnaire 2004. Paris.

SNCF (2002). *Mémento de statistiques 2001. Exercice 2001 et rappel de l'exercice 2000*. Paris: SNCF.

A.1. APPENDIX

A.1.1. Assumptions and Supplementary Data

Traffic data: Road traffic is based on the audit data (IGF and CGPC, 2003). Local traffic is considered to be domestic while transit traffic corresponds to traffic both entering and leaving France. It is assumed that 80 per cent of road freight traffic is local.

Occupancy rate: For roads, the hypothesised occupancy rates for the two case studies are 8.6 tonnes for freight vehicles and 1.9 passengers for passenger vehicles. For rail, the occupancy rates are 400 passengers/train and 280 tonnes/train (LTF, 2003).

Congestion periods: Congestion duration is estimated for each new motorway project according to average figures based on road capacity and daily traffic load (see DATAR, 2002). Eight per cent of passengers (cars and rail) and 2 per cent of heavy goods vehicles (HGV) and freight trains use infrastructures in peak periods (see ECMT (2003) case study for France).

Monetary costs: For passenger road vehicles, resource costs are made up of purchasing expenditures, financial expenditures, car insurance, fuel (without taxes), maintenance (without taxes) and parking. For freight road vehicles, they include fuel, tyres, maintenance, insurance, purchasing or hiring expenditures and staff wages. Central taxes essentially consist of fuel taxes for passenger road vehicles. For road freight vehicles only fuel tax and axle tax have been considered (because VAT is recovered by firms). The level of "central taxes on transit freight" has been considered to be 40 per cent of "central taxes on local freight" (Girault, Fosse, & Jeger, 2000).

Speeds: For both road and rail modes, maximum speed is considered to be identical for passenger and freight. The same applies to peak period speed. In the road case study, the maximum speed is higher for motorways (based on the 130 km/h speed limit on French motorways) than for "conventional" highways (between 70 km/h and 100 km/h according to local topography). The speed under congested conditions is fixed (as required by the MOLINO model) at 60 km/h. In some cases mountainous terrain reduces both motorway and highway speeds.

Time costs: The value of time for passengers amounts to €12/h for distances up to 310 km (Commissariat Général du Plan (CGP), 2001). Operating costs are estimated at €31.4/h for road freight vehicles, i.e. €3.65/tonne/h.

Share of household expenditure devoted to transport: 15.02 per cent of household income is considered to be spent on transport consumption (DAEI/SES-INSEE, 2004). These data make no distinction between low and high income. Concerning freight transport expenditures, a single rate of 3.4 per cent has been applied.

Elasticities of substitution: With regard to transport/other consumption and peak/off-peak elasticities the values of the TRENEN model (interregional model) and reported in (ECMT, 2003) have been used. For peak and off-peak elasticities between motorways and highways we have sought values that give an acceptable value for the toll-elasticity of demand on toll motorways: investigation of some projects has given empirical values of between -0.4 and -0.5 on French toll motorways. The values of elasticities of substitution are shown in Table A.1.

Other costs: For the road case study, the motorway costs are based on data from the ASF (2004) report. In France, there is no distinction between the motorway manager and the motorway operator, so maintenance costs have been included with ‘‘variable operating costs’’. The detailed financial charges for the new road projects are not known: from this it follows that the ‘‘profit’’ (of road operators and infrastructure managers) used in the MOLINO model is actually the GOS. However since the scenarios have only been compared with each other, the variations in welfare that result from variations in the GOS are correct.

External costs: In 2001, transport pollution was estimated at €0.009/vehicle/km for passenger vehicles and €0.062/vehicle/km for freight (CGP, 2001).

Welfare parameters and weights: When the operator is private, as is the case with road and rail operators, the profit tax ratio is equal to 0.35. Central government revenues are considered as equally allocated between the high- and the low-income users.

Public subsidies vs other funding: MOLINO is used in a *static* approach, that is to say as if the new projects are implemented and in operation. MOLINO computes the traffic and financial flows on a daily basis. In order to establish a common daily basis of comparison between construction subsidies and private or transport fund financing, it is assumed for each project that public subsidies would be financed by a government loan of a 30-year duration with a 4 per cent interest rate. In this way it is possible to compute a daily level of public subsidy broken down for the duration of the concession. This rule is also applied in the rail–road (Lyon–Turin) study.

Assumptions and additional data specific to the rail–road case study: The rail operator (SNCF) is private while the rail manager (RFF) is public. The road operator and manager are both private.

The traffic data given by the two motorway operators (ATMB and SFTRF) through the Mont Blanc and Fréjus tunnels (both between France and Italy) in 2003 have been aggregated. An increase of 3 per cent per year was applied in order to obtain traffic forecasts for 2015.

Rail-traffic data are estimated as rail passenger traffic in Modane (night and day trains) according to the ‘‘Lyon–Turin Ferroviaire’’-scenario V4 (the scenario features an annual growth rate in demand of 1.8 percent, a tunnel speed of 220 km/h, €20 of additional tax on international trains, and a 10 percent decrease in air fares triggered by the advent of low-cost air carriers). Freight traffic forecasts have been considered at the border crossing (LTF, 2003).

High- and low-income passengers are distinguished for both road and rail traffic. It has been assumed that 33 per cent of passenger traffic is drawn from low income households (Hivert, 2000) (incomes of €1,900 per "consumption unit" per month or less).

Road toll levels are those applied for a trip between Lyon and Turin according to the type of vehicle. Rail tolls for passengers correspond to the ticket price paid to the railway operator SNCF by travellers. Low-income passengers are considered as travelling second class while high income travel first class. Freight rail tolls are based on the SNCF price for freight traffic.

For the maximum rail speed, the maximum passenger speed (220 km/h) and freight speed (120 km/h) were weighted in proportion to the share of each, i.e. a maximum speed of 130 km/h. For peak period speed the assumption was that travel time for passenger and freight trains is increased by 1 h. As stated above, the passenger speed (128 km/h) and the freight speed (86 km/h) were weighted in proportion to the shares of passenger and rail trains, giving 90 km/h as an average.

Motorway costs were computed from specific operating costs for Alpine motorways, based on financial reports (see ATMB, 2004; MINEFI, 2004). Rail operating costs are taken from the SNCF (2002) report. They include operating costs without infrastructure tolls and make a separation between freight trains and high-speed trains.

The rail infrastructure cost is the infrastructure tolls paid by the rail operator (in our case SNCF) to the infrastructure manager (RFF): data come from the LTF study.

Table A.1. Elasticities of Substitution.

		Transport/Other	Peak/Off-Peak	Peak Highway or Rail/Motorway	Off-Peak Highway or Rail/Motorway
Passengers	Low and high income	0.4	0.8	18	6
Freight	Local and transit	0.2	0.5	4	3.5

CHAPTER 12

SYNTHESIS OF CASE STUDY RESULTS AND FUTURE PROSPECTS

André de Palma, Robin Lindsey and Stef Proost

ABSTRACT

This chapter summarises the results of the case studies and assesses their main insights. It also draws on recent experience with congestion pricing in London and elsewhere to consider the prospects of successful implementation of efficient pricing and revenue-use schemes.

1. INTRODUCTION

The research covered in this book has two main objectives: first to develop guidelines for good use of the revenues from marginal social cost pricing, and second to compare current practice with the guidelines using a set of case studies. Part I of this volume focused on the first objective, while Part II concentrated on the second. This chapter concludes Part II by summarising the case studies and the insights they yield. It also draws on recent experience with congestion pricing in London and elsewhere to consider the prospects of successful implementation of efficient pricing more widely.

Investment and the Use of Tax and Toll Revenues in the Transport Sector
Research in Transportation Economics, Volume 19, 269–297

ISSN: 0739-8859/doi:10.1016/S0739-8859(07)19012-0

2. CASE STUDY RESULTS

The main features of the case studies are summarised in Table 12.1 and the major results are listed in Table 12.2. As discussed in Chapter 1 of this volume, the studies assess a range of scenarios that span the status quo, policy proposals and policies developed for the case studies that may be more efficient and/or more acceptable than either existing or proposed policies. Chapter 1 identifies and analyses nine problems related to transport user charges and revenue use that are addressed in this volume. Of the nine problems, six were taken up by at least one of the case studies. The summary of results presented here is organised around these six problems.

2.1. Questions about Efficiency

This section summarises the main results of the case studies regarding efficiency of charging policy and revenue use that were posed as Problems 1–4 in Chapter 1.

2.1.1. What Charges Should Users Pay? (Problem 1)

2.1.1.1. Oslo. The Oslo study takes as its starting point the user charge and revenue-use policies of Oslo Packages 1 and 2, and considers variations on them. In Oslo Packages 1 and 2, the cordon toll and public transit fares are not based on either first-best or second-best pricing principles, but are rather set at politically ‘‘acceptable’’ levels that yield sufficient revenues to finance selected investment projects. Only in the prospective Oslo Package 3 is marginal social cost pricing (MSCP) implemented by replacing the cordon toll with a road-pricing scheme.

The first alternative policy experiment considered involves jointly optimising public transport (PT) fares and service frequency within the framework of Oslo Package 1 while holding the cordon toll and road investment program fixed.[1] This policy calls for a reduction in fares and a large increase in frequency. Doing so results in a substantial welfare gain that is attributable to the severe underpricing of peak-period car travel and low PT service frequency in Oslo Package 1. The study shows that, if car travel were efficiently priced, fares should rise rather than fall and the welfare gain from optimising public transport would be appreciably smaller.

Similar results are obtained for Oslo Package 2 although the welfare gains are higher than for Package 1 because, with a higher cordon toll, more revenues are available to improve PT services. Oslo Package 3 yields

Table 12.1. Case Study Characteristics.

	Oslo (Chapter 7)	Edinburgh (Chapter 8)	Switzerland (Chapter 9)	Germany (Chapter 10)	France (Chapter 11)
Subject of study	Oslo Packages 1, 2, 3 successive pricing, earmarking and investment packages for Oslo	Proposed double inbound cordon toll of £2/day for each cordon	Efficiency and equity effects of rail fund financed from the HVF[a]	Alternative uses of revenues from German motorway HGV[b] toll	Cross-financing from existing motorways to either new motorways (Pt 1) or rail projects (Pt 2)
Investments, charges and financing schemes (*alternative schemes identified in italics*)					
Investments	PT[c] infrastructure and rolling stock	Base: 2 tramlines *Base+additional: 3 tramlines & higher bus frequency*	Two new railway tunnels *New rail tunnels+road capacity*	Roads: motorways, trunk roads Rail: network capacity, facilities, rolling stock	Pt 1: 10 new motorways Pt 2: Lyon–Turin rail link
Charge types	PT[c] fares Cordon toll Road pricing	Inner and outer cordon ring tolls	Fuel and vehicle taxes, car vignettes, HVF[a], railway track charges	HGV[b] motorway toll Hypothetical road, rail and inland waterway charges for passenger and freight traffic	Pt 1: tolls on new motorways. Pt 2: tolls on existing motorways
Charge levels	Oslo Package 1: cordon toll €1, existing PT[c] fares MSCP[d] of PT within Package 1 *Oslo Package 2: cordon toll €1.25, PT[c] fare hike of €0.10* *Oslo Package 3: MSCP*[d]	Base: No cordon charge Base + additional: £2 Welfare-max for region *Welfare-max for Edinburgh residents*	Status quo *Internalisation of congestion and environmental externalities, fuel taxes in place* *MSCP*[d] *of congestion, environmental and infrastructure costs, all existing taxes removed*	Status quo (average cost pricing) Average cost *pricing of infrastructure costs for all vehicles on for all vehicles on all roads using current taxes* *Privately set tolls* *MSCP*[d] *of all modes*	Pt 1: planned tolls on new motorways *MSCP*[d] *on new and (currently toll-free) competing motorways+motorways+abolition of fuel tax* Pt 2: current tolls *Higher (+25% or +80%) motorway tolls*
Funds	–	–	Rail fund (FINÖV): financed from HVF[a], fuel taxes, VAT. Road fund: Financed from HVF[a] in rest of SWZ	Transport fund (in MOLINO model only)	Pt 1: transport fund for motorways Pt 2: Alpine motorway fund for cross-financing rail *Transport fund as Pt 1*
Matching funds	National government matches toll ring revenues	–	Federal government has provided loans to FINÖV	–	–

Table 12.1. (*Continued*)

	Oslo (Chapter 7)	Edinburgh (Chapter 8)	Switzerland (Chapter 9)	Germany (Chapter 10)	France (Chapter 11)
Earmarking of revenues	Oslo Package 1: 80% to road infrastructure, 20% to PT[c] and slow modes Oslo Package 2: PT[c] fare increase to PT[c] rolling stock; toll increase to PT[c] infrastructure *Oslo Package 3: flexible*	Revenue earmarked to local transport sector *Revenue used to reduce general taxes*	HVF[a] revenues: 7% to operating costs; for remainder 2/3 to rail fund, 1/3 to local government Three alternative scenarios with HVF[a] revenues to rail fund, road fund and local government in varying proportions	20% to toll operations; 40% to roads; 30% to rail; 10% to inland Waterways Roads only Roads and rail General budget	Pt 1: to new motorways Planned tolls on new motorways to concessionaire. Public subsidies cover remaining construction costs *As above+tax on existing toll motorways to new motorways* Pt 2: to rail
Institutional set up (status quo and alternatives)					
Management of infrastructure	Public	Public	Public	Roads: private Rail: private	Pt 1: roads private Pt 2: rail private (RFF[e]), road private
Operation of services	Public	Public	Public	Roads: private Rail: public under AC pricing, private under MSCP	Pt 1: road private Pt 2: rail private (SNCF[f]), road private
Setting charges	–	CEC[g] *Higher authority*	–	Federal govt except private in private-sector scenario	Pt 1&2: federal government
Collecting revenues	–	CEC[g] (through Transport Initiatives Edinburgh)	–	State	Pt 1&2: federal government
Allocation of revenue	–	CEC[g]	–	State (transport infrastructure financing society)	Pt 1&2: federal government
Multiple governments	Yes	Yes	No	No	No
Alternative revenue sharing arrangements	Yes	CEC[g]: 54%, Lothians: 46% *All to CEC[g]*	–	–	–

Perceived political acceptability constraints	Political acceptability dominated efficiency during formulation of Oslo Package 2	By Scottish law all revenues earmarked for transport. CEC[g] subject to approval by Scottish Executive, proposal subject to public inquiry and approval	Earmarking of revenues to rail critical for acceptability of HVF[a] Internalisation pricing scheme assumes fuel taxes remain in place	Earmarking of revenues to transport encouraged by EU. Also necessary in Germany for acceptability	–
Model and case-study design					
Model(s)	FINMOD (strategic transport model)	MARS (strategic land use and transport model)	MOLINOinGAMS (Variant of MOLINO model that accounts for lag between investment costs incurred and project completion)	1. MOLINO (partial equilibrium assessment tool for pricing, investment, financing and regulatory schemes) 2. ASTRA (system dynamics model with interactions between transport sector and rest of economy)	MOLINO (partial equilibrium assessment tool for pricing, investment, financing and regulatory schemes)
Time horizon	2004	2006–2036	2000–2040	2000–2020	Representative future year Pt 2: 2015
Discount rate	7% per annum	3.5% per annum	2% per annum	2.5% per annum	4% per annum
Marginal cost of public funds (MCPF)	Base case: 1.25	Base case: 1	Labour tax: 1.35 Rail and road funds: 1	Base case *Sensitivity: 1.35*	Pt 1: 1.1 (public subsidy), 1.023 (tolls, endogenous) *Sensitivity: 1.5 for public subsidy* Pt 2: MCPF[h] not specified
Assessment of acceptability	Internet survey of politicians and admin.	Interviews with stakeholders and analysis of the policy process	–	Internet survey of road hauliers, shippers, logistic providers and combined transport operators	–

Source: Case study chapters.
[a]HVF = Heavy vehicle fee.
[b]HGV = Heavy goods vehicle.
[c]PT = Public transport.
[d]MSCP = Marginal social cost pricing.
[e]RFF = Réseau Ferré de France.
[f]SNCF = Société Nationale des Chemins de Fer.
[g]CEC = City of Edinburgh Council.
[h]MCPF = 1 + marginal excess burden.

Table 12.2. Case Study Results.

	Oslo (Chapter 7)	Edinburgh (Chapter 8)	Switzerland (Chapter 9)	Germany (Chapter 10)	France (Chapter 11)
Welfare effects of pricing reform	Oslo Package 1: raising PT[a] fares beneficial Oslo Package 2: higher tolls and PT[a] fares welfare-enhancing Oslo Package 3: road pricing yields highest welfare gains. Comes with much higher peak-period PT[a] fares	Base cordon toll highly beneficial. But optimal toll yields much higher welfare gains and revenues	Welfare increases with higher road prices and with higher rail prices if revenues used to reduce existing distortionary taxes	Welfare gain from MSCP[b] dominates gain from efficient earmarking. But road and rail tariffs are very high with MSCP[b]	Pt 1: MSCP[b] + abolition of fuel taxes raises welfare slightly Pt 2: 25% and 80% toll increases raise welfare
Welfare effects of earmarking to transport vs. the general budget	–	Earmarking to transport possibly welfare-reducing due to few options for worthwhile transport investments	–	Earmarking to general budget superior according to MOLINO model. But ASTRA model favours transport investments	–

Welfare effects of cross-subsidisation between modes	Earmarking toll revenue for PT[a] service improvement beneficial for Oslo Package 2 and Package 3.	–	Welfare increases mono. with per centage earmarking of HVF[c] to rail tunnel investments rather than to local government	Conditional on earmarking to transport, efficient to invest in roads rather than rail or transit	Pt 2: replacing public subsidies with Alpine motorway toll revenues enhances welfare slightly. Supplementing with national fund eliminates public subsidies but yields little further benefit.
Welfare effects of cross-subsidisation within modes	Existing allocation of revenue between PT[a] modes and regions is near-optimal for all three Oslo packages	–	–	Maintenance preferable to investments for motorways.	Pt 1: replacing public subsidies for new motorways with toll revenues from existing motorways enhances welfare slightly
Trade-offs between efficiency, equity and acceptability	Oslo Package 2 embodied many compromises to gain universal acceptance, yet still yields large welfare gains	Raising tolls boosts revenue, but impairs acceptability. Also possibly welfare-reducing given limited transport investment options	–	Trade-off between charges and revenue use driven by acceptability. MSCP[b] yields large gains, but tariffs are very high	–
Loss of efficiency or acceptability with multiple governments or constituencies	–	Maximising welfare for Edinburgh residents results in substantial welfare loss for region	–	–	–

Table 12.2. (*Continued*)

	Oslo (Chapter 7)	Edinburgh (Chapter 8)	Switzerland (Chapter 9)	Germany (Chapter 10)	France (Chapter 11)
Main policy recommendations and/or lessons	1. Road pricing is superior to public funds as a revenue source for PT[a]. 2. Earmarking is a justifiable concession to gain policy approval, particularly if enhanced efficiency becomes a priority after approval	1. Charge-setting and revenue allocation decisions should be vested with higher authority 2. Earmarking all revenues to transport could be relaxed 3. Regional support for a policy must be gained with an agreed and clearly committed use of revenue that is seen as efficient and fair	1. Raising prices for one mode to cross-subsidise another mode can be welfare-enhancing 2. Pricing, investment and revenue use decisions should be made jointly rather than in isolation	1. Earmarking motorway revenues to the road sector enhances acceptability. But allocation to the general budget may well be welfare-superior 2. Efficiency would be enhanced by implementing MSCP[b] with the HGV[e] charge rather than charging average costs	1. With a high MCPF[d] cross-financing either between modes or within modes can enhance welfare by reducing subsidies from general fund 2. New investments are not self-financing under MSCP[b] if congestion levels low, although investments may still be warranted by enhanced speed and safety

Source: Case study chapters.
[a]PT = Public transport.
[b]MSCP = Marginal social cost pricing.
[c]HVF = Heavy vehicle fee.
[d]MCPF = 1 + marginal excess burden.
[e]HGV = Heavy goods vehicle.

considerably higher gains yet. Three factors account for this: the greater efficiency of car travel induced by MSCP, the scope to raise PT fares and fare box revenues without diverting travellers to car and the greater revenues available to improve PT services without increasing the public subsidy.

Perhaps the most important lesson from the Oslo pricing experiments is that even crude pricing schemes can be highly beneficial if existing prices and service levels are far from optimal in the status quo.

2.1.1.2. Edinburgh. The congestion-charging scheme planned for Edinburgh featured a £2 toll for the city centre cordon from 7:00 to 18:30 and a £2 toll for the outer cordon from 7:00 to 10:00, with a maximum daily charge of £2 regardless of the number of crossings of either cordon. The case study determined that these charges are far below levels that maximise welfare for the region, and yield only about half the maximum potential welfare gain. It also found that a further modest welfare gain could be realised by eliminating the toll during the off-peak period (although toll revenues would drop precipitously). Thus, as in the Oslo study, the proposed charging scheme is welfare-enhancing relative to the status quo, but far from optimal.

2.1.1.3. Switzerland. Three pricing regimes are assessed for the Swiss study: (1) the status quo, (2) an "internalisation" scheme in which Pigouvian charges are imposed on congestion and environmental externalities while existing transport taxes remain in place, and (3) pure MSCP with existing taxes abolished. Although the MSCP regime is optimal by definition under first-best conditions, the internalisation scheme actually produces a higher welfare gain because it generates more revenues, which are assigned a premium in the objective function by the marginal cost of public funds (MCPF). Thus, the welfare loss in the internalisation regime from pricing transport above social marginal cost is more than offset by the reduction in excess burden from subsidies paid out of the general revenue fund.[2]

Another notable finding of the study is that, for both the internalisation and the MSCP schemes, low-income and high-income passengers end up better off whereas local and transit freight traffic end up worse off. This is a consequence of the assumption that revenues are used in ways that benefit passengers but not freight transporters.

2.1.1.4. Germany. The German study examines three pricing regimes: (1) the existing heavy goods vehicle (HGV) charge which is limited to vehicles above 12 tonnes gross vehicle weight using motorways, (2) charging all vehicle types on all roads for average infrastructure costs using existing tax instruments, and (3) MSCP of all modes to internalise infrastructure, congestion and environmental costs. The study finds that a shift from regime (1) to regime (2) benefits users but causes welfare to drop, whereas a shift from regime (1) to regime (3) has the opposite effects. MSCP entails large price increases relative to (1) or (2) that reflect the large marginal external costs of transport.

2.1.1.5. France. Part 1 of the French study examines alternative ways to finance construction of 10 preselected motorways. Two pricing schemes for the motorways are analysed: (1) planned tolls on the new motorways, and (2) abolition of the fuel tax in combination with MSCP on the new motorways as well as MSCP on competing motorways that are presently not tolled. Shifting from planned tolls to MSCP causes tolls to fall because the new motorways are not heavily congested and the planned tolls exceed first-best Pigouvian levels. The efficiency gain on motorways from MSCP is approximately offset by the loss of revenues from abolition of the fuel tax (total revenues decrease by over half). Furthermore, the reduction in new motorway tolls attracts some traffic off competing highways. The impact of this shift varies across motorway projects and a slight welfare loss occurs on some corridors.

Part 2 of the French study explores the use of toll revenues on existing motorways to cross-subsidise rail. Raising tolls (either by 25 per cent according to European Commission (EC) guidelines, or by 80 per cent to maximise revenues) increases welfare because public subsidy requirements are reduced.

2.1.1.6. Summary. Overall, the case studies show how the effects of transport pricing reform vary with the *status quo* pricing regime as well as the accompanying investment and funding policy. More pointedly, they reveal that the welfare effects of MSCP depend on whether charge revenues rise or fall, and on the value assumed for the MCPF. The Swiss and French case studies demonstrate how MSCP is not optimal in the presence of fiscal distortions.

Problems 2 and 3 concern how to allocate revenues from user charges. Following Fig. 1.1 in Chapter 1, which is reproduced here as Fig. 12.1, the allocation process can be broken down into three steps: allocation between transportation and the rest of the economy (i.e. to the general budget), allocation of the transport revenues to modal funds and allocation of the fund revenues between specific alternative expenditures. The first and second steps are treated here under Problem 2 and the third step in Problem 3.

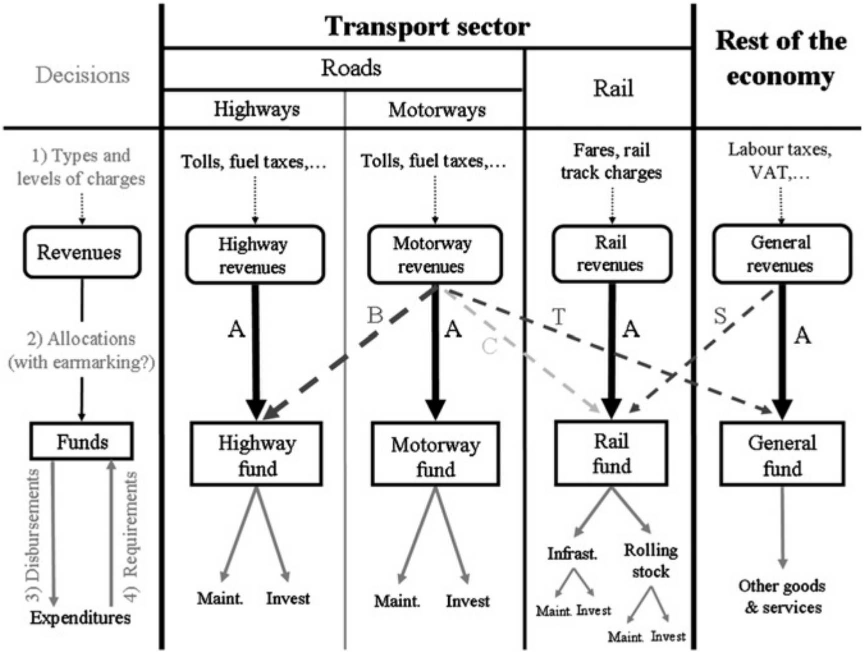

Fig. 12.1. Schema for Transport User Charge and Revenue Use Decisions. *Source:* Authors' Creation.

2.1.2. How Should Revenues be Allocated to Funds? (Problem 2)

2.1.2.1. Allocating Revenues between Transport and the General Budget. The relative merits of allocating user charge revenues between the transport sector and the general budget are assessed in the Edinburgh and German studies.

2.1.2.1.1. Edinburgh. By law, revenues from the proposed Edinburgh congestion-charging scheme would have been earmarked to finance local transport, and a particular set of investments had been identified. Although earmarking funds in this way may have enhanced public acceptability of the scheme (see Section 2.2 below), the Edinburgh study determined that the number of economically worthwhile public transport projects in the region is limited and all the revenues could probably not be efficiently used.[3] Thus, earmarking of revenues to local public transport appears to be inefficient in the Edinburgh setting.

2.1.2.1.2. Germany. Whereas the Edinburgh study examines the merits of particular transport investments, the German study considers allocating revenues in varying proportions between the transport sector and the general budget in order to lower direct taxes. Both the MOLINO and ASTRA models were used for the assessment. The conclusions were at odds, with the MOLINO model favouring the allocation of all revenues to the general budget, whereas according to the ASTRA model revenues should be devoted wholly to transport sector investments. One reason suggested in Chapter 10 for the discrepancy is that, in the ASTRA model, transport service improvements contribute over time to productivity gains throughout the economy whereas no such feedback occurs in the neo-classical MOLINO model. Since the two models differ in basic structure, other factors may also underlie the diverging results. In any case, the discrepancy is unsettling and calls for further research.

2.1.2.2. Allocating Revenue between Transport Modes. At the second stage of the allocation process, a given amount of revenues has been set aside for the transport sector and must then be apportioned between modes. One choice, examined in the Oslo and German studies, is between recycling the money to the mode on which the charges are levied and using the revenues to cross-subsidise other modes.

2.1.2.2.1. Oslo. In Oslo Package 1, revenues from the Oslo toll ring were earmarked for new road investments. But in Oslo Package 2, dedication was shifted to public transport investments, which introduced a form of cross-subsidisation between modes. Despite the fact that public transport is also funded from other sources, including fare box revenues and extraordinary state funds, the study concludes that cross-subsidisation of public transport from the toll ring is warranted on efficiency grounds because the level of PT service in Oslo is well below optimum.

2.1.2.2.2. Germany. The German study assesses whether revenues from the HGV motorway charge should be allocated to road or to rail transport on the assumption that expenditures would be divided equally between maintenance and investment. Although trucking generates higher environmental costs per tonne-kilometre than rail, road transport turns out to be the preferred choice because of the heavy traffic on German roads and the corresponding large benefits in congestion relief from road investments. Thus, the German study does not support the use of motorway charges to cross-subsidise interurban rail freight transport.

In summary, the Oslo study finds that urban road user charges should be used to cross-subsidise urban public transport, whereas the Germany study concludes that interurban road charges should not be used to cross-subsidise rail. The contrast in these results should not be taken as reflecting any endemic differences between urban and interurban transport, but simply differences in the relative benefits of expenditures on roads and specific competing modes in the particular case-study settings.

2.1.3. How Should Funds be Spent? (Problem 3)

The final step in the revenue allocation process is to spend the funds assigned for particular transport modes. This step was analysed in the Oslo and German studies.

2.1.3.1. Oslo. The FINMOD model used for the Oslo study includes three public transport modes (bus, metro and tram) and two regions (Oslo and Akershus). Oslo Packages 1–3 feature particular allocations of user charge revenues and other funds between the modes and regions. The Oslo study examines whether alternative allocations might improve the overall efficiency of the system, and concludes that little would be gained for any of the three packages. Thus, whereas transport prices in Oslo Packages 1 and 2 appear to be far from optimal (see Problem 1), the distribution of funds between modes and regions appears to be near-optimal.

2.1.3.2. Germany. The German study assumes that revenues from the HGV charge are allocated to motorways and assesses whether road maintenance or capacity expansion is preferable. Maintenance turns out to be the preferred choice because new roads occupy valuable land, induce more traffic and increase long-run maintenance requirements. However, users benefit more from investment than maintenance – in part because they do not incur the external costs associated with new roads.

2.1.4. How Should Predetermined Expenditures be Paid for? (Problem 4)

Problems 2 and 3 concern how to allocate revenues from user charges. In contrast, Problem 4 assumes that investment decisions have been made and examines how to finance them at least social cost. Problem 4 was posed in the Swiss and French studies.

2.1.4.1. Switzerland. Two funding sources have been considered to pay for the two New Alpine Rail Tunnels in Switzerland: the heavy vehicle fee (HVF) and the general fund. According to current Swiss policy, two-thirds of HVF revenues are dedicated to the rail fund for construction, and the remaining third goes to local government. The study finds that, regardless of which pricing regime is implemented, welfare increases with the fraction of HVF revenues dedicated to the rail fund. The reason is that the welfare loss to the local government (the Swiss cantons) from reduced local public spending is more than offset by the fiscal benefits from lower labour taxes.

2.1.4.2. France. Part II of the French study examines how to finance the Lyon–Turin rail link at minimum social cost. The alternative funding sources considered are the Alpine motorway toll and the general fund. Raising the mark-up on the motorway toll to maximise revenue is found to be welfare enhancing, although it does not eliminate the need for public subsidies.

2.1.4.3. Summary. All three of the interurban case studies examine the merits of using motorway toll revenues to cross-subsidise rail investments. As just noted, cross-subsidisation is supported by the Swiss and French study results, but not by the German results. The contrast in results may be driven in part by differences between the German road and rail networks, and the networks in Switzerland and France.[4] However, the contrast may arise simply because the studies investigated different questions. In the Swiss and French case, the particular rail investments were treated as given and the question was how to fund them (Problem 4). In the German study, rail investments were not predetermined and it turned out that motorway toll revenues were better allocated to road investments than to rail (Problem 2).

2.2. *Trade-offs between Efficiency, Equity and Acceptability (Problem 5)*

Prospective transport policy reforms have efficiency and welfare-distributional effects that influence whether the reforms will be implemented. Although the theoretical case is not clear-cut, policies that promise the largest efficiency gains are likely to make some groups appreciably worse off, and therefore may be unattractive on equity and/or acceptability grounds. When formulating and choosing between alternative policies, it is clearly useful to know whether a trade-off between efficiency and equity/acceptability exists and, if so, how policies can be made more palatable without eviscerating their effectiveness. The Oslo, Edinburgh and German case studies provide evidence on such trade-offs from the policy assessment and modelling exercises they conducted as well as from the acceptability surveys. The two sets of results are summarised in the following two subsections.

2.2.1. Evidence from Policy Assessment and Modelling Exercises

2.2.1.1. Oslo. One of the main conclusions of the Oslo study is that politics and bargaining rather than economic considerations dominated the formulation of Oslo Packages 1 and 2. The process of formulating Oslo Package 2 can be characterised as planning through negotiation, with all major parties having the power to veto the package. The modelling results indicate that potential efficiency gains were indeed sacrificed to satisfy acceptability constraints, although earmarking of revenues appears to have been less at fault than departures from efficient pricing. Given the widespread horse-trading that took place, it is heartening that Oslo Package 2 actually did yield some efficiency improvement over Oslo Package 1. Interestingly, the case study attributes this partly to a shift of emphasis after the package was approved from political acceptability to efficiency.

2.2.1.2. Edinburgh. The modelling analysis for Edinburgh concluded that the proposed congestion-charging scheme was inefficient in two respects. First, the £2 toll for each cordon was well below the welfare-maximising level for the region. Second, the requirement that all revenues be earmarked for local public transport improvements was wasteful because of the limited number of economically viable projects. Both the low tolls and the earmarking are interpreted as attempts to gain public support.[5,6] In particular, low charges may have been seen as less harmful to Edinburgh's retail sector as well as conducive to gaining necessary approval from the Scottish Executive. The model also suggests that the City of Edinburgh Council's (CEC) proposed distribution of revenue between the city and outside regions was broadly equitable. Thus, the results clearly indicate that the proposed scheme sacrificed efficiency in an effort to boost its acceptability.[7]

2.2.1.3. Germany. As mentioned in Section 2.1.3, the German study concludes that revenues from the HGV charge are more efficiently invested in maintenance than new capacity but new capacity is more beneficial to users. This suggests a trade-off between efficiency and acceptability.

2.2.1.4. Summary. A trade-off between efficiency and acceptability is evident in all three of the Oslo, Edinburgh and German studies. This adds to the growing body of evidence that first-best pricing schemes (which dominated the early literature on road pricing) are unlikely to be practically feasible, and need to give way to messier second-best schemes that stand a better chance of mustering public and political approval.

2.2.2. Survey Evidence

The Oslo, Edinburgh and German case studies include surveys of attitudes towards transport user charges and revenue uses. These surveys differ with respect to the types of respondents as well as the characteristics of the charging scheme in question. After briefly reviewing the highlights of each survey, a few common findings are identified.

2.2.2.1. Oslo. The Oslo study features a self-administered Internet survey of local politicians and administrators to elicit their preferences for measures such as increased cordon tolls, parking restrictions and reduced transit fares to encourage a modal shift from car to public transport. Respondents indicated that they favoured improvements in transit service, and saw them as more important than car-related measures for assembling acceptable policy packages. They also approved of having public transport users pay for better services, but expressed greater support for earmarking toll ring revenues to help fund public transport.

2.2.2.2. Edinburgh. A range of stakeholders were interviewed to assess attitudes towards the Edinburgh congestion-charging proposal with a view to determining why it was rejected in the February 2005 referendum. The main contributing factors appear to have been flaws in the policy package and uncertainties. Four possible flaws were identified. First, a double cordon ring may have been inferior to a single cordon that would have been easier for the public to understand, and more similar to London's successful area-based scheme. Second, a common complaint of residents was that tolling would have begun before the public transport enhancements were completed. In addition, two of the three tramline investments were to be funded independently of the congestion charge, and thus were not contingent on approval of the charge. Another possible strategic flaw was that the referendum was held before the scheme was launched rather than after a trial period, as had been (successfully) done in Norway and in Stockholm in the autumn of 2006. Finally, residents of adjacent local authorities considered the investment package to be more suited to the needs of city residents than to their own needs.

As far as uncertainties are concerned, Edinburgh's proposal would have been just the second congestion-charging scheme in Britain after London[8] and its double cordon design has not been implemented anywhere in the world. One likely concern for residents and other stakeholders are charge levels. The two cordon tolls were to have been set at a relatively modest £2 with a maximum daily charge of £2. In contrast, London's scheme began in February 2003, with a charge of £5, which was raised, in July 2005, to £8 despite a statement by Ken Livingstone that an increase in the charge was unlikely in the foreseeable future. This pending development may have increased the perceived threat of similar hikes for Edinburgh.[9] Another uncertainty for the local authorities outside Edinburgh was whether revenues to fund their own public transport improvements would be forthcoming since the CEC had no legal obligation to deliver on its revenue-sharing plan. There was also uncertainty about the extent to which the scheme would succeed in alleviating congestion or generate enough revenues to fund the planned investments.[10]

2.2.2.3. Germany. The German study included a stated-preference Internet-based survey of road hauliers. Their responses indicated that they considered the time and monetary costs of congestion to be a severe problem. On balance, their assessment of the HGV motorway charge was relatively favourable – in large part, because they felt it reduced the unfair competitive advantage of foreign truckers who are better able to purchase fuel outside Germany at lower prices. But the hauliers had several objections or concerns about the HGV charge scheme and how it might evolve.

First, they were concerned about the lack of interoperability of the charging technology, and considered TollCollect's 20 per cent share of toll revenues for operating the system to be too expensive.[11] Second, they were opposed to extending charges to the whole road network or to altering the charging scheme to a form of congestion-charging or real-time pricing. They were also sceptical that tolls would reduce traffic because of the limited scope for truckers to modify their choices of route, vehicle size or load factor.[12] Finally, the hauliers were strongly against allocating toll revenues to the general budget rather than to the road sector. They also expressed a preference for revenue collection and allocation decisions to be made either by an independent motorway operator or by two separate government agencies.

2.2.2.4. Summary. Despite the differences between the case studies in the types of actors surveyed and the concerns they raised, several common findings are discernible:

1. The user-pays principle is widely accepted (Oslo, Edinburgh, Germany).
2. There is general support for earmarking (Oslo, Germany).

3. It is imperative to assemble a charging and revenue use package that is both efficient and equitable to users (Edinburgh, Germany).
4. Public transport improvements are essential to secure acceptability of urban road pricing (Oslo, Edinburgh).

2.3. How to Assign Responsibilities between Governments (Problem 8)

When their jurisdictions overlap, governments often have conflicting interests regarding transport policy. An interesting instance of this is examined in the Edinburgh study. The proposed congestion-charging scheme featured an inner toll cordon around the city centre and an outer toll cordon encompassing a much larger area. By law, responsibility for setting the tolls, collecting the revenues and determining how they would be used had to be made by a single local authority, which was proposed to be the CEC. Since the outer cordon toll would have been paid for largely by residents outside Edinburgh, the CEC had an incentive to benefit city residents at the expense of outsiders by setting the outer cordon toll at a high level.

This possibility was formally examined in the Edinburgh study by assuming that the CEC could set tolls freely and use the revenues exclusively for the benefit of city residents. The city-optimal cordon tolls turn out to be far higher than either the proposed tolls or the welfare-maximising tolls for the region as a whole. Indeed, they are so high that they reduce welfare relative to not charging at all. As discussed in Section 2.2.2, a fear that the CEC would raise tolls above the planned level once the scheme was established may have contributed to the proposal's defeat in the referendum.

3. CONCLUSIONS

The research described in this volume originated from a growing realisation that the efficiency, equity and acceptability of transport infrastructure charging policies depend on how the revenues generated are used. Part I is devoted to developing guidelines for appropriate use of revenues with an emphasis on the underlying theory. Part II contains case studies.

This chapter has endeavoured to review and synthesise the principal findings of the case studies on the efficiency and welfare-distributional impacts of transport pricing and revenue-use schemes. By design, the studies share some common features, and the three interurban studies use the

MOLINO model. But the studies also differ in many ways: the modes of transport involved, the institutional settings, the nature of the questions that were addressed and so on. Conclusions on some of the questions also differ. These differences raise the question whether general policy insights can be gleaned from the studies that carry over to other jurisdictions, transport modes, institutional environments and so on.

There are clearly limits on the transferability of such findings as the economic merits of a particular type of investment or the degree of public support for a given policy package. Results at this level of detail are sensitive to case-study-specific factors such as congestion levels, the marginal cost of public funds, public trust in local governments and so on. The value of the case studies lies more in illustrating how pricing and revenue-use policy packages can be analysed and the range of results that are possible. They are also useful for illustrating general principles such as:

- Interdependence between pricing, earmarking and investment decisions;
- Dependence of policy impacts on characteristics of the *status quo* such as levels of fuel taxes, quality of public transport service and so on;
- Sensitivity of welfare impacts to assumptions about how the benefits from expenditures are incident across population or traveller groups (e.g. low-income v. high-income households and local freight transporters v. transit freight operators).

The remainder of this concluding section elaborates on three of the questions that were addressed in this research: the merits of earmarking, acceptability of charging and revenue use policies and assignment of responsibilities for charging and revenue allocation.

3.1. Merits of Earmarking[13]

Although widely practiced, earmarking is controversial. Various arguments for and against it were reviewed in Chapter 1 and Chapter 3 of this volume. Earmarking may entail returning the money to the facilities on which the charges are levied, or it may call for cross-subsidisation of other facilities or other modes. The case studies identify circumstances in which revenues are best allocated to particular uses, in which case earmarking the revenues for them is justified. The case studies also report survey and other evidence that earmarking enhances acceptability. Earmarking may increase efficiency too if it deters politicians from making self-interested decisions that are socially wasteful. But it can harm efficiency by preventing money from going to the most economically worthwhile uses. A clear example of this is the aborted plan for Edinburgh's cordon charge scheme to earmark all revenues to local public transport. The apparent wastefulness of this policy led to the case study's policy recommendation:

> If we were to maximise welfare from the viewpoint of the region then the optimal cordon charges produce a significant financial surplus which could not be spent in the transport sector without investing in inefficient projects. A relaxation of current legislation to permit any financial surplus to be spent in other sectors would therefore be necessary. (Laird, Nash, & Shepherd, this volume, p. 184)

As an alternative to earmarking, local decision makers could be allowed to spend revenues as they see fit as long as they can justify the expenditures on economic grounds. This is the approach taken in the 1999 Greater London Authority Act, which permits local authorities to retain the net revenues from congestion charging and workplace parking levies for 10 years provided the revenues are used for transport purposes *and that they provide value for money*.[14] Enforcing the provision "value for money" is not a trivial matter since it calls for some kind of cost–benefit analysis but it does offer a safeguard against misallocation of revenues that may occur with earmarking.

In some circumstances earmarking may channel revenues to both economically efficient and publicly acceptable uses. Yet even well-targeted earmarking schemes will be undermined if funds from other sources are reduced in an offsetting way.[15] As Richards (2005) points out, this has been a concern in Britain regarding local road user charges:

> it may well be that road user charging was seen as an opportunity, in that the charges created an additional source of revenue, and with the Treasury's effective control of local government expenditures and its ability to claw back a part of such revenues simply by reducing central funding, there was little downside risk. Indeed, in evidence to the London Assembly on Livingstone's plans to extend the central London scheme westwards, the Confederation of British Industry explained: "our disappointment is with … the extent to which parallel funding from central government appears to have been reduced, almost to the exact amount of revenue raised from the congestion charge" (London Assembly, 2003).[16] (Richards, 2005, p. 83)

Yet another potential drawback with earmarking, also identified by Richards, is that it may work against acceptability if the benefits are perceived to be too narrowly distributed:

> it can be argued that directly associating revenues with a specific project, benefiting a particular sector, rather than spreading the benefit across the community, can reduce the general acceptability of a charging scheme. (Richards, 2005, p. 229)

Implementation of London's congestion charge was greatly facilitated by the fact that a large majority of London commuters use public transport. Earmarking the revenues for improved bus services was therefore virtually guaranteed to enhance acceptability of the charge. Devising a popular earmarking scheme is likely to be much more difficult almost anywhere else – as Edinburgh's experience demonstrates.

3.2. Acceptability

There is now abundant evidence from various countries that acceptability is a *sine qua non* of transport policy reform. Acceptability appears to have been a major consideration in the design of the pricing and revenue-use policy packages (both implemented and proposed) that are examined in the case studies. Policy reforms embody numerous elements that either contribute to, or detract from, their acceptability. The brief discussion here deals with factors that relate to clarity and credibility.

3.2.1. Clarity in Objectives and Operation

To convince decision makers or the public that a charging *cum* revenue scheme is warranted, it is necessary to explain to them five things: (a) the scheme's goals and why they are worth pursuing, (b) how the scheme is designed to achieve the goals, (c) why the scheme is superior to (or complementary with) other measures, (d) the impacts of the scheme on major stakeholder groups, and (e) why the impacts are beneficial. All this is easier to do if the scheme is simple and transparent.

Harsman (2003) argues that congestion pricing has progressed further on interurban roads than in cities partly because traffic is confined to a limited number of well-defined links. This makes it easier to predict the effects of tolls during the planning stage, to measure the effects once the system is operating, to correct problems and to communicate the results to the public. In contrast, Harsman (2003, p. 143) notes, "[t]he urban transport network is considerably more complex with several partly competing, partly complementary modes of transport". Inadequate public information campaigns have contributed to a number of failed attempts at urban road pricing including electronic road pricing in Hong Kong (Borins, 1988), Rekening Rijden (Bill Riding) for the Randstad area in the Netherlands (Small & Gómez-Ibáñez, 1998, Section 10.5.1), Cambridge UK (Oldridge, 1995; Ison, 2004), several schemes in the US and Edinburgh (see Chapter 10 of this volume).

Revenue use is easier to understand if the revenues are earmarked for just a few purposes or only one – consistent with Bös's (2000) definition of earmarking. Edinburgh's proposed scheme meets this criterion reasonably well. So do the High Occupancy Toll lanes on Interstate 10 and Interstate 394 in the US for which toll revenues are earmarked to operations and bus services on the same right-of-way.[17]

Simplicity in pricing is even more important for acceptability than is simplicity in revenue use because individuals have to understand a pricing scheme in order to adapt optimally to it. It is well documented that people dislike complex tolling systems and public transport tariff structures, as well as an excessive range of options (Bonsall, Shires, Matthews, Maule, & Beale, 2004; UK Department for Transport, 2004, Appendix D; Seidel et al., 2005). And the survey of road hauliers conducted for the German study reveals opposition to altering the existing HGV charging system to a form of congestion charging or to real-time pricing.

One promising approach to enhance acceptability is to link toll levels directly to service quality in a way that is transparent to users. In Singapore, toll schedules on expressways and arterials are varied every three months to maintain speeds within specified ranges. And tolls on US HOT lane facilities are varied by time of day to maintain free-flow or near free-flow conditions on the tolled lanes. Pricing according to value of service in this way departs from second-best pricing, and therefore sacrifices some efficiency, but has advantages with respect to transparency and acceptability as well as computational ease.[18]

Charging schemes can be used for demand management, revenue generation and other objectives, and it is natural from an operations-research perspective to pursue them all. But doing so works against clarity. As Jones (2003) points out, most successful operational schemes have emphasised either congestion relief or revenue generation combined with earmarking.[19] Yet most proposed schemes promote both objectives. The objectives of road pricing in Oslo have evolved from generating revenue for new road investments (Oslo Package 1) to generating revenue for public transport investments (Oslo Package 2) to congestion charging (Oslo Package 3). However, these shifts took place gradually in response to changing circumstances, and after much debate, and do not seem to have adversely affected public acceptability.

In concluding it should be acknowledged that clarity does not invariably contribute to acceptability. Transparent schemes may be more vulnerable to opposition because the identities of the losers and/or the magnitude of their losses are evident. As an example, Richards (2005, p. 236) notes that consultation on the extension of the London congestion-charging scheme drew more than 10 times as many representations from the public as did the initial scheme. Richards attributes this to the focus of attention on a single issue and a particular area, and the ability of local councils to launch a strong (although ultimately unsuccessful) campaign against the extension.

3.2.2. Credibility

Even a policy that has clear objectives and operational characteristics is unlikely to be accepted unless it is reasonably assured that the policy will actually be carried out.[20] Stakeholders must be convinced that charges will be imposed fairly and evolve (or remain fixed as the case may be) as promised. And if revenues are earmarked, there must be assurance that moneys will be allocated as intended and without offsetting reductions from other sources (see Section 3.1 above). The Edinburgh study remarks on how a lack of legal obligation for the CEC to share revenues undermined the confidence of residents outside the city. And the decision to raise the London congestion charge from £5 to £8 soon after the scheme began operation was perceived as a change in the rules that could undermine the credibility of a future national road-pricing scheme (Richards, 2005, p. 240). It also raised the fear that a scheme purportedly designed for congestion management had a hidden agenda to raise revenue. The perception that the scheme might be a ''cash cow'' was reinforced by statistics showing that penalty charges for failing to pay the toll on time amounted to more than a third of total revenues (Richards, 2005, p. 211).

As discussed above in connection with value-of-service pricing, deviations from second-best pricing may be warranted on acceptability grounds. But significant departures may undermine the credibility of a scheme. London's congestion charge includes discounts or exemptions for various categories of vehicles and travellers that collectively account for nearly half the vehicles on the road. Edinburgh's proposed scheme featured a similar list. Efficiency does call for toll differentiation with respect to certain characteristics of vehicles (e.g. size and axle load) and trips (e.g. time of day). But granting exemptions to buy off opposition compromises congestion-relief objectives.[21] So do discounts for block payments or passes (which were a feature of Singapore's Area Licensing Scheme, the Norwegian toll rings and London' congestion charge) and ceilings on payments such as the £2 maximum daily charge in Edinburgh, the SEK 60 maximum daily charge in Stockholm's 2006 congestion-pricing experiment and other proposed toll cordon schemes.[22]

3.3. Assignment of Responsibilities between Governments and the Private Sector

National, regional and local governments all play a role in transport policy. The appropriate assignment of responsibilities for user charging and revenue use decision making depends on various considerations. One is local knowledge about congestion, the merits of alternative infrastructure investments and so on, which favours assignment of responsibility to local governments. By contrast, spill over problems between regions related to inter-regional traffic, pollution and so on, call either for centralised government control or coordination between neighbouring regional governments. The Edinburgh case study illustrates the dangers of delegating decision making to a local authority and, correspondingly, the need to develop proposals on a consensus basis between authorities.[23]

It has become popular to delegate transport infrastructure financing and operation to the private sector (see Chapter 4 of this volume). The primary motivation is to lower costs. In other cases, responsibilities have been devolved to independent bodies. The Swiss railway investment fund FINÖV and the French funding agency AFITF are two examples in the case studies. In this case the motivation usually derives from government failure. Politicians pursue hidden agendas or succumb to regulatory capture, which causes a loss of trust or credibility (see Section 3.2.2 above). Lack of trust in government is evident from the German survey of road hauliers who expressed a preference for revenue collection and allocation functions to be made either by an independent motorway operator or by separate government agencies. However, the survey also revealed a lack of support for charging and revenue collection by a private company. The case study attributes this attitude to adverse experience with TollCollect.

One potential drawback of devolving responsibility is a loss of control over integrated transport policy. Again, London provides a useful illustration. As Richards (2005, pp. 225–226, 276) explains, Transport for London was able to coordinate bus service improvements with the congestion charge because bus services remain regulated in London. Local authorities outside London lack similar control because services have been deregulated. Similarly, a case study of France for the TIPP project (Pahaut, 2005) concluded that (partial) devolution of responsibility for regional passenger rail services to regional governments resulted in a failure to control for external effects or to accommodate long-distance passenger services.

3.4. Future Prospects

Three of the case studies of this volume (Oslo, Switzerland and Germany) describe transport charging *cum* revenue-use schemes that have been implemented and appear to be meeting their goals relatively well. In the case of Oslo, passage of the legislation for Oslo Packages 1 and 2 was enabled by the ability and willingness of the major stakeholders to bargain towards a consensus. The German HGV toll was facilitated by the perception that congestion and infrastructure deficiencies were a major problem, and by a design scheme that addressed the problem and was also seen to be fair. The reasons for London's triumph with congestion charging have been extensively analysed (e.g. Richards, 2005; Litman, 2005). A few other systems have also been successfully launched around the world, but there have also been many failures. On balance, it appears that transport charging can be both efficient and politically feasible only if accompanied by a revenue-use plan and an information/marketing campaign that meet five conditions closely related to the conditions identified above in connection with acceptability:

(a) Goals that are worth pursuing;
(b) A design that can achieve the goals;
(c) Advantages over (or at least in complementarity with) other measures;
(d) A clear and credible explanation of how the scheme impacts major stakeholder groups; and
(e) A clear and credible explanation of why these impacts are beneficial.

Satisfying all these conditions is a tall order. And yet it seems important, if not essential, that the next few steps for road pricing in Europe master these hurdles if the current momentum towards efficient transport pricing and revenue use is not to stall. One reason for concern is wavering support for road pricing in Britain. Another is Richards' (2005) view that transport pricing research has not been directed in the most useful directions:

> Although much recent academic and EC-funded research has been directed towards the development of "optimal" road pricing designs, the principle of optimality was never an issue in London. There was little political debate about charging systems or the level of the charge, and comparatively little debate on the location of the cordon. Neither was there much political debate about the revenues, costs, net social costs and benefits, or overall levels of traffic reduction; the orders of magnitude forecast were considered reasonable, given the uncertainties. Indeed, it is difficult to avoid the conclusion that there is a substantial gap between issues pursued by the research community and the issue of primary interest to those at the front end of scheme delivery. (Richards, 2005, p. 225)

Richards' point is certainly well taken. It is worth noting, however, that several recent EU projects have addressed a number of practical considerations of road-pricing implementation, including barriers to user charging (AFFORD and MC-ICAM), acceptability (PRIMA and PATS) and institutional factors (TIPP). The REVENUE project took an important further step by demonstrating the importance of revenue-use and its relationship with charging, earmarking and investment decisions.

NOTES

1. Since fares and service frequencies are optimised together, it is not possible to isolate the effects of the fare changes alone.
2. This is the case even though pre-existing taxes are left unchanged rather than being adjusted optimally (i.e. following Ramsey pricing rules) to raise the extra revenues at minimum cost.
3. This is consistent with an earlier case study of Edinburgh for the AFFORD project by Shepherd (2003), who found that allocating all revenues to transport ''would clearly be impractical'' (p. 429).
4. Link et al. (2005, p. 12) argue that cross-subsidisation may be appropriate for Switzerland, but not Germany, because of the higher market share of rail transport in Switzerland. They also suggest that representation of the entire German road and rail networks as two links (in order to fit the MOLINO model) may have introduced serious aggregation bias. Aggregation bias is less likely to be a problem in the Swiss and French studies because they considered specific projects involving simple transport networks.
5. Setting low tolls and earmarking are complementary insofar as low tolls yield less revenue that can potentially be wasted on unfavourable projects (see Chapter 3 of this volume).
6. Some case studies in the PRIMA project also concluded that charges need to be introduced at low levels to gain public acceptance. Indeed, this was Singapore's strategy when it introduced electronic road pricing in 1998.
7. Consistent with this assessment, McQuaid and Grieco (2005) note that a zonal charging scheme would have been more effective than the proposed cordon scheme at reducing congestion. They follow up with the remark: ''This suggested that reducing congestion may have been a secondary motive to the pressure to gain political support for the referendum to introduce congestion pricing and that congestion pricing for Edinburgh had a high revenue raising component''. (McQuaid & Grieco 2005, p. 476).
8. There is also a charge in Durham, but it applies only to a single road leading to the city's historic centre and therefore offers no significant precedent.
9. Charges for the long-established schemes in Norway and Singapore have been raised infrequently. Frequent and substantial toll hikes have occurred on some toll roads including State Route 91 in Orange County, California and Highway 407 in Toronto. But it is doubtful that these schemes, which are outside Europe and quite different in design, have had any appreciable influence on public attitudes in Europe.

10. This last point is raised by McQuaid and Grieco (2005, p. 476).

11. This share is larger than for most other existing road-pricing schemes. As a fraction of gross revenues, collection costs for the Swiss heavy vehicle charge are about 7 per cent and for the Austrian scheme 10–15 per cent (Richards, 2005, pp. 71–72). For the Norwegian urban toll rings the figures are 9.8 per cent for Oslo, 19 per cent for Bergen and 15 per cent for Trondheim (Ramjerdi, Minken, & Ostmoe, 2004). Only in London are annual operating costs a larger fraction of toll revenues (Transport for London, 2005).

12. Truckers may also be constrained in their ability to reschedule deliveries in response to time-of-day tolls. Holguín-Veras (2005) reports persuasive evidence of such constraints for freight shippers in New Jersey and New York.

13. Some of the ideas in this section were suggested by Chris Nash.

14. As Richards (2005, p. 88) remarks, the requirement of "value for money" is not stipulated in the 2000 Transport Act for authorities outside London. The legislation for Scotland that applies to the Edinburgh scheme is different again.

15. According to Bös's (2000) definition of earmarking (see Chapter 1), dedicated revenues provide only part of the total revenues required to fund a public good. Other sources are therefore tapped as well, and funding from them can be reduced or withdrawn.

16. Richards later (p. 214) cites a claim that the UK Treasury reduced London's transport grant by more than the net proceeds from the congestion charge. Richards (p. 229) is more sanguine about innovative financing schemes in which project funds are borrowed against the stream of future charge revenues.

17. Singapore is a notable counterexample since toll revenues are not earmarked but rather absorbed into the general budget.

18. De Palma, Kilani, and Lindsey (2005) investigate value-of-service pricing in the form of third-best tolls on a simple road network.

19. An exception is the German HGV toll (Chapter 12), which has the twin goals of rationing road capacity and raising funds for investment.

20. Credibility also matters for efficiency. For example, adjustments in vehicle fuel efficiency, residential location and other long-run travel-related decisions will unfold as intended only if the decision makers involved believe that future costs will evolve as claimed (UK Department for Transport, 2004, Appendix F, F.30).

21. Discounts or exemptions for residents may be justifiable on efficiency and/or acceptability grounds in the case of crude schemes, such as cordons or area charges, for which payments do not vary smoothly with distance travelled.

22. Maximum charges are generally criticised in the theoretical literature on road pricing, but they seem to be supported by the public.

23. This point was suggested by Chris Nash.

ACKNOWLEDGMENTS

Financial support from the European Commission (EU Fifth Framework research project REVENUE (DG TREN, Contract No. GMA2-2002-52011)) is gratefully acknowledged.

REFERENCES

Bonsall, P., Shires, J., Matthews, B., Maule, J., & Beale, J. (2004). *Road user charging – pricing structures*. Draft final report for the Department for Transport, University of Leeds: The Institute for Transport Studies.

Borins, S. F. (1988). Electronic road pricing: An idea whose time may never come. *Transportation Research A*, *22A*, 37–44.

Bös, D. (2000). Earmarked taxation: Welfare versus political support. *Journal of Public Economics*, *75*(3), 439–462.

De Palma, A., Kilani, M., & Lindsey, R. (2005). A comparison of second-best and third-best tolling schemes on a road network. *Transportation Research Record*, *1932*, 89–96.

Harsman, B. (2003). Success and failure: Experiences from cities. In: J. Schade & B. Schlag (Eds), *Acceptability of transport pricing strategies* (pp. 137–151). Amsterdam: Elsevier.

Holguín-Veras, J. (2005). *The truths, the myths, and the possible in freight road pricing in congested urban areas*. Working Paper, Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute.

Ison, S. (2004). *Road user charging: Issues and policies*. Hampshire, UK: Ashgate Publishing Ltd.

Jones, P. (2003). Acceptability of road user charging: Meeting the challenge. In: J. Schade & B. Schlag (Eds), *Acceptability of transport pricing strategies* (pp. 27–62). Amsterdam: Elsevier.

Link, H., & Stewart-Ladewig, L. with contributions from partners (2005). *Deliverable 4: Report on the implementation of interurban case studies*. Project REVENUE (Revenue Use from Transport Pricing), Funded by the European Commission, Competitive and Sustainable Growth Programme, Contract: GMA2-2002-52011.

Litman, T. (2005). London congestion pricing – implications for other cities. *Dice Report: Journal of Institutional Comparisons* (http://www.cesifo-group.de/portal/page?_pageid = 36,34692&_dad = portal&_schema = PORTAL), *3*(3), 17–21.

McQuaid, R., & Grieco, M. (2005). Edinburgh and the politics of congestion charging: Negotiating road user charging with affected publics. *Transport Policy*, *12*, 475–476.

Oldridge, B. (1995). Congestion metering in Cambridge city, United Kingdom. In: B. Johansson & L-G. Mattsson (Eds), *Road pricing: Theory, empirical assessment and policy* (pp. 131–140). Boston: Kluwer.

Pahaut, S. (2005). *Intermodal competition in French intercity passenger markets*. Project TIPP (Transport Institutions in the Policy Process: http://www.strafica.fi/tipp), Funded by the European Commission, Contract No: GMA2/2001/52043 S07.15419 TIPP.

Ramjerdi, F., Minken, H., & Ostmoe, K. (2004). Norwegian urban tolls. In: G. Santos (Ed.), *Road pricing: Theory and evidence. Research in transportation economics* (9, pp. 237–249). Elsevier Science.

Richards, M. G. (2005). *Congestion charging in London: The policy and the politics*. Houndsmill: Palgrave Macmillan.

Seidel, T., Matthes, A., Wieland, B., Schlag, B., Schade, J., Verhoef, E. T., Ubbels, B., Tánczos, K., Kosztyó, A., & Mészáros, F. (2005). *Deliverable 4: Political acceptability and perceived legitimacy of transport policy implementati*on. Project TIPP (Transport Institutions in the Policy Process: http://www.strafica.fi/tipp), Funded by the European

Commission, Competitive and Sustainable Growth Programme, Contract No: GMA2/2001/52043 S07.15419 TIPP.

Shepherd, S. P. (2003). Towards marginal cost pricing: A comparison of alternative pricing systems. *Transportation*, *30*(4), 411–433.

Small, K. A., & Gómez-Ibáñez, J. A. (1998). Road pricing for congestion management: The transition from theory to policy. In: K. J. Button & E. T. Verhoef (Eds), *Road pricing, traffic congestion and the environment: Issues of efficiency and social feasibility* (pp. 213–246). Cheltenham, UK: Edward Elgar.

Transport for London (2005). Congestion charging: Third annual monitoring report. April, http://www.tfl.gov.uk/tfl/cclondon/cc_monitoring.shtml, accessed July 22, 2005.

UK Department for Transport (2004). Feasibility study of road pricing in the UK – report. http://www.dft.gov.uk/stellent/groups/dft_roads/documents/divisionhomepage/029709.hcsp, accessed July 28, 2004.

This page intentionally left blank

SUBJECT INDEX