VATT-TUTKIMUKSIA
63
VATT-RESEARCH REPORTS

David Milne
Esko Niskanen
Erik Verhoef

OPERATIONALISATION OF MARGINAL COST PRICING WITHIN URBAN TRANSPORT

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Helsinki 2000

ISBN 951-561-310-8
ISSN 0788-5008

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Hämeentie 3, 00530 Helsinki, Finland

Email: dmilne@its.leeds.ac.uk
esko.niskanen@vatt.fi
everhoef@econ.vu.nl

J-Paino Oy
Helsinki, 2000

Operationalisation of Marginal Cost Pricing within Urban Transport

AFFORD

Deliverable 1
January 2000

David Milne[*]
Esko Niskanen[**]
Erik Verhoef[***]

Project No: PL97-2258
Project Coordinator: VATT

Funded by the European Commission
4th Framework - Transport RTD

---

[*] Institute for Transport Studies (ITS), University of Leeds
[**] Government Institute for Economic Research (VATT)
[***] Free University of Amsterdam (FUA)

The AFFORD Consortium

Government Institute for Economic Research (VATT), Coordinator, *Finland*
University of Leeds (ITS), *United Kingdom*
Universidad Politecnica de Madrid (UPM), *Spain*
Institute of Transport Economics (TOI), *Norway*
Technische Universität Dresden (TUD), *Germany*
TRIAS SA Consulting (TRIAS), *Greece*
MIP - Politecnico di Milano (MIP), *Italy*
University of York (UYORK), *United Kingdom*
LT Consultants Ltd (LTCON), *Finland*
Center for Interdisciplinary Systems Research (C.I.S.R), *Greece*
Free University of Amsterdam (FUA), *The Netherlands*

MILNE, DAVID - NISKANEN, ESKO - VERHOEF, ERIK: OPERATIONALISATION OF MARGINAL COST PRICING WITHIN URBAN TRANSPORT. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, Government Institute for Economic Research, 2000, (B, ISSN 0788-5008, No 63). ISBN 951-561-310-8.

Abstract: Charges and taxes for transport have traditionally had little connection to costs, instead being part of broader fiscal policies of raising revenue or directly promoting other goals, industrial, social and environmental. The gap between the costs and actual charges is particularly evident in urban road transport where current pricing mechanisms typically make little or no attempt to reflect concentrations of transport activity in time and space and hence of transport induced costs. Economic theory shows that, under the market approach, marginal cost pricing is a condition for economic efficiency. Still, a huge gap exists between the lessons of economic theory and the possibilities of current technology on one hand, and the achievements in implementing marginal cost pricing thus far in practice on the other. In relation to the broader socio-economic context of marginal social cost pricing, determined by various technological, institutional, legal and political constraints, this report highlights three important aspects or distinctions: the distinction between policy situations with different coverage, the distinction between first-best and second-best situations, and the need for policy packaging.

Key words: urban transport, marginal cost pricing, first-best, second-best

MILNE, DAVID - NISKANEN, ESKO - VERHOEF, ERIK: OPERATIONALISATION OF MARGINAL COST PRICING WITHIN URBAN TRANSPORT. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, Government Institute for Economic Research, 2000, (B, ISSN 0788-5008, No 63). ISBN 951-561-310-8.

Tiivistelmä: Liikenteen hinnoilla ja veroilla on perinteisesti ollut vain vähän yhteyttä kustannuksiin. Ne ovat ennemminkin määräytyneet osana laajempaa vero- ja finanssipolitiikkaa tai niihin ovat suoraan vaikuttaneet muut tavoitteet kuten teollisuus-, sosiaali- ja ympäristötavoitteet. Kustannusten ja todellisten hintojen välinen kuilu on erityisen ilmeinen kaupunkien henkilöautoliikenteessä, missä nykyiset hinnoittelukäytännöt tyypillisesti vain vähän jos ollenkaan pyrkivät heijastamaan liikennesuoritteiden jakautumista yli ajan ja paikan ja täten liikenteen aiheuttamia kustannuksia. Talousteorian mukaan rajakustannushinnoittelu on markkinoihin perustuvassa ratkaisussa taloudelliselle tehokkuudelle välttämätön ehto. Suuri kuilu on kuitenkin yhä olemassa yhtäältä talousteorian johtopäätösten ja nykytekniikan mahdollisuuksien ja toisaalta rajakustannushinnoittelun käytännön soveltamisen välillä. Rajakustannushinnoittelua on tarkasteltava laajemmassa yhteiskuntataloudellisessa yhteydessä, jossa otetaan huomioon erilaisten teknologisten, institutionaalisten, lakiin perustuvien ja poliittisten rajoitteiden olemassaolo. Tässä tutkimuksessa korostetaan kolmea tärkeätä näkökohtaa: erilaisten politiikkatilanteiden ja päätöksentekotasojen huomioon ottaminen, first-best ja second-best -tilanteiden erottelu sekä tarve kehittää toimenpidepaketteja.

Asiasanat: kaupunkiliikenne, rajakustannushinnoittelu, first-best, second-best

Acknowledgements

This report has benefited from the inputs of the whole AFFORD consortium. The following persons have provided background texts: Heide Coenen (sections 3 and 6), Lasse Fridstrom (5), Mustapha Ghali (5, 6 and 7), Anton Goebel (2), Paavo Moilanen (5, 6 and 7), Simon Shepherd (5, 6 and 7) and Michael John Smith (5, 6 and 7). Michael John Smith was also a co-author of the first version (version 1.1 in December 1998) of this report. Astrid Beier, Ionna Giannouli, Ferdinando Stanta and Arild Vold have provided background reports regarding their country specific situation. Anita Niskanen has provided technical assistance. In particular, we want to thank Catharina Sikow-Magny and Amihai Glazer for helpful comments. Many other persons have contributed through their comments and suggestions at various AFFORD workshops. However, the authors alone are responsible for the views presented in the report.

Executive summary

Introduction

1. Charges and taxes for transport have traditionally had little connection to costs, instead being part of broader fiscal policies of raising revenue or directly promoting other goals, industrial, social and environmental. The gap between the costs and actual charges is particularly evident in urban road transport where current pricing mechanisms typically make little or no attempt to reflect concentrations of transport activity in time and space and hence of transport induced costs.

2. Economic theory shows that, under the market approach, marginal cost pricing is a condition for economic efficiency. Still, a huge gap exists between the lessons of economic theory and the possibilities of current technology on one hand, and the achievements in implementing marginal cost pricing thus far in practice on the other. There is an obvious tension or paradox between the economic theory, which suggests marginal cost pricing is the right solution, and practical experience, which suggests that such pricing measures are hard to implement. In particular, road pricing appears to enjoy notably scant support among the population in general and is a controversial topic among politicians and professional policymakers.

3. The reasons for this situation can be manifold:

(i) the gap between the underlying economic theory of marginal cost pricing and the requirements for practical implementation may be too large;
(ii) important institutional and legal barriers to implementing marginal cost pricing may exist;
(iii) concerns of the distributional effects of marginal cost pricing may overshadow considerations of the positive efficiency gains; and
(iv) public and political acceptance of marginal cost pricing may be low because of various non-economic factors.

Item (i) can be said to refer to problems of *operationalisation* of marginal cost pricing. Items (ii), (iii) and (iv) can be called *implementation* problems.

4. This deliverable will address issues of *operationalisation*. Implementation problems will be addressed in Deliverable 2 of the AFFORD project. The term *operationalisation* means taking actions with the aim to narrow a gap between the theory and the practice of policy making. Such a gap may exist in urban transport because:

(i) the relevant marginal cost concepts and items, including the correct level of (dis)aggregation, may not be identified carefully enough and estimated with sufficient accuracy;

(ii) the economic merits of marginal cost pricing are typically derived and illustrated in highly simplified graphical and analytical settings, ignoring many complexities surrounding urban transport pricing in reality, thus casting doubt on the transferability of insights and results to practical cases;
(iii) the institutional context in which marginal cost pricing would be implemented may not be fully accounted for in the policy recommendations thus far; and
(iv) the two major strands of research in the field, conceptual economic approaches and real-world modelling approaches, have previously not been sufficiently integrated.

This study will address these sets of issues. However, because type (i) issues have already received quite a lot attention in other recent projects, this report focuses on type (ii), (iii) and (iv) issues.

Marginal cost concepts and categories

5. A precondition for marginal cost pricing is that the relevant marginal cost items have been identified and defined and their magnitudes estimated. A natural starting point for the categorisation of different cost types in urban transport is to identify the different vested interests and affected parties (i.e. those actors who generate and/or incur the costs). They are:

(i) private infrastructure users;
(ii) transport service operators;
(iii) infrastructure operators;
(iv) government; and
(v) third parties.

6. Marginal cost in transport can in principle refer to:

(i) the cost of an additional user of transport infrastructure;
(ii) the cost of providing an additional unit of transport service (e.g. freight haulage or public transport trips); or
(iii) the cost of providing an additional unit of infrastructure capacity.

This study will focus on marginal cost pricing of the use of transport infrastructure. Therefore, the central issue will be the marginal cost of using the infrastructure (case i). It is the inefficiencies related to the use of infrastructure that most urgently call for application of marginal cost pricing in urban transport.

7. Restricting to the costs related to the use of transport infrastructure, an important distinction views the relevant actors (i.e. private infrastructure users and transport service operators) as generating two types of costs:

(i) internal costs; and
(ii) external costs.

Internal costs refer to costs which are accounted for and included in transport-related decision making, either through monetary resource costs or the time taken during travel. External costs refer to costs which are not included in such decision making (i.e. they are external to the system). Within both broad categories of costs, marginal costs and total costs can be considered separately.

8. All internal costs are not necessarily confronted by the users of transport infrastructure when making transport-related decisions. That is, they are not necessarily internal to the primarily benefiting actor, in the sense that (s)he will cover or pay for them in full (in accordance to the user-pays principle). For example, parking is often considered as causing internal resource costs which are not covered in the sense of this principle. Either the infrastructure operator providing the parking facilities incurs these uncovered costs, or parking is subsidised from outside the transport market by employers or by the government.

9. The major external cost categories in the context of urban transport are:

(i) congestion costs;
(ii) infrastructure damage;
(iii) external accident costs;
(iv) noise;
(v) visual intrusion and barrier effects;
(vi) local emissions; and
(vii) global emissions.

Items (i) and (ii) are intra-sectoral externalities, items (v), (vi) and (vii) are inter-sectoral externalities. Item (iii) contains both elements. Notice that infrastructure damage (item ii) can (in principle) be well defined for property rights, but is not an externality in the same strict sense as are the other items. However, the implications for marginal cost pricing to be addressed in this study are the same.

10. A prerequisite for marginal cost pricing is the ability to make estimates of relevant internal and external marginal costs. A theoretically sound identification and estimation of different marginal costs can only be based on deep understanding of the mechanisms behind the generation of such costs. This, in turn, requires identification of the different types of activities in which the users of transport infrastructure are involved. In this study, these activities are called *dimensions of behaviour*.

Marginal cost pricing: basic categorisations

11. Transport infrastructure users are involved in a wide range of different activities or dimensions of behaviour (cf. 10). Correspondingly, they are also involved in a wide range of decision-making situations: from short-run to long-run decisions, and from strictly transport-related decisions to those which relate to and affect their transport behaviour only indirectly. Understanding how marginal cost pricing and other related pricing types affect these decisions is critical to successful implementation.

12. Marginal cost pricing and other related pricing in urban transport can, in principle, be applied by:

(i) transport service operators;
(ii) infrastructure operators; and
(iii) the government/regulator.

The transport service and infrastructure operators are typically commercial firms or public agencies.

13. The purpose of pricing, when applied by transport service operators and infrastructure operators, is to collect revenue. Besides marginal cost pricing, such operators can also adopt a number of other pricing approaches. The following is a rough categorisation of possibilities:

(i) marginal cost pricing;
(ii) average cost pricing;
(iii) monopoly pricing;
(iv) two-part tariffs; and
(v) Ramsey pricing.

Items (iv) and (v) extend the basic idea of marginal cost pricing. Sometimes, when appropriate to make a distinction, item (i) is called *pure marginal cost pricing.*

14. When applied by the government/regulator, the goal of marginal cost pricing is to internalise transport related external costs, both intra- and inter-sectoral, and, thus, to secure that each activity by each user will be extended to the point where the social benefit of the last unit equals the social cost. Otherwise, in urban transport, since marginal private costs and benefits can differ significantly from the corresponding social costs and benefits, the resulting market allocation would typically be inefficient. Marginal cost pricing which aims to correct distortions due to discrepancies in marginal private and social costs is called *marginal social cost pricing* or *marginal externality cost pricing.*

Marginal social cost pricing

15. A useful categorisation of the practical pricing measures for implementing marginal social cost pricing is the following:

(i) charges for using particular stretches of road;
(ii) charges for parking;
(iii) public transport fares and subsidies;
(iv) taxes upon the purchase and licensing of vehicles and upon associated commercial services; and
(v) taxes on the purchase of fuel.

The ordering here is from microscopic measures (item i) to macroscopic measures (item v), the latter having obvious direct implications to the other sectors of the economy.

16. Besides a means of marginal social cost pricing with the aim of internalising external costs, pricing measures (i)-(v) can also be used to promote other economic and social goals. In particular, measures (ii) and (iii) can be used for charging the relevant internal resource costs. Indeed, this is their primary task or purpose. Measures (iv) and (v), in turn, are also typically used for fiscal taxation and thus, via the use of the revenues, for promoting wider social goals.

17. Considering marginal social cost pricing, an important question concerns its coverage: what are the relevant behavioural dimensions to be included? A related important question is the relevant level of (dis)aggregation. In this study, in order to secure a full and balanced coverage of these issues, the following *settings* are distinguished:

(i) focusing on road transport;
(ii) covering multimodal transport;
(iii) covering interactions with inter-urban transport; and
(iv) covering interactions with land use.

The motivation for a separate treatment of settings (i) and (ii) is obvious. A separate consideration of settings (iii) and (iv) reflects the high relevance of interactions of urban transport with other spatio-economic structures.

18. Another important distinction relevant to marginal social cost pricing concerns the extent and detail in which different technological, institutional, legal and political constraints be allowed. A *first-best optimum* refers to a full, unconstrained social optimum, while a *second-best optimum* is defined as a constrained optimum, taking account of the range of practical complications. Correspondingly, as representations of marginal social cost pricing, economists speak of *first-best pricing* and *second-best pricing* as the charges required to bring about an optimum in each case.

Modelling approaches

19. A wide range of modelling approaches is available in order to represent marginal external costs and marginal social/external cost pricing within urban transport and to analyse its potential effects. An important distinction exists between *conceptual modelling approaches*, mostly from an economic background, and *real-world model applications*, which incorporate economic principles alongside other social, mathematical and engineering concepts.

20. A basic classification of conceptual model approaches is the following:

(i) the conventional static single-link model of congestion;
(ii) increased modelling detail related to the introduction of dynamics;
(iii) the introduction of spatial networks; and
(iv) the representation of randomness and uncertainty.

Characteristics to the conceptual models is that they each focus on single specific aspects of the urban transport problem and aim to use analytically solvable and, thus, highly stylistic models.

21. The real-world model applications are simulation models which enable planners and decision makers to test the impacts of changes to transportation systems against a variety of criteria, often within a single software package. In many cases, these model applications enable a wide range of possible schemes to be assessed within the same modelling framework. An ideal real-world model application would address the full range of transport and related land use policy issues at all levels of detail. At present, no such application exists due to limitations of behavioural research, data availability, modelling techniques, computing power and cost. Current applications rely on trade-offs, focusing resources on particular chosen issues and levels of detail at the expense of others.

22. Broadly speaking, four categories of currently available real-world model applications may be identified:

(i) detailed simulation models;
(ii) tactical network models;
(iii) strategic transport models; and
(iv) geographic models.

Here the model types are arranged from microscopic (item i) to macroscopic (item iv) models. An individual real-world model application may not fit wholly and exclusively within any one of these categories and may involve interactions between models with quite different features within a common framework. Examples from each of these categories are available for use within the AFFORD project.

23. An important observation is that the real-world model categories (i)-(iv) are distinguished along the same lines as the settings (i)-(iv) defined earlier (cf. 17). Detailed simulation and tactical models are typically used to analyse road transport, strategic transport models usually focus on multimodal transport and sometimes cover interactions with inter-urban transport and geographic models usually cover the inter-urban setting and, by definition, address interactions with land use. Evidently, the policy relevance of these settings has, to some extent, affected the development of real-world model applications to coincide.

24. The integration of conceptual economic analyses and real-world model applications has not previously received the attention it deserves. In particular, greater interaction of the two approaches is needed to render the real-world model applications to better allow for economic efficiency as a goal, alongside the traditionally well represented broader objectives. This is a precondition for the use of these real-world models for analysing marginal social cost pricing.

First-best pricing

25. Marginal social cost pricing considers economic *efficiency* as the sole criterion of performance. A widely accepted way to represent economic efficiency as a welfare criterion in conceptual economic model approaches is through the idea of the net economic or social benefit. This criterion ultimately reflects the consumers' willingness to pay and is defined as the sum of consumers' and producers' surpluses. However, this efficiency concept is based on the standard static demand-supply framework and, thus, represents short-run or static efficiency only.

26. Whether also long-run efficiency is true has to be verified separately. This requires that the implications of various factors behind the supply and demand curves of the relevant actors in the transport market be investigated, in addition to the long-run effects of possible adjustments in such factors. It turns out that, under first-best conditions, short-run marginal cost pricing provides optimal incentives for long-run optimality too. That is, first-best prices not only optimise transport *given* the shape and position of the demand and supply curves, but also create optimal incentives to change those aspects of behaviour that affect the actual shape and position of these curves in the long-run.

27. Such incentives may be related to:

(i) individuals' long-run decisions behind their spatial behaviour, determining the shape and position of the demand curve;
(ii) technology choice decisions, determining the marginal external environmental cost and marginal private cost curves; and
(iii) optimal investments in road infrastructure under conditions of congestion.

In cases (i) and (ii), the relevant demand and supply (cost) curves are affected through induced adjustments in the behaviour of the users of infrastructure. In case (iii), the long-run adjustment occurs through the behaviour of the infrastructure operator.

28. An important property of conceptual economic model analyses is that the objective function and behavioural assumptions within the model are consistent. This is due to the fact that the efficiency criterion reflecting individuals' behaviour has such a central role. It is self-evident that net economic or social benefit as the social welfare criterion (cf. 25) reflects the behavioural decisions governed by the demand and supply functions.

29. In most existing real-world model applications, in contrast, the economic efficiency criterion has, traditionally, not been given such a major role. The criteria used for evaluation typically reflect the full range of practical objectives faced by planners and politicians within the setting covered by the model. While this may include economic efficiency, it is also likely to focus on more pragmatic social, political and engineering goals. Also, the use of real-world models to justify public funding for large scale investment projects often has the implication

that economic evaluation is carried out externally, according to a government approved framework. As a result, the objective functions and behavioural assumptions within the model may not be as closely connected as in the conceptual modelling case and the link between them may be implicit.

30. Any consideration of marginal social cost pricing within real-world model applications has to presuppose an objective function which is consistent with and which reflects the notion of economic efficiency. This is required because the very concept of marginal cost pricing is based on it. In general, all real-world models have an underlying economic basis which conforms to this. However, if the existing structure of a model's (implicit or explicit) objective function causes other goals to dominate and economic efficiency has only a minor role, then it is not suited to analysing marginal cost pricing issues in its prevailing form.

31. While various other aspects of pricing policy have been investigated widely within the existing real-world model applications in urban transport, rather less has been said specifically about marginal cost pricing issues. However, notable exceptions exist (e.g. the EC DGVII OPTIMA and TRENEN projects).

32. The provision of a *benchmark* is the ultimate purpose of the notion of first-best. A first-best optimum provides an idealised benchmark only: the assumptions behind it are rather unrealistic and, typically, cannot be satisfied in real life. However, in providing such a reference state, a first-best optimum enables the analyst and the policy maker to place practical solutions and their internal differences into perspective in relation to an idealistic optimum and, thus, to better evaluate the realistic options.

33. The determination of a first-best benchmark is extremely straightforward in conceptual economic analysis. The simplicity derives from the fact that the conceptual models and their assumptions, as well as the social welfare criterion, are typically highly stylised and unambiguously defined. In contrast, in practical real-world model applications, the analyst typically faces a great number of open choices. Which institutional constraints should be included and which not? Which variables should be optimised, and which should be treated as given? And so on. Making such choices and, hence, deciding on a first-best benchmark is to a large extent a matter of judgement. The first-best optimum should represent an ideal situation which, however, should not be too unrealistic. Indeed, it would make little sense to compare potential practical outcomes to something that everyone knows can never ever be achieved.

34. A closely related question concerns the choice of the appropriate level of detail in the characterisation of a first-best benchmark. An important feature of first-best analysis should be the power of the pricing measures to differentiate between different links and different types of users of the transport infrastructure, as well as between the externalities they are generating. However, the operational definition of first-best can still be very different depending on the setting and real-world model application adopted. The definition of first-best in this respect

has to be consistent with the level of (dis)aggregation adopted in the analysis at large.

Second-best pricing

35. The distinction between the notions of first-best and second-best is a central feature in much of the literature on conceptual economic models. The literature dealing with second-best issues in terms of conceptual economic models is large. The wide range of second-best situations addressed reflects the variety of complications and imperfections brought by different technological, institutional, legal and political constraints affecting urban transportation that must be taken into account in practical policy making.

36. The *second-best pricing rules* have received ample attention in the literature on transport pricing and on road pricing in particular. A joint conclusion from various studies is that, in second-best situations, economic efficiency requires pricing instruments to be applied according to different and more complicated rules than those that apply for the first-best policy.

37. In this study, the following five major types of second-best situations are considered:

(i) insufficient power of pricing measures to differentiate;
(ii) distortions in other routes;
(iii) distortions in other modes;
(iv) distortions in other sectors; and
(v) shadow price of public funds.

38. The reasons for second-best in cases (iv) and (v), of course, are beyond the scope of transportation policy. Therefore, typically, when comparing first-best and second-best solutions on partial equilibrium approaches, focusing on the transport sector (rather than being general equilibrium and covering all sectors), they are accepted as *facts of life*. This reflects the degree of freedom that the analyst has and the need for judgement that the (s)he has to exercise when defining a first-best benchmark (cf. 33 and 34).

39. A general result applying to all cases (i)-(v) (in 37) is that, while first-best prices fully internalise marginal external costs and are equal to *direct* marginal externality cost, the second-best optimal pricing rules include correction terms reflecting the relevant and explicitly assumed second-best distortions.

40. Within real-world applications, in contrast, the distinction between the first-best and second-best has received much less attention thus far. Also, separate treatment of different second-best situations like (i)-(v) above has often not been carried out explicitly. Therefore, very little has been said about second-best pricing issues either.

41. However, the existing real-world model applications have a great potential for representing different second-best cases and for contrasting the second-best

optima with the corresponding relevant first-best benchmark optima. Given the wide range of different second-best situations and the wide range of differences in coverage and detail of the different real-world model applications (cf. 22), it is obvious that different applications are best suited to analysing different situations.

Marginal cost based policy packaging

42. The task of marginal social cost pricing, in providing optimal incentives to transport infrastructure users to change their behaviour, can be extremely complex. This is due to the great number of behavioural dimensions and categories of external costs to be accounted for. A further complication is that the different behavioural dimensions can simultaneously affect several cost categories. Rather than considering overall solutions, the literature typically has addressed individual measures in taking care of isolated problems. This reflects the general feature of conceptual economic model approaches in focusing on individual issues at a time.

43. The AFFORD study is investigating the concept of *policy packaging*, that is the joint use of different practical measures to achieve the marginal social cost pricing goal. An important feature of the policy packages considered is that they are specifically *marginal cost based*, so that first-best and second-best pricing rules are at their core, rather than the much broader range of issues which have often been used as a justification for policy packages in practice.

44. If first-best pricing measures were possible, the various behavioural dimensions determining marginal external costs in urban transport would, by definition, be affected optimally. However, in real life, the available pricing measures typically have only limited power to differentiate (cf. 37) between different types of infrastructure users and the different categories of external costs they are generating. Therefore, such second-best measures can only reproduce partially the behavioural responses that first-best pricing would have. In the construction of marginal cost based policy packages, the full set of first-best incentives, with respect to all behavioural dimensions, should ideally be covered.

45. A useful framework for the design of practical policy packages contains a categorisation of different package types. These should be sufficiently comprehensive so that they together reflect the main aspects of pricing in urban transport. Typically, the packages to be considered should include:

(i) a reference or scenario package;
(ii) a first-best benchmark package;
(iii) a best-practice second-best package; and
(iv) a package reflecting acceptability concerns.

There may be more than one relevant package in each category. Examples of this approach will be considered in Deliverable 2 of the AFFORD project.

Theory vs. current approaches to pricing in urban transport

46. While the general goal of promoting efficiency is the sole criterion behind the notion of marginal social cost pricing, politicians and planners of practical pricing systems are constrained by the need to take account of a broad range of other objectives (such as equity, revenue needs, sustainability etc.), the prevailing technological mechanisms available for levying charges and the human issues relating to whether users perceive the system and react to it as intended. As a result, somewhat paradoxically, in those few cases where road pricing measures have been implemented in practice, the objectives which have motivated the move have not always been in line with the efficiency objectives.

47. On the other hand, concentration on practical policy objectives at the expense of an explicit recognition of economic efficiency and equity does not necessarily mean that there should be a major contradiction between theory and practice. Often the practical objectives used simply provide more pragmatic goals which fall within the efficiency and equity objectives described in the theory. Better integration of economic theory, real-world modelling and practical transport planning has the potential to provide more consistent and transparent connections in this respect in the future.

Concluding comments

48. In relation to the broader socio-economic context of marginal social cost pricing, determined by various technological, institutional, legal and political constraints, this report has highlighted three important aspects or distinctions: the distinction between policy situations with different coverage (referred to as *settings*), the distinction between first-best and second-best situations and the need for policy packaging (rather than considering individual measures at a time).

49. All these aspects are extremely important for conceptual clarity and theoretical structure of the analysis. In particular, actual pricing rules suggested by the marginal social cost pricing principle can be very different depending on circumstances, for example on the level of coverage and (dis)aggregation chosen and on the explicit inclusion of various constraints and policy instruments in the analysis.

50. This report has focused on operationalisation problems of marginal cost pricing. Deliverables 2a, 2b and 2c of the AFFORD project will investigate the implementation issues. This means analysis of institutional and legal barriers to marginal cost pricing, of the actual magnitudes of the efficiency and equity impacts that the policy might have (and any trade-off between them) and of the acceptability issues.

Contents

Executive Summary

1 Introduction

Cities and urban areas in Europe (and beyond) suffer from heavy congestion and environmental degradation as a result of prevailing levels of car use. Without remedial action, this situation is expected to continue and additional problems to arise in the future, as reported in the Commission's Green Paper "*Towards Fair and Efficient Pricing in Transport*" (1995) and White Paper "*Fair Payment for Infrastructure Use*" (1998) and in many other reports. One important reason for this is that the economic costs on which many transport decisions are based are currently inefficient, because private costs (those paid by the user) differ from social costs (those borne by society as a result of the travel decision). This situation occurs both within and between the different travel modes and, therefore, affects private cars, public transport and freight. From the social point of view, there is an obvious need to introduce corrective policy measures to account for this discrepancy.

Charges and taxes for transport have traditionally had little connection to costs, instead being part of broader fiscal policies of raising revenue or directly promoting other goals, industrial, social and environmental. The gap between the costs and actual charges is particularly evident in urban road transport where current pricing mechanisms typically make little or no attempt to reflect concentrations of transport activity in time and space and hence of transport induced costs. This has led to inefficient use of the existing transport system and, in the longer run, to inefficient provision of infrastructure capacity, inefficient location of activities and distorted land use development. In contrast, a general view is that in inter-urban transport road covers its cots, albeit, not always in a behaviourally ideal way.

Economic theory shows that, under the market approach, marginal cost pricing is a condition for economic efficiency. The principle of marginal cost pricing derives from the general notion of social optimisation: at the social optimum the social benefit of the last or marginal unit of any product or service must equal the social cost of producing that unit. The theory argues that the economic welfare of society will be maximised if economic agents, i.e. consumers and producers, are faced with prices reflecting the true social costs of their activities. The principles of marginal cost pricing have been studied in the economic literature in relation to different kinds of services, both public and private.

The principle of marginal cost pricing and the practical measures through which it may be implemented have already attracted considerable interest in policy discussions within the transportation sector. The principle is well known and widely accepted as a general rule. However, in many instances its practical implementation is far from obvious. This is especially true in urban transportation, where the textbook conditions for competitive private goods markets are seldom if ever satisfied. In this situation, transport pricing in practice is nearly always based on average costs rather than marginal costs. As a result, prices are typically in some cases too high and in others too low.

Technological research has already resolved many of the practical requirements for the introduction of flexible marginal cost pricing schemes. Still, a huge gap exists between the lessons of economic theory and the possibilities of current technology on one hand, and the achievements in implementing marginal cost pricing thus far in practice on the other. There is an obvious tension or paradox between the economic theory, which suggests marginal cost pricing is the right solution, and practical experience, which suggests that such pricing measures are hard to implement. In particular, road pricing appears to enjoy notably scant support among the population in general and is a controversial topic among politicians and professional policymakers.

The reasons for this situation can be manifold:

(i) the gap between economic theory behind the general idea of marginal cost pricing and practice may be too big;
(ii) important institutional and legal barriers to implementing marginal cost pricing may exist;
(iii) concerns of the distributional effects of marginal cost pricing may overshadow considerations of the positive efficiency gains; and
(iv) public and political acceptance of marginal cost pricing may be low because of various noneconomic factors.

Item (i) can be said to refer to problems of *operationalisation* of marginal cost pricing. Items (ii), (iii) and (iv) can be called *implementation* problems.

This deliverable will address issues of operationalisation. (Implementation problems will be addressed in Deliverables 2a, 2b and 2c of the AFFORD project.) The term operationalisation means taking actions with the aim to narrow a gap between the theory and the practice of policy making. Such a gap may exist in urban transport because:

(i) the relevant marginal cost concepts and items, including the correct level of (dis)aggregation, may not be identified carefully enough and estimated with sufficient accuracy;
(ii) the economic merits of marginal cost pricing are typically derived and illustrated in highly simplified graphical and analytical settings, ignoring many complexities surrounding urban transport pricing in reality, thus casting doubt on the transferability of insights and results to practical cases;
(iii) the institutional context in which marginal cost pricing would be implemented may not be fully accounted for in the policy recommendations thus far; and
(iv) the two major strands of research in the field, conceptual economic approaches and real-world modelling approaches, have previously not been sufficiently integrated.

This study will address these sets of issues. However, because type (i) issues have already received quite a lot attention in other recent projects, this report focuses on type (ii), (iii) and (iv) issues.

The structure of this report is the following:

(i) sections 2, 3 and 4 review the relevant cost concepts and pricing approaches, and lay down basic distinctions in relation to marginal cost pricing;
(ii) sections 5, 6 and 7 present an overview of available modelling approaches, and consider the relevant marginal cost pricing situations within urban transport in terms of these models;
(iii) section 8 shows directions for the design of effective marginal cost based policy packages for application in practice in urban transport; and finally
(iv) sections 9 and 10 evaluate the current approaches to pricing in practice in the light of the theory and provide concluding comments.

2 Marginal cost concepts and categories

A precondition for marginal cost pricing is that the relevant marginal cost items have been identified and defined and their magnitudes estimated. Since there is not much use in considering marginal cost pricing in isolation (i.e. from other pricing types and from a general transport policy context) also marginal costs must be considered in the wider context of other cost types. Similarly, categorisation of different cost types only makes sense when at the same time identifying the vested and affected parties, i.e. those actors who are generating and/or incurring the costs.

A careful description of the basic setup and categorisations is needed. A natural next step is the description of the mechanisms between the different marginal cost categories and their causes. This requires identification of the relevant dimensions of behaviour.

Section 2.1 reviews the relevant parties i.e. those actors who generate and/or incur the costs. Section 2.2 discusses the important distinction between internal vs. external costs. Section 2.3 reviews the different types of external costs. Section 2.4 illustrates the insights behind introducing the notion of behavioural dimensions.

2.1 Relevant parties

A natural starting point for the categorisation of different cost types in urban transport is to identify the different vested and affected parties:

(i) private infrastructure users;
(ii) transport service operators;
(iii) infrastructure operators;
(iv) government; and
(v) "third parties".

Items (i)-(iii) refer to actors operating within the transport markets; (iv) and (v) are outside actors. The term *private infrastructure users* refers to consumers travelling for their own direct benefit. *Transport service operators* are profit-maximising commercial firms and non-profit public agencies operating within the urban transport market and producing transportation services, i.e. freight haulage and public transport trips. *Infrastructure operators* providing infrastructure can in principle be profit-maximising commercial firms and non-profit public agencies or the government itself directly. The *government* often subsidises transportation and collects fiscal taxes on it; it also acts as a regulator and imposes regulatory taxes or alternative non-price measures. Finally, the term *"third parties"* refers to economic actors, i.e. consumers and firms in other sectors, who are affected by transport-related decisions but are not directly involved in them.

Marginal cost in transport can in principle refer to:

(i) the cost of an additional use of transport infrastructure;
(ii) the cost of providing an additional unit of transport service (e.g. freight haulage or public transport trips); or
(iii) the cost of providing an additional unit of infrastructure capacity.

This study will focus on marginal cost pricing of the use of transport infrastructure. Therefore, the central issue will be the marginal cost of using the infrastructure (case i). It is the inefficiencies related to the use of infrastructure that most urgently call for application of marginal cost pricing in urban transport. However, the other two cases (ii and iii) will be covered too, where important interrelationships with the marginal cost of urban infrastructure use exist.

Evidently, sufficiently comprehensive, logical and theoretically sound further categorisations of the different costs generated in transport have to build on the basic distinctions above. This is particularly true when the ultimate goal of the categorisation is to provide information for the purposes of marginal cost pricing.

2.2 Internal costs vs. external costs

Restricting to the costs related to the use of transport infrastructure, an important distinction views the relevant actors (i.e. private infrastructure users and transport service operators) as generating two types of costs:

(i) internal costs; and
(ii) external costs.

Internal costs refer to costs which are accounted and included in transport-related decision making, either through monetary resource costs or the time taken during travel. *External costs* refer to costs which are not included in such decision making (i.e. they are external to the system). Within both broad categories of costs, marginal costs and total costs can be considered separately.

All internal costs are not necessarily confronted by the users of transport infrastructure when making transport-related decisions. That is, they are not necessarily internal to the primarily benefiting actor, in the sense that (s)he will cover or pay for them in full (in accordance to the user-pays principle). For example, parking is often considered as causing internal resource costs which are not covered in the sense of this principle. Either the infrastructure operator providing the parking facilities incurs these uncovered costs, or parking is subsidised from outside the transport market by employers or by the government.

Internal costs can be divided into *resource costs* and *time costs*. Resource costs are typically subdivided into *variable costs* and *fixed costs*; relating to these concepts, a distinction is often made between *short-run marginal costs* and *long-run marginal costs*. For the variable cost concepts, certain factors (infrastructure capacity, car fleet, etc.) are given; for the fixed and long-run concepts, all cost categories can be variable and can be adjusted. A most famous example of long-run marginal costs in transport relates to capacity expansion of a congested road: the long-run marginal resource cost of an additional car on a congested road allows that the road capacity will be optimally adjusted.

Fixed cost categories do not vary automatically with the level of traffic or transport. However, it should be understood that the distinction between the variable and fixed costs, and the short-run and long-run costs, is to some extent analytical. In any practical policy making situation, and in analysis, which cost categories are fixed and which variable depends on the circumstances, on the time frame etc. For instance, in the extreme short run even route choice and route specific costs may be given.

External costs refer to those costs that private infrastructure users, transport service operators and infrastructure operators directly impose on other actors, either within the transport sector (typically within the same mode) or on "third parties" outside the sector. The term "directly" here means physical interaction: the external costs are in no way accounted in the transport-related decision making, and therefore are not covered either. (A similar definition, of course, applies to external benefits.)

A more analytical definition is the following: *An external cost exists when an agent's utility or production function contains a real argument whose actual value depends on the behaviour of another agent, who does not take this effect of his/her behaviour into account in his/her decision.* This definition, covering physical externalities, excludes other unpriced effects or activities such as pecuniary externalities, barter trade, etc. (see e.g. Baumol and Oates, 1988; Verhoef, 1996; Rennings et al, 1999).

This definition in principle applies to private infrastructure users, transport service operators and infrastructure operators. However, external costs attributable to the provision or existence of infrastructure unrelated to its usage are not relevant to marginal cost pricing. Such costs may include visual intrusion and barrier effects to communities and/or eco-systems. They should be dealt with by other means, including other pricing measures. In the case of new infrastructure, they should be taken into account in the cost benefit analysis. (Visual intrusion and barrier effects, in so far as they are related to traffic levels, of course are relevant to marginal cost pricing.)

External costs related to the usage of transport infrastructure use can be subclassified into *intra-sectoral externalities* and *inter-sectoral externalities*. The former are contained wholly within the transport market; the latter cover the effects on the "third parties" (cf. section 2.1) and may be seen as an "unpaid bill" which transport poses upon society at large. The former may be further subdivided to distinguish between *intra-modal externalities*, describing costs which users of a single mode impose upon each other, and *inter-modal externalities*, which users of one transport mode impose on users of another. Intra-sectoral external costs are external to users of the infrastructure but internal to the infrastructure operator. Within these categories, a further distinction could be made to address the time dimension: some externalities are instantaneous, while others materialise in the long-run.

2.3 Categories of external costs

The major external cost categories in the context of urban transport are:

(i) congestion costs
(ii) infrastructure damage
(iii) external accident costs
(iv) noise
(v) visual intrusion and barrier effects
(vi) local emissions
(vii) global emissions

Items (i) and (ii) are intra-sectoral externalities, items (v), (vi) and (vii) are inter-sectoral externalities. Item (iii) contains both elements. Notice that infrastructure damage (item ii) can (in principle) be well defined for property rights, but is not an externality in the same strict sense as are the other items. However, the implications for marginal cost pricing to be addressed in this study are the same.

Items (iv)-(vii) can be referred to as *environmental externalities*, with (iv) and (v) representing *human quality of life* impacts and (vi) and (vii) *atmospheric and ecological* impacts.

Theoretical and empirical research on the quantification and valuation of the different external cost categories have appeared in large quantities during the last decade. Broadly speaking, two different approaches for analysing these cost categories exist: a bottom-up approach and a top-down approach. Generally speaking, the former is more preferable on theoretical grounds; however, because of data problems, the latter may in many cases be more realistic. Summary presentations of the state-of-the-art are provided by Maddison et al (1996) and Rennings et al (1999).

A most recent up-to-date review of these issues is provided by a series of reports prepared for the High Level Group on Transport Infrastructure Charging during spring 1999. Also see European Conference of Ministers of Transport ECMT (1998). Therefore, here we just shortly review a few basic observations.

The external *congestion cost* is the total value of time losses that users of a transport network impose on each other. It is sometimes argued that road traffic congestion is not a genuine externality, because it is internal to the road transport sector (i.e. the view that "road users only hinder each other, and no-one else suffers"). This is not correct, and is not consistent with sound economic welfare analysis, in which the relevant level of (dis)aggregation is the level of the individual decision maker (in this case, the traveller). Indeed, if this reasoning were extended, practically all external costs could be classified as internal effects, since they are all ultimately internal to the world population!

Infrastructure damage refers to the costs of wear and tear caused by vehicles to road infrastructure. These costs are imposed upon the operator of that infrastructure. An important issue in determining marginal infrastructure damage costs is the wide variety of different users. Apart from the approximately linear dependence of damage costs on kilometres travelled, damage rises extremely rapidly with increased axle weight by a power of four. Therefore, heavier vehicles with fewer axes cause considerably greater damage than passenger cars, per kilometre travelled.

The external *accident costs* are those material and immaterial costs which motorists involved in accidents impose upon others. These may involve both intra-sectoral and inter-sectoral external costs and may include: physical damage to vehicles, transport infrastructure, personal property and the natural environment; costs incurred by legal, policy-related and emergency services; financial costs of injuries and fatalities, such as medical and funeral costs; psychological costs of pain and suffering; values associated with lives; and production losses. Valuation of some of these elements is fraught with considerable ethical problems.

Noise is an important external cost, especially in urban areas. Calculations for The Netherlands estimate that in 1987 approximately 54% of the population suffered traffic noise exceeding 55dB(A) (Maddison et al, 1996). If the analysis were restricted to urban areas, one would expect the figure to be even higher. Determining the marginal value of external noise costs is complicated by the fact that the marginal external cost function may be falling over a rather large range: at higher traffic volumes, the addition of a new driver will cause less extra noise annoyance than the first driver did.

Visual intrusion and barrier effects can also relate to traffic levels. For example, there may often be a valid "marginal cost" community severance effect caused by a road that has been there for decades because of increased traffic levels over time. Equally, visual intrusion is often seen to be traffic related, linked to composition of traffic (e.g. a sudden increase in HGVs). Also parking can generate external costs in the form of visual intrusion and barrier effects.

Local emissions and *global emissions* cover the damage caused to health and the environment (both natural and built) by the airborne pollutants produced during motorised travel. Important pollutants emitted by the internal combustion engine include NO_x (52%), CO (90%), VOC's (37%), CO_2 (19%), SO_2 (2%), particulate matters (27%), and 'black smoke' (42%), where the percentages show the road transport contribution to total emissions in the UK (1991) as presented by Maddison et al (1996). For instance, the EC DGVII ExternE project has provided estimates of the emission costs (External Costs of Transport, 1998).

2.4 Dimensions of behaviour

A prerequisite for marginal cost pricing is the ability to make estimates of relevant internal and external marginal costs. A theoretically sound identification and estimation of different marginal costs can only be based on deep understanding of the mechanisms behind the generation of such costs. This, in turn, requires identification of the different types of activities in which the users of transport infrastructure are involved. In this study, these activities are called *dimensions of behaviour*. (The notion "cost drivers", used e.g. by Link and Maibach (1999), refers to the same phenomenon.)

Table 2.4 illustrates the situation, considering road transport as an example. The table distinguishes between the following broad dimensions of behaviour: car use, car ownership and spatial behaviour. "Spatial behaviour" refers to the choice of residence and the location of other activities. These broad categories are further subdivided into more refined dimensions. In particular, the inclusion of the "number of trips" in addition to "vehicle kilometers" indicates situations where, in a given "market", a larger number of cars making proportionally shorter trips would cause higher total external costs (practically: cold starts, queuing up, etc.). The purpose of the inclusion of the item "place of driving" is to distinguish between externalities that are local in the sense that they vary over space, and externalities that are route-specific.

Regarding congestion externality, Table 2.4 makes an important distinction between two archetypes, bottleneck congestion and flow congestion, both of which may be relevant for congested transport networks. The main distinguishing feature is that bottleneck congestion is caused by the existence of physical bottlenecks in the network, such as bridges, tunnels, ramps, etc. Flow congestion, on the other hand, refers to the capacity of roads in general, and can occur also in the situation where no link has a relatively speaking smaller capacity than the other links. In real networks, observed congestion will often represent a mixture of bottleneck and flow congestion. As shown in the table, however, the contribution to pure bottleneck congestion depends only on the question of whether a user wants to pass the bottleneck, and is independent of the total length of the trip. The more closely we approach the other extreme of pure flow congestion, the more closely the overall marginal external congestion costs become dependent upon total distance travelled. (For more details see section 5.1.2.)

The table indicates the relevance of each dependence on a three point scale. The assigned stars are merely indicative and debatable; one could evidently find counter-examples. That is also the reason for using a three-point scale only. However, the table is illustrative in drawing explicit attention to the dependence between various externalities and behavioural dimensions. (A similar illustration could be given for freight transport and public transport.)

Table 2.4 *Dependence of various external costs of road transport on behavioural dimensions.*

	Car use					Car ownership		Spatial behaviour**
	Vehicle kilometres	Number of trips	Time of driving (peak or off-peak)	Place of driving (area or route)	Driving style	Fleet size*	Vehicle technology	
Intra-sectoral externalities:								
– Flow congestion	*	-	**	**	**	*	-	**
– Bottleneck congestion	-	**	**	**	-	*	-	**
– Infrastructure damage	**	-	-	-	-	*	*	**
– Accidents	*	-	*	*	**	*	*	*
Inter-sectoral externalities:								
– Noise	*	-	*	**	**	*	**	**
– Local emissions	**	*	*	**	**	*	**	**
– Global emissions	**	*	-	-	**	*	**	**

** particularly strong and direct relation

* possibly strong indirect relation, or moderately strong direct relation

– no particular strong or direct relation

[*] Also allows for car size.

[**] Location of residence vs. work and leisure activities.

3 Pricing in urban transport: basic categorisations

Transport infrastructure users are involved in a wide range of different activities or dimensions of behaviour. Correspondingly, they are also involved in a wide range of decision-making situations: from short-run to long-run decisions, and from strictly transport-related decisions to those which relate to and affect their transport behaviour only indirectly. For instance, Table 2.4 focusing on road transport, distinguishes between decisions relating to car use, car ownership, and spatial behaviour. Understanding of how marginal cost pricing and other related pricing types affect these decisions is critical to successful implementation.

Section 3.1 identifies the relevant actors, i.e. those actors which may apply marginal cost pricing in urban transport. Sections 3.2 and 3.3 then discuss marginal cost pricing as applied by transport service operators and infrastructure operators and by the government/regulator respectively.

3.1 Relevant actors

Marginal cost pricing and other related pricing in urban transport can, in principle, be applied by:

(i) transport service operators;
(ii) infrastructure operators; and
(iii) the government/regulator.

When adopted by transport service operators and infrastructure operators, i.e. by profit-maximising commercial firms and non-profit public agencies, marginal cost pricing aims to collect revenue. When applied by the government/regulator, the goal of marginal cost pricing is to internalise transport related external costs. In this case, it is often referred to as *marginal social cost pricing* or *marginal externality cost pricing.*

3.2 Pricing approaches by transport service and infrastructure operators

The purpose of pricing, when applied by transport service operators and infrastructure operators, is to collect revenue. Besides marginal cost pricing, such operators can also adopt a number of other pricing approaches. The following is a rough categorisation of possibilities (cf. e.g. Tirole, 1990; Cornes and Sandler, 1996):

(i) marginal cost pricing;
(ii) average cost pricing;
(iii) monopoly pricing;
(iv) two-part tariffs; and
(v) Ramsey pricing.

Items (iv) and (v) extend the basic idea of marginal cost pricing. Sometimes, when appropriate to make a distinction, item (i) is called *pure marginal cost pricing.*

Under perfect competition, a basic economic textbook result says that profit-maximising commercial firms adopt marginal cost pricing as their pricing rule. A perfectly competitive firm is a *price-taker*: it sets the level of production so that marginal cost equals price, rather than choosing the price (the firm faces a horizontal demand curve and "chooses to take" the given price). Analogously, on the demand side of the market, utility-maximising consumers demand additional units as long as the price they have to pay (which they, again, take as given) is lower than their willingness to pay. In a perfectly competitive world, with no government intervention, the resulting market allocation is efficient (e.g. Varian, 1992; Rosen, 1995). In such a world, goods and services will be produced and consumed up to the point where consumers' willingness to pay for the last unit just covers its social cost.

The textbook model of a perfectly competitive firm, which produces a private good, is in principle suited to representing a situation with a perfectly competitive transport service operator (provided, of course, that such operators exist). In order to model perfectly competitive transport infrastructure operators (again: assuming they exist), a model of perfectly competitive clubs may be a more appropriate approach. In this model, the club owner adopts marginal cost pricing. Such a model has been developed and analysed in the literature, analogously to the concept of a perfectly competitive firm (see Scotchmer, 1985a, 1994; Cornes and Sandler, 1996). In club theory, the *club* means a congestible, i.e. impure public good, in relation to which can be determined optimal usage, optimal capacity, and optimal membership size.

However, in practice in urban transport, the assumptions underlying the perfectly competitive model are rarely if ever satisfied, for well-known reasons (see e.g. Nash et al, 1999). A typical situation in transport service markets, as well as in infrastructure provision, is a natural monopoly, where the average cost of production decreases as output rises, so that marginal cost is unambiguously smaller than average cost and, therefore, marginal cost pricing if adopted as the pricing rule would for sure yield a deficit to the operation. A different type of reason preventing a transport service or infrastructure operator from implementing marginal cost pricing can be data problems: the calculation of the relevant marginal costs may simply be too a laborious task.

Under *average cost pricing*, the price paid by the user equals the average cost of production, i.e. total production cost divided by the number of units produced and sold. If only operating costs are considered, average cost pricing means that neither profits nor losses can be made, because the revenue collected (the product of the number of units sold and the price charged) covers total cost exactly. The calculation of average cost based prices requires only information regarding total costs and total output, making it relatively simple to compute and implement compared with the more complex concept of marginal cost prices.

This procedure is often considered as providing a reasonably good approximation to marginal cost prices in practical environments (cf. the bottom-up vs. top-down approaches in relation to the estimation of costs in section 2). From the viewpoint of the non-profit public agency, the average cost pricing rule possesses a further attractive feature in that (assuming all costs accrue solely to the operation) it breaks even automatically and, thus, avoids any possibility of requiring subsidies.

Monopoly pricing is adopted by profit-maximising firms, when possessing monopoly power. A monopoly firm faces a downward sloping demand curve, meaning that marginal revenue is less than price. In contrast to the price-taking behaviour of a perfectly competitive firm, a monopoly firm rather chooses its price and chooses this higher than marginal cost. The mark-up, over and above the additional cost of producing the last unit sold, is equal to difference between the price and marginal revenue.

Two-part tariffs involve charging a marginal cost price for each unit of service produced plus an additional fixed fee. The fixed fee may be considered as a service access charge: having paid an initial flat fee, the user is only charged for the marginal cost price of the units actually consumed. Under this approach, any operating deficit incurred by marginal cost pricing can be retrieved through the fixed fee, without losing consumers' incentives to acquire the service in efficient quantities. For instance, telephone companies commonly impose a fixed service charge on a periodic basis, which subscribers pay in addition to fees which are directly related to system use.

A two-part tariff can sometimes be adopted by a profit-maximising firm. Oi (1971), in his classical paper, when considering a perfectly discriminating monopoly firm, shows that the firm adopts a two-part tariff and that the resulting allocation is efficient (see also Tirole, 1990). Scotchmer (1985b) shows that an owner of a club who has market power charges a two-part tariff, with the fixed part reflecting the extent of his/her market power. For two-part tariffs as applied in urban transport, see Nash et al (1999).

Ramsey pricing is another departure from pure marginal cost pricing, explicitly designed to meet budget-related constraints in a socially efficient way. Under Ramsey pricing (assuming zero cross-elasticities), the gap between price and marginal cost is inversely proportional to the demand elasticity of the commodity in question (Varian, 1992; Stiglitz, 1988; Rosen, 1995). The underlying rationale is that commodities with relatively inelastic demand are less sensitive to price changes, so distortions due to deviations from marginal cost are minimised for any given volume of revenue which needs to be raised. A prerequisite for implementing Ramsey pricing is the ability to differentiate between different sub-markets. In the context of transport, this typically means differentiating between different modes, routes and/or travelling times.

This case differs from the other cases above in that social optimisation is explicitly assumed as a goal. The approach is thus especially relevant in relation to public agencies, which have close ties to the government. This pricing rule, and its equivalent rule for fiscal taxation, reflect the fact that welfare loss is minimised if prices/taxes are set so that the induced percentage deviation from quantities which would be attained under marginal cost pricing is the same for all commodities.

3.3 Pricing approaches by the government/regulator

Marginal cost pricing as applied by the government/regulator, with the aim to correct distortions due to discrepancies in marginal private and social costs, and called *marginal social cost pricing* or *marginal externality cost pricing*, covers both intra- and inter-sectoral externalities. The goal is to secure that each activity by each user will be extended to the point where the social benefit of the last unit equals the social cost. Otherwise, in urban transport, since marginal private costs and benefits can differ significantly from the corresponding social costs and benefits, the resulting market allocation would typically be inefficient.

Another approach for pricing by the government/regulator, or, more generally, for determining pricing/taxation/subsidy measures and related policies at the governmental level, is the so-called *transport accounts approach*. This is a top-down approach, as contrasted to the bottom-up nature of marginal social cost pricing. The approach is typically advocated by non-economists. See e.g. DIW et al (1998).

The rest of this study will focus on marginal social cost pricing. The notion of marginal social cost pricing in principle applies to all the categories of external costs as discussed in section 2.3: congestion, infrastructure damage, accident costs, noise, visual intrusion and barrier effects, and local and global emissions. Having estimated the appropriate marginal costs, an important question is to determine to what extent each external cost category should be internalised through corrective pricing/taxation measures and to what extent through non-price administrative means.

4 Marginal social cost pricing

A natural starting point for considering marginal social cost pricing is to identify the types of the actual *practical pricing measures*, which in principle are available for its practical implementation. Which measures are relevant in any practical situation depends on the broader socio-economic context: determined by a number of technological, institutional, legal and political constraints. For realistic and successful implementation of marginal social cost pricing, it is extremely important to carefully identify and allow for such constraints.

An important qualification in respect to the broader socio-economic context of marginal social cost pricing concerns the actual coverage of policy making and analysis: the range of the behavioural dimensions and of the technological, institutional, legal and political constraints to be included. In this study, alternative specifications in this respect are called *settings*.

Another important distinction relevant to marginal social cost pricing concerns the notions of *first-best pricing* and *second-best pricing*: the former representing marginal social cost pricing as implemented in a full social optimum, the latter as implemented in a constrained optimum.

Section 4.1 reviews the available practical pricing measures. Section 4.2 introduces the relevant settings and section 4.3 the notions of the first-best and second-best pricing.

4.1 Practical pricing measures

The different types of pricing measures, i.e. the actual charges/taxes for implementing marginal social cost pricing, can be classified in many ways. A standard and useful categorisation is the following:

(i) charges for using particular stretches of road;
(ii) charges for parking;
(iii) public transport fares and subsidies;
(iv) taxes upon the purchase and licensing of vehicles and upon associated commercial services; and
(v) taxes on the purchase of fuel.

The logic here is from microscopic measures (item i) to macroscopic measures (item v), the latter having obvious direct implications to the other sectors of the economy. The "associated commercial vehicles" in (iv) refer, for example, to vehicle insurance premiums.

Besides a means of marginal social cost pricing with the aim of internalising external costs, pricing measures (i)-(v) can also be used to promote other economic and social goals. In particular, measures (ii) and (iii) can be used for charging the relevant internal resource costs (cf. section 2.2). Indeed, this is their primary task or purpose. Measures (iv) and (v), in turn, are also typically used for fiscal taxation and thus, via the use of the revenues, for promoting wider social goals.

Another important distinction may be made between:

(i) direct demand management measures;
(ii) indirect demand management measures; and
(iii) supply oriented measures.

Category (i) may be taken to include charges for using particular stretches of road, charges for parking and taxes on the purchase of fuel (items i, ii and v in the previous list). Category (ii) includes public transport fares and subsidies and taxes upon the purchase and licensing of vehicles and upon associated commercial services (items iii and iv). Finally, category (iii) includes taxes upon both vehicles and associated commercial services and taxes on fuel (items iv and v). More particularly, non-differentiated taxes upon the purchase and licensing of vehicles and associated commercial services, which affect total car-ownership only, are primarily indirect demand management. Similarly, non-differentiated fuel taxes, while primarily affecting the number of kilometers driven, are direct demand management. The corresponding differentiated taxes are supply oriented measures, for differentiation enables one to affect supply-side characteristics like technologies used.

4.2 Settings

Considering marginal social cost pricing, an important question concerns its coverage: what are the relevant behavioural dimensions to be included? A related important question is the relevant level of (dis)aggregation. In this study, in order to secure a full and balanced coverage of these issues, the following *settings* are distinguished:

(i) focusing on road transport;
(ii) covering multimodal transport;
(iii) covering interactions with inter-urban transport; and
(iv) covering interactions with land use.

The motivation for a separate treatment of settings (i) and (ii) is obvious. A separate consideration of settings (iii) and (iv) reflects the high relevance of interactions of urban transport with other spatio-economic structures.

In particular, the importance of the interaction between urban transport systems and land use seems to be currently undervalued in transport policy (and research). Both areas are too often considered in isolation, and particularly so in relation to transport pricing. However, the adoption of an integrated approach to deal with transportation pricing and land use issues jointly, may be critical to the achievement of an optimal and sustainable urban system overall.

Besides securing the full coverage of the different behavioural dimensions and institutional contexts of policy making, the explicit consideration of settings (i)-(iv) is useful also because the existing conceptual and real-world model applications available for analysing these issues are incomplete and cannot handle all the relevant information simultaneously. In fact, the separate consideration of the different settings would be useful on that account alone: as will be demonstrated below (section 5), the different real-world model applications have originally been developed to handle issues specific to settings like (i)-(iv).

In principle, the different types of pricing measures (i)-(v) as discussed in section 4.1 are relevant in all settings (i)-(iv). However, there is a very practical problem: the determination of the appropriate level of (dis)aggregation regarding the differentiative power of the measures. Pricing measures in policies which focus on road (setting i) typically tend to be rather detailed, whereas policies within wider settings (ii-iv) and with wider coverage use rougher measures. This will be an important issue within the different real-world model applications to be discussed below (section 5).

4.3 First-best vs. second-best pricing

The notions of first-best and second-best are central concepts in economists' policy and welfare analyses. A *first-best optimum* refers to a full, unconstrained social optimum, while a *second-best optimum* is defined as a constrained optimum, taking account of the range of practical complications. Correspondingly, as representations of marginal social cost pricing, economists speak of *first-best pricing* and *second-best pricing* as the charges required to bring about an optimum in each case.

The major types of different marginal social cost pricing situations in urban transport can be categorised as follows:

(i) first-best;
(ii) second-best due to insufficient power of pricing measures to differentiate;
(iii) second-best due to distortions in other routes;
(iv) second-best due to distortions in other modes;
(v) second-best due to distortions in other sectors; and
(vi) second-best due to shadow price of public funds.

This categorisation is adopted in this study. Section 6 below will address the first-best situation (i), and section 7 the second-best situations (ii)-(vi). However, before these discussions, section 5 reviews the model types available for analysing the different situations.

5 Modelling approaches

Sections 2 to 4 reviewed basic concepts and classifications in relation to the types of costs in urban transport and the different approaches to pricing these costs. The rest of this study will focus on marginal social cost pricing (or marginal external cost pricing), which the government/regulator imposes with the aim to internalise marginal external costs.

In order to structure the treatment of marginal social cost pricing, we introduced a few important concepts: the *dimensions of behaviour* for analysing the mechanisms behind marginal external costs; the *settings* for accounting the broader socio-economic context of social marginal cost pricing; and the *first-best* and *second-best* for accounting the technological, institutional, legal etc. constraints. An overall solution to the externality problem in urban transport requires joint consideration of all these aspects.

This section will discuss the range of modelling approaches which are available, in order to represent marginal external costs and marginal social cost pricing within urban transport and to analyse the effects of such pricing. Section 5.1 reviews standard conceptual model approaches, mostly from an economic background. Section 5.2 describes real-world model applications, which incorporate economic principles alongside other social, mathematical and engineering concepts. Section 5.3 summarises the previous discussions and shows the way forward towards an integrated approach.

5.1 Conceptual modelling approaches

Economic conceptual model approaches to transportation pricing problems are based on the proposition that transport services are not essentially different from other goods or services and, hence, can be studied using the conventional demand-supply framework. However, within this general framework, certain specific features of transport markets have been emphasised.

Concerning the demand side of the market, it is common to treat the demand for transport as a derived demand which results from spatial and temporal mismatch of the demand and supply for other goods and services. For freight transport, this is straightforward. For private car traffic and public transport, it also usually holds true, such as in the case of differing locations of labour and employment resulting in peak-hour commuting. An exception is tourist traffic, where the trip itself may often be the final good that individuals want to consume. The derived character of demand is not itself a source of market failure in transport. Rather, the sources of market failure and economic distortions within urban transportation lie on the supply side. Here, urban transport markets differ from other goods markets, in particular, through the pronounced role that congestion plays in the determination of market equilibria and through the network structure of the market.

A basic classification of conceptual model approaches is the following:

- (i) the conventional static single-link model of congestion;
- (ii) increased modelling detail related to the introduction of dynamics;
- (iii) the introduction of spatial networks; and
- (iv) the representation of randomness and uncertainty.

Characteristics to the conceptual model approaches is that they focus on single specific aspects of the urban transport problem and aim to use analytically solvable and, thus, highly stylistic models.

Section 5.1.1 presents the conventional static single-link model of congestion. Sections 5.1.2 through 5.1.4 address the complications which arise through increased modelling detail related to the introduction of dynamics, the introduction of spatial networks and the representation of randomness and uncertainty. These discussions focus primarily on road transport, because the existing models do. Section 5.1.5 broadens the scope to allow for interactions with public transport, inter-urban transport and land use (i.e. to take account of the broader settings defined in section 4.2).

5.1.1 Conventional static single-link model of congestion

The traditional textbook justification for marginal cost pricing in urban road transport is based on a static model of a single transport link with a single lane, single entry, single exit and constant capacity. The "supply side" of the model builds on the so-called *Fundamental Diagram of Road Traffic Congestion*, which depicts how vehicle speed (meters per second) decreases with increasing homogeneous vehicle density on the road (vehicles per meter). On the "demand side", the model assumes road users which are homogeneous and differ only in willingness to pay to use the road.

A particularly interesting aspect of this model is the representation of *hypercongestion.* As traffic flow (vehicles per second) is the product of density and speed, it obtains a maximum for some combination of these. Allowing for the inverse relationship between speed and travel time, this gives rise to the familiar backward-bending travel time - flow curve. Furthermore, in economists' analyses, assuming that generalised user costs are based solely on travel time costs (or, alternatively, that monetary costs correlate with the time costs), such a travel time - flow curve can be transferred to obtain a backward-bending average cost curve. Such a curve is depicted as AC in Figure 5.1.1.

With the exception of zero flow and the maximum value, the backward-bending average cost curve implies that all levels of flow are obtainable at two costs: a low cost, on the upward-sloping section of the curve, where speeds are relatively high and travel times relatively low; and a high cost, on the backward-bending section, representing *hypercongestion* situations, where the reverse holds. This observation has led to heated debates in the literature, because the intersections with a typical demand curve (D) may lead to puzzling and counterintuitive results. These include the possibility of multiple equilibria (denoted x, y and z) and the fact that, in the decentralised optimum with tolling, the traffic flow may actually exceed the non-intervention flow.

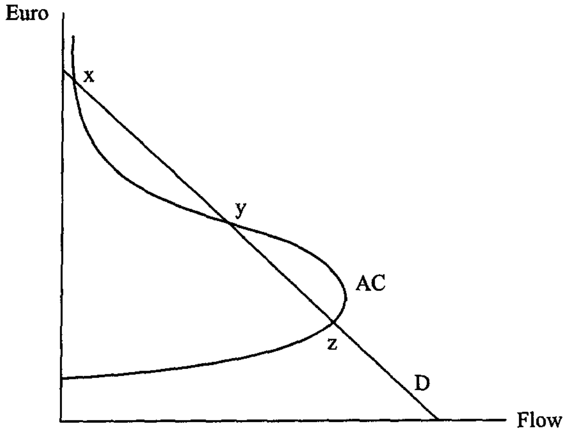

Figure 5.1. The backward-bending average cost curve.

For further details of this debate see Else (1981,1982), Nash (1982), De Meza and Gould (1987), Alan Evans (1992), Andrew Evans (1992,1993), Hills (1993), Chu and Small (1996), Yang and Huang (1998), and Verhoef (1998a,1998b). As a kind of resolution to this debate, Chu and Small and Verhoef show that the backward-bending segment of the average cost curve is dynamically unstable and, hence, irrelevant as a representation of steady state equilibria.

Though hypercongestion is extremely interesting as a theoretical phenomenon and may have some empirical relevance too, typically, in practical applications, it suffices to restrict to the upward-sloping section of the average cost curve.

5.1.2 Dynamic modelling approaches

The attraction of the static model discussed above (section 5.1.1) is its simplicity: it avoids many of the mathematical complexities associated with dynamic modelling approaches. Nevertheless, since traffic and congestion are intrinsically dynamic phenomena, it should ideally be modelled as such. Evidently the debate referred to above reflects this point. The road transport system involves a wide range of complex dynamic and spatially differentiated processes, caused by and affecting a large variety of heterogeneous agents. An increasing number of dynamic economic models have been proposed and analysed in the literature.

Dynamic conceptual models of road traffic congestion typically focus on peak congestion and explicitly describe endogenised equilibrium patterns of variables such as speeds, densities and arrival rates over time, during the peak. Two main driving forces in these models are: a given, usually common, desired arrival time; and a scheduling cost function for early departures from the origin or late arrivals at the destination. The duration of the peak and the distribution of equilibrium travel delays and scheduling costs over it are then determined endogenously, following the equilibrium condition that no user should be able to become better off by changing departure time.

Another important distinction is between bottleneck congestion and flow congestion (cf. Table 2.4). Bottleneck congestion is caused by the existence of physical bottlenecks in the network, such as bridges, tunnels, ramps, etc. Flow congestion refers to the capacity of roads in general, and can occur also in the situation where no link has a relatively speaking smaller capacity than the other links.

A bottleneck can be defined as a link in the network that has a notably smaller capacity than the joint capacity of the direct feeding links. This means that, although the congestion may spread out over significant parts of the network, in particular to upstream feeding links, its cause is to be found in the limited capacity at certain well-defined points. The overall level of congestion depends first and foremost on the capacity of the bottleneck in relation to the desired flow through that bottleneck, defined as the number of users per unit of time that would like to pass the bottleneck. The contribution of a driver to pure bottleneck congestion depends only on the question of whether (s)he wants to pass the bottleneck, and is independent of the total length of the trip. Optimal pricing, too, should then primarily depend on the capacity of the bottleneck, as opposed to the capacity of the links on which congestion (i.e. queues) is physically observed.

In real networks, observed congestion will often represent a mixture of bottleneck and flow congestion. The more closely we approach the extreme of pure flow congestion, the more closely the overall marginal external congestion costs become dependent upon total distance travelled.

The existing conceptual dynamic models reflect these distinctions. The models can be classified into three different approaches: (i) the *bottleneck* approach, (ii) the *flow congestion* approach, and (iii) the *finite group-velocity* approach. A useful state-of-the-art review of these approaches is given by Arnott, De Palma and Lindsey (1998).

The *bottleneck* approach was originally developed by Vickrey (1969), and it is later refined and extended in various directions by Arnott, De Palma and Lindsey (1993, 1998), Braid (1989), and many others. This model assumes that a bottleneck exists in the road system, which will not affect travel times so long as the arrival rate of users is below its capacity, but before which a queue will develop when capacity is exceeded. In the *flow congestion* approach, proposed by Henderson (1974, 1981), various types of flow congestion can be distinguished. For instance, Henderson (1974, 1981) and Chu (1995) assume *zero group-velocity*, meaning that a driver will maintain a constant speed during his trip, so that different speeds may occur simultaneously along the road. The opposite extreme is *infinite group-velocity*, where all drivers who are present on the road simultaneously have the same speed at each instant, so that drivers will speed up or slow down during their trips (Agnew, 1973). Finally, the *finite group-velocity* approach has recently been proposed, based on kinematic wave and car-following theories. The greater degree of realism of these models has to be traded off against the increased mathematical complexity caused by the first-order or second-order differential equations underlying them, usually precluding analytical closed-form solutions (Newell, 1988; Verhoef, 1998b).

5.1.3 Conceptual network models

The extension of the conventional single-link models described above to the consideration of spatial networks introduces a different set of complexities. Most studies have focused on very simple networks, aiming to identify the basic economic processes in transportation networks by finding analytical solutions for given network problems.

The most popular network used for such conceptual studies concerns a standard two-link case, where two parallel roads connect a single journey origin and destination (commonly referred to as an *O-D pair*), providing a simple route choice for travellers. Such networks are considered by Lévy-Lambert (1968) and Verhoef, Nijkamp and Rietveld (1996), in a static environment, and by Arnott, De Palma and Lindsey (1990, 1992) and Braid (1996) for bottleneck models. These analyses allow for homogeneous and heterogeneous travellers, with their different willingness to pay for travel, as well as for individual links of the network. For larger network models, it may often become difficult to obtain analytical solutions – at least tractable solutions. Therefore, extensions to larger networks with traffic congestion usually involve numerical simulation models (see e.g. De Palma, Marchal and Nestorov, 1995).

The principles underlying network equilibria and optima in conceptual models include Wardrop's first principle. This states that, for every O-D pair, the costs for all routes (or *paths*) which are used must be the same and that no unused path can offer lower costs. Additional equilibrium conditions become relevant when elasticity of demand is considered: the marginal benefits should then be equal to the marginal private costs (including tolls) for each used path.

5.1.4 Modelling randomness and uncertainty

The approaches outlined above normally assume deterministic congestion, with complete certainty and perfect information regarding all costs and benefits (private and social) of road usage, including external costs. However, in practice, demand and cost functions are not known with certainty. In recent years, there have been extensive studies of randomness and uncertainty, particularly related to Advanced Traveller Information (ATI) Systems and pricing, both in isolation and combined. See Verhoef, Emmerink, Nijkamp and Rietveld (1996) and Arnott, De Palma and Lindsey (1995) for recent contributions using static and dynamic models, respectively.

Important questions in these studies concern the welfare effects of providing information on actual traffic conditions. Opposite effects may often play an important role for such questions. On the one hand, it can be expected that, based on the information, road users will try to avoid severe congestion, which in itself has a positive impact on this congestion and, hence, a welfare improvement can be expected. On the other hand, the quality improvement that results from the improved information may attract latent demand, which, in turn, will lead to worse congestion. Moreover, various disequilibrium behavioural aspects, such as overreaction, may further reduce the beneficial impacts of information provision.

5.1.5 Extensions to broader settings

Sections 5.1.1 to 5.1.4 above focused entirely on road transport, reflecting the fact that, in relation to implementing marginal social cost pricing, the road is the most important mode in urban transport. However, the focus on road transport is also a result of the fact that most of the existing literature addresses this case. Conceptual economic models to consider the broader settings (i.e. multimodal urban transport, interactions with inter-urban transport, and interactions with land use, as defined in section 4.2) are fewer.

One useful approach to such analysis is provided by *club theory* (cf. section 3.1.2), which has been basis for many economists to advocate congestion pricing, and quite independently from the transportation literature. Club theory addresses the same issues as the static standard model of road congestion, but in a more general context than just focusing on roads. Besides representing a single road, a *club model* can represent a public transport system, a city etc: they can all be viewed as a club (see Cornes and Sandler, 1996; Glazer, Niskanen and Scotchmer, 1997). The theory presents conditions for profit-maximising and socially optimal pricing, capacity provision and membership. It is particularly suited for analysing interactions between pricing, investment, financing and membership issues within congested facilities.

Broader types of approaches describe the provision and consumption of local public goods within models of spatial structure as well as represent usage of scarce land (see e.g. Small, 1992; Scotchmer, 1994; Ramjerdi, 1996; Anas, Arnott and Small, 1998).

5.2 Real-world model applications

Real-world transport model applications are generally based around long-standing conceptual models of demand, supply and cost, as described in Section 5.1. They are typically designed to enable planners and decision makers to test the impacts of changes to transportation systems (usually referred to as *schemes*) against a variety of criteria, within a comprehensive software package.

The main features which distinguish real-world model applications from the conceptual models are size and scope. Whereas conceptual models typically consider very simple, idealized and symmetrical environments, real-world applications include realistic representations of geographic complexity; and while conceptual models may focus all attention on a single problem, real-world applications face demands for addressing a range of different problems together, often within tight time and budget constraints. In many cases, it is necessary for a wide range of possible schemes to be represented within the same modelling framework, to allow comparative assessment of benefits, so that the best of the options tested may be considered by politicians for financial support and implementation. Therefore, real-world applications are characterised by software packages which incorporate a number of conceptual models together and which provide input/output interfaces for large volumes of local data.

Section 5.2.1 presents an overall taxonomy of real-world applications. Sections 5.2.2-5.2.5 discuss different categories of model application in more detail. Examples of specific software packages are used during the discussion, based on the real-world model applications which are being considered within the AFFORD project.

5.2.1 Taxonomy of real-world applications

An ideal real-world transport model application would address the full range of transport and related land use policy issues at all levels of detail: from predicting the short-run welfare effects for individuals of minor local traffic management measures to forecasting the long-run consequences for economic activity, efficiency and sustainability of major shifts in national economic policy. At present, no such application exists due to limitations of behavioural research, data availability, modelling techniques, computing power and cost (cf. the discussion about settings in section 4.2). Current real-world model applications rely on trade-offs, focusing resources on particular chosen issues and levels of detail at the expense of others.

Broadly speaking, four categories of currently available real-world model applications may be identified:

(i) detailed simulation models;
(ii) tactical network models;
(iii) strategic transport models; and
(iv) geographic models.

Here the model types are arranged from microscopic (item i) to macroscopic (item iv) models.

The majority of applications currently available fall into categories (ii), (iii) and (iv). These models are normally *static*, in that they assume a steady state within the system represented, and they attempt to reach an equilibrium between forces of supply and demand based on some definition of economic generalised cost. Detailed simulation models (category i) are a more recent development and have been designed specifically to address issues which are precluded by the other traditional approaches. Most of these models are *dynamic* (i.e. they include time as an explicit variable). Table 5.2.1a summarises the most significant differences between typical real-world model applications under the four categories identified.

Detailed simulation models provide disaggregate short-run analysis of supply issues affecting roads for a compact network. *Tactical network models* are mainly concerned with short-run analysis of supply issues on road networks, but for larger networks (e.g. for a whole city). *Strategic transport models* focus primarily on long-run transportation system and travel demand issues across an urban region. *Geographic models* focus primarily on long-run issues of land use and of the economy, taking account of their relationship to the transportation system and to travel demand, across an urban region.

Therefore, *detailed* and *tactical* models are better able to represent road-based congestion and the travel choices and variability in travel conditions affecting drivers, but are less well equipped to analyse changes in travel demand and long-run issues. They are also more constrained in their potential to cover large areas than *strategic* and *geographic* models and are typically designed to focus on operational statistics relating to the performance of the transport system rather than measures of economic benefit. It is important to understand that any individual real-world model application may not fit wholly and exclusively within one of the categories identified and may involve interactions between a number of separate models (and/or software packages) with quite different features, within a common overall analysis framework.

Table 5.2.1b summarises the software packages and their local applications being considered in the AFFORD project. These provide a representative selection of real-world urban transport model applications available within Europe and will be used as examples in sections 5.2.2-5.2.5 below. Of course, there is no shortage of alternative software packages and case study locations. Most major towns and cities within the European Union might be expected to have access to a local model application similar to those described for practical transport planning purposes. While there are no comprehensive reviews available covering the use of use of real-world model applications for investigating urban transport pricing issues, an extremely informative review of *detailed* and *tactical* modelling software packages has been carried out by Watling (1994) in relation to their ability to represent ATI Systems.

Table 5.2.1a Summary of main features of real-world applications.

Category of model:		Detailed	Tactical	Strategic	Geographic
Features of typical applications:					
Private road transport:	included	***	***	***	**
	explicit network	***	***	**	-
	detailed junction delays	***	**	-	-
	time dynamics of congestion	***	-	-	-
Public transport:	included	**	**	***	**
	explicit network	**	**	**	-
	road congestion effects	**	**	-	-
	patronage effects	-	-	***	-
Disaggregation of travel demand:	in space	***	***	***	***
	by time of day	**	**	***	**
	by departure time	***	**	-	-
	by user type	**	**	***	**
Travel choices:	route choice	**	***	**	-
	departure time choice	*	*	-	-
	variable travel demand:	-	**	***	***
	– time of day choice	-	*	***	**
	– mode choice	-	*	***	**
	– origin/destination choice	-	-	**	***
	– frequency of travel	-	-	**	**
Long-run responses:	travel choices	-	-	***	**
	public transport services	-	-	**	**
	car ownership	-	-	-	**
	location & land use choices	-	-	-	***
Spatial limitations:	regional	-	**	***	***
	city-wide	**	***	-	-
	local	***	-	-	-
Nature of outputs:	operational statistics	***	***	**	-
	travel demand impacts	-	**	***	**
	economic indicators	-	**	***	***

*** an important feature of most applications
** included in some applications
* included in some applications
- not usually included

Table 5.2.1b Real-world model applications available in Project AFFORD.

Category of model:	Detailed	Tactical	Strategic	Geographic
Real-World Case Study:				
Athens	RONETS	SATURN ASCOT		
Edinburgh	RONETS	SATURN ASCOT DREAM+	START	
Helsinki		EMME2		MEPLAN
Oslo		EMME2	RETRO	

Two major studies have recently used real-world model applications to investigate urban transport pricing.

The EC DGVII TRENEN project developed a macroscopic model for the determination of optimal transport prices in a (near) general equilibrium setting. Although similar to *strategic* and *geographic* models in the sense that the focus is primarily on demand issues in a long-run perspective, this approach does not fit neatly into the categories illustrated in Table 5.2.1a. Transport consumption, in a multimodal framework, and consumers' expenditures on non-transport goods are determined endogenously in the model, so that the full economy-wide equilibrium impacts of different pricing policies can be determined and so that the equilibrium levels of taxes can be assessed. This model has been applied to various cities and regions in Europe and is documented in De Borger and Proost (1999).

The London Congestion Charging Research Programme (LCCRP) was the largest practical study of urban road pricing to be carried out in Europe in recent years (The MVA Consultancy, 1995). It made use of four independent real-world model applications within a common framework: LASER, a *geographic* model for predicting long-run regional land use and transport impacts; APRIL, a specially constructed transport version of the MEPLAN *geographic* model; LTS, a four-stage generation, distribution, mode choice and assignment model (a combined *strategic* and *tactical* approach which has lost favour due to high costs of data collection); and SATURN, one of the most detailed *tactical* modelling software packages.

5.2.2 Detailed simulation model applications

Detailed simulation model applications focus almost entirely on road transport and normally assume that all aspects of travel demand, other than, perhaps, route and departure time choice, remain fixed. They are typically *dynamic*, in that they include time as an explicit variable, some treat each vehicle on the road network individually (referred to as *micro-simulation*) and some provide a more complex treatment of user behaviour than the conventional steady state (equilibrium) assumption. A reasonably comprehensive review of the alternative software packages for micro-simulation has been provided by the EC DGVII SMARTEST project (Algers et al, 1997) and a number of these models are also considered by Watling (1994).

The RONETS model being considered during the AFFORD project is a detailed traffic simulation and assignment model developed at York University (Ghali and Smith, 1992a, 1992b). The basic model is similar to the well established CONTRAM software (Taylor, 1990) and retains the conventional equilibrium assumption. Its main areas of detail are dynamic traffic assignment and control. Vehicles are treated as individual entities which choose a single route in each assignment, rather than as aggregate flows proportioned between routes during equilibration. The model may be applied to realistic networks with many origins and destinations and contains two distinct sub-models: (i) a road network assignment sub-model, which estimates driver route choices on the basis of economic generalised cost and user optimal equilibrium assumptions; and (ii) a road network simulation sub-model, which represents the interactions between vehicles, both on links and at junctions, and calculates resulting traffic flows and network performance statistics. These two procedures are used iteratively until satisfactory levels of stability and convergence are achieved. The study period is discretised into a series of time slices during which link-based attributes, such as entry and exit flow rates, are assumed to be constant. The road network simulation is based on bottleneck modelling approaches and the *link travel cost* encountered by each vehicle on any given link is the sum of the uncongested travel cost required to arrive at the downstream junction and the bottleneck cost encountered as a result of queuing. The latter will vary between individual vehicles dependent on their position in the queue. As with most detailed simulation modelling software packages, existing practical applications are limited and usage to date has been mainly for research purposes (such as within the EC DGVII MUSIC and AIUTO projects). Therefore, RONETS has the ability to access and customise local data from other (more common) existing tactical network model applications.

5.2.3 Tactical network model applications

Most tactical network model applications focus solely on road transport and only a minority incorporate some form of demand variability other than route choice. Typically, they rely on the assumptions of a steady state within the modelled period and treat vehicular traffic as aggregate flows. Therefore, as compared to detailed simulation models, they provide a less detailed treatment of the time dynamics of congestion. However, they do possess the ability to represent traffic interactions in space on a realistic network, and they provide coverage of much larger geographical areas within practical constraints of computational capacity, compared to detailed simulation models. Variability of road travel demand and of road network travel conditions may be catered for to some extent by the subdivision of demand into a series of time slices to be applied within sequential steady state model runs, with the initial travel conditions for any time slice being based on information from the end of the preceding period. Beyond these general points, different tactical network models can have quite different features. A wide range of alternative software packages exist, the most commonly applied of which are probably EMME/2 (INRO, 1992), SATURN (Van Vliet, 1982) and TRIPS (MVA Systematica, 1991).

The SATURN software package being considered in AFFORD is a long-standing network based traffic assignment model which was developed at Leeds University. It contains two separate sub-models: one for road network assignment and the other for simulation (as in RONETS). Travel costs in the simulation sub-model are calculated on the basis of aggregate cost-flow relationships, where all drivers making a particular movement experience a single average cost, rather than separately for individual vehicles (cf. RONETS). It extends the conventional conceptual static model of congestion for single links, described in section 5.1.1, to provide turning movement specific cost-flow relationships at junctions which are continuously modified to reflect interactions with flows making other movements. Therefore, the travel costs incurred at junctions differ by turning movement and are affected by the flows of all other relevant movements at the intersection, rather than being solely dependent on the flow for the movement concerned. This is achieved by the iterative recalculation of turning movement related cost-flow relationships within the simulation sub-model and underpins the concept of *congested assignment* modelling approaches.

In its conventional form, SATURN assumes fixed road travel demand. However, the capability exists to introduce variable demand through an *elastic assignment* algorithm, which allows road travel demand to vary in response to own-price elasticity relationships. This allows the representation of changes in road travel demand which occur as a direct result of changes in travel costs experienced on the road network. Each origin to destination movement in the network is assumed to possess an additional connection to those available through explicit roads and trips are transferred between this *pseudo-link* and the rest of the network on the basis of an own-price elasticity function responding to cost changes between a base and a forecast situation. Changes to trip volumes may be positive or negative and there is currently no link to availabilities and costs of alternative travel choices, although this could clearly be incorporated through interaction with an independent demand model. The Edinburgh SATURN application has been developed specifically for the AFFORD project at a level of detail appropriate for interaction with the START strategic model application.

The ASCOT software package is a steady state equilibrium assignment model for road traffic developed at York University (Ghali, 1992). It is similar to the assignment sub-model of SATURN. However, rather than employing a specific junction simulation sub-model to estimate movement specific delays, the issue of flows on other links affecting link costs is addressed through a reformulation of the link cost calculation to incorporate all flows entering an intersection, based on evidence in the literature for different junction types (Kimber, 1976, 1980; Kimber and Hollis, 1979). There are two other principal distinctions. First, the ASCOT model considers a wide range of predetermined road travel demand levels, by gradually increasing demand from zero during the assignment procedure. Second, it incorporates the optimisation of network control strategies, through modifications to traffic signal settings, as an integral part of the assignment algorithm. Currently, the ASCOT model does not allow the demand for road travel to vary in response to changes in travel cost, although it would be perfectly possible to incorporate an elastic assignment algorithm similar to that used in SATURN.

The DREAM+ software package has recently been developed at York University, in conjunction with The MVA Consultancy and with University College, London. It is a steady-state equilibrium model with multimodal assignment and incorporates inherent travel demand variability and optimisation. The DREAM+ model is a steady state multimodal equilibrium assignment model which includes facilities for both fixed and variable travel demand and for bilevel optimisation procedures. It contains two types of network: a *base network*, in which roads and junctions are represented as a bottleneck model, and a series of *multi-copy networks*, specific to origin-destination movements, in which each link comprises a route taken from the base network. It is the copy networks which are used during the assignment procedure, using conventional principles of Wardrop user equilibrium, with costs being derived from route costs calculated in the base plus appropriate additional elements, such as parking charges and bus fares. For any given origin to destination movement, separate copy networks may be constructed to represent disaggregations of demand, such as different modes or vehicle types. The copy networks also include explicit representations of all signal-controlled junctions encountered on a route to allow the optimisation of signal settings. Switching between modes is controlled by an elastic assignment algorithm.

The EMME/2 software package is a long-standing network based traffic assignment model (Spiess, 1984; INRO, 1992). It is a *conventional assignment* model, relying on assumptions of steady state equilibrium and predicting traffic flows on the basis of economic generalised cost and link-based speed-flow relationships. It does not incorporate the more detailed approach to junction simulation available in SATURN. However, it does possess explicit facilities which aid interactions with more macroscopic strategic and geographic models and which allow users to specify their own demand model relationships.

Both the SATURN and EMME/2 software packages have many existing practical applications besides those being considered in AFFORD. In contrast, the ASCOT and DREAM+ models have primarily been used to date for research purposes and, as with RONETS, have the facility to derive local data from other tactical network model applications.

5.2.4 Strategic transport model applications

Strategic transport model applications focus primarily on the distribution of travel demand between the various travel choice options available (origin-destination, mode, time of day, trip frequency etc). They share with detailed and tactical models the basic principle of attempting to achieve an equilibrium through calculations related to economic generalised cost. However, there are at least three fundamental differences in their approach to the problem.

First, travel demand is specified in person-trips rather than in vehicle units or vehicular flows, providing more natural links to trip generation theory, to disaggregation by user type (e.g. household type, income level and journey purpose) and also to economic analysis of the society overall. Second, the mathematical approaches within the model are typically based on logit techniques rather than assignment algorithms. The former can more easily provide simple representations of the complex behavioural processes underlying travel choice decisions, while the latter are more efficient for providing simple representations of route choice behaviour on a network. Third, the representation of spatial and supply side issues within strategic models is typically much coarser. Whereas detailed and tactical models include an explicit topological representation of the road network, normally requiring a similar spatial disaggregation of travel demand, strategic transport models typically assume a more aggregate spatial demand structure (e.g. at the level of local administrative boundaries) and derive travel costs either from an aggregate cost-flow based network sub-model or by interfacing with an independent tactical model.

The ability of the existing strategic models to represent the range of travel choices available is variable. Choice of mode and time of day may be handled quite adequately within the logit formulation, based on coefficients derived from observed data. However, changes of origin, destination and trip frequency are more difficult to estimate because they depend on the availability of goods and services at alternative locations and the way in which travel choices interact with lifestyles, activity schedules and trip-chaining opportunities. For this reason, these options are often considered externally and assumed to be fixed within the transport model. Some strategic models attempt to include route choice on the road network, but where the network and road travel demand are defined at an aggregate level this may grossly underestimate the impacts of the choices available, leading to poor predictions of road travel costs (Milne, 1997). In some cases, the quality of these predictions may also have a significant impact on the other travel choices represented. The disaggregation of travel time within strategic models is usually insufficient to allow consideration of departure time choice.

The START software package being considered by AFFORD is a strategic transport model developed by The MVA Consultancy in the UK (Bates, 1991). It contains two sub-models: an *external forecast model*, used to predict changes of demand due to factors unrelated to the transport system (e.g. land use, household structure and income), and a *transport model*, to represent travel choices. The latter in turn contains two submodels: a *demand model*, which calculates changes in travel patterns based on changes in generalised cost, and a *supply model*, which deals with both *system effects* (e.g. road congestion and public transport overcrowding) and *operator effects* (e.g. responses of a bus company to changes in patronage).

The external forecast model and the transport model are usually run incrementally for any given policy scenario and there is no automatic feedback or formal land use prediction model. The supply model produces cost information for the demand model and the two elements iterate until equilibrium is achieved. The demand model represents travel choice for destination, mode and sub-mode (e.g. car trips to a public transport terminal), time of day and route. Origins and trip frequencies are assumed to be fixed. The supply model covers car and public transport, but excludes non-motorised trips. Road travel costs are calculated using an aggregate cost-flow based network and the route choice available for each origin to destination movement is restricted to a small number of predetermined alternatives. Car parking is also subject to capacity restraints, affecting travel cost through changes in time spent both searching for a space and accessing the final destination. In-vehicle travel costs for road-based public transport are affected by changes in cost on the aggregate road network, but those for non-road modes are fixed. All public transport costs are affected by service frequencies, number stopping points and effects related to patronage. The service levels of each public transport mode are able to respond to changes in patronage to a predetermined extent, across a range from maintaining current levels to adjusting in proportion to demand. This, in turn, will affect public transport travel costs.

The main outputs of the START model include both standard (aggregate) operational and environmental statistics (e.g. modal split, total vehicle kilometres, total fuel consumption, vehicle emissions, predictions of casualties from accidents etc.) and economic and financial indicators (e.g. *net present value* and *present value of finance*). The Edinburgh START application has been used widely to test packages of measures, including road pricing, designed to form part of an integrated transport strategy and has played a central role in the EC DGVII OPTIMA and FATIMA projects.

The RETRO (*Re*gional *Tr*ansport model for *O*slo and Akershus) software package was developed specifically for analysing passenger transport in the greater Oslo area. It contains two sub-models: a car ownership sub-model (Ramjerdi and Rand, 1992; Rand and Rekdal, 1996) and a travel demand sub-model. The car ownership sub-model is a national model which operates independently to predict car availability in the region based on changes in income and travel costs. The travel demand sub-model uses a nested logit approach to predict travel choices for the population (mode and trip frequency) based on travel costs. It operates in conjunction with an EMME/2 tactical network model application, which predicts route choice and travel cost on the transport network. The travel demand and route choice models are run as an iterative loop until satisfactory convergence is achieved. As with the START Edinburgh application, the RETRO model for Oslo has been used during the EC DGVII OPTIMA and FATIMA projects.

5.2.5 Geographic model applications

Besides transport issues, geographic model applications also include modelling the location of different activities that generate the traffic. The main characteristic of geographic models is that the location of various actors in the urban structure and, therefore, the resulting travel patterns can change. These models are also sometimes referred to as *integrated land use and transport models* or simply *urban models*. Geographic models are typically spatially more coarse than strategic transport models. There are some variations amongst the operational practises in the world. The structure of the model can consist either of interconnected subsystems or a unifying single system. The most commonly used theories are the random utility or discrete choice theories, equilibrium modelling and microeconomic methods. Usually, the dynamic structure is quasi-dynamic as the simulation period consists of individual cross sectional parts (Wegener, 1998).

The MEPLAN software package being considered in AFFORD is a geographic model application framework for integrating location, land-use and transport models. MEPLAN is based on a principle that transport is not a direct demand in itself but a necessary outcome of the interactions between people and workplaces that are distributed in an urban environment. As a consequence of this principle, the economic activity needs to be modelled as the source of the transport demand. The demand model structure usually consists of different types of households and employment sectors and the floorspace that they use for locating in a particular zone. Typically, the model structure consists of large number of separate categories in different stages of the model. The categorisation is very disaggregate but the individual formulae can then be simple with fewer variables.

MEPLAN uses land rent as the basis for constraining the location model. This solves the problem of implementing doubly constrained land use model in a similar way to transport trip distribution models (Harris, 1996). The land use model estimates the *crowding* effect in land use similarly as flow-delay functions create the congestion on roads in the transport model. The economic rent in a zone increases to balance excessive demand if it rises over the floorspace supply. This, in turn, reduces the attractiveness of the zone for location. The model is doubly constrained to meet the observed volumes of travel (origins and destinations) in the calibration year. The dual variables enable the model to calibrate itself to match observed land use when the constraints on it are known for the calibrated time period. These zone-specific constants can then be carried forward in an incremental fashion to the forecast period, when the economic rent can change in order to attract or push away activities, so that an equilibrium can be achieved in the land use model simultaneously with the transport demand and supply model.

Trip frequency is modelled with a general representation of economic interaction between two zones using an input-output formulation. An important feature of the MEPLAN location model is a logit type formulation for the distribution function. Nested logit modelling approaches have been adopted extensively throughout, from trip distribution to stochastic traffic assignment. This gives a strong theoretical foundation of utility maximising theory and leads to a consistent evaluation based on consumer surplus calculations (Williams, 1977; McFadden, 1978). The level of detail of the traffic assignment can be the same as for tactical transport models but the detail of the land use model often restricts this in practice to a more strategic level. The MEPLAN application in Helsinki derives traffic data from an existing EMME/2 tactical network model application which has been used previously during the EC DGVII OPTIMA and FATIMA projects.

5.3 Towards an integrated approach

Sections 5.1 and 5.2 discussed the existing conceptual model approaches and real-world model applications for representing marginal social costs and marginal social cost pricing in urban transport. The discussion reflected the fact that no single model can, in practice, address all relevant issues. A great variety of models is needed, because of the great variety of issues to be considered. In particular, the different real-world model applications can represent physical networks and travel behaviour in varying levels of detail. Economists' conceptual analyses, in turn, can provide basic insights, frameworks and building blocks for the further development of the real-world models. Another typical economists' contribution here, of course, is to provide numerical estimates for the various model parameters such as elasticities, which need to be fixed within the real-world applications.

An important observation is that the real-world model categories (i)-(iv) are distinguished along the same lines as the settings (i)-(iv) defined earlier (cf. 17). Detailed simulation and tactical models are typically used to analyse road transport, strategic transport models usually focus on multimodal transport and sometimes cover interactions with inter-urban transport and geographic models usually cover the inter-urban setting and, by definition, address interactions with land use. Evidently, the policy relevance of these settings has, to some extent, affected the development of real-world model applications to coincide.

The different real-world model application types are far from a homogeneous group with comparable properties and transferable results. In this situation, transport planners attempting to address transportation and related problems most sensibly adopt a *horses for courses* approach by choosing the type of model most suited to each situation. Where the consideration of a policy issue demands modelling strength across the categories, it is increasingly common to incorporate more than one independent model application within a common framework, sharing input data sources and passing appropriate output information between different levels.

The integration of conceptual economic analyses and real-world model applications has not previously received the attention it deserves. In particular, greater interaction of the two approaches is needed to render the real-world model applications to better allow for economic efficiency as a goal, alongside the traditionally well represented broader objectives. As will be discussed below in section 6, this is a precondition for the use of these models for analysing marginal social cost pricing.

6 First-best pricing

Section 4, when discussing *marginal social cost pricing*, distinguished between *first-best pricing* and *second-best pricing*. This section addresses first-best pricing. (Section 7 below will address second-best pricing.)

Section 6.1 discusses the concepts of first-best optimum and first-best pricing within conceptual models. Section 6.2 discusses long-run properties of first-best pricing, still relying on simple conceptual comparative static analysis. Section 6.3 discusses first-best pricing issues within different real-world model applications. Section 6.4 makes a point of the role of the first-best as a benchmark.

6.1 First-best pricing within conceptual models

Section 5.1 reviewed standard conceptual economic model approaches, however, focusing on road congestion. We now consider the representation of marginal external costs in these models and the determination of the corresponding first-best prices with the aim to internalise such costs.

To begin with, it is important to understand that the presence of external congestion costs (this evidently applies to other inter-modal externalities too) does not necessarily justify intervention by the government/regulator. It is possible that profit-maximising (or non-profit) infrastructure operators already optimally internalise such externalities. Results demonstrating this important property have e.g. been presented based on club theory (cf. sections 3.1.2 and 5.1.5).

One important result says that when membership and capacity in a club are at their optimal levels, then optimal congestion toll yields revenues such that they exactly cover the capacity cost, irrespective of the production technology (of both the club facility and club services). Also, a profit-maximising club owner secures socially optimal resource allocation concerning the use of the club, its capacity (i.e. investment, and membership size). This is true no matter whether the club owner has monopoly power or not (Scotchmer, 1985a, 1985b). Moreover, as already stated in section 3.1.2, when the club owner has market power then (s)he charges a two-part tariff, with the fixed part reflecting the extent of his/her market power.

However, these kind of results require that the whole population can be divided into identical clubs and the integer problem can be ignored (see Scotchmer, 1985a, 1985b; Scotchmer, 1994; Glazer and Niskanen, 1995). When this condition is not satisfied, then the results are quite different. When an essential characteristic of the model is that population divides between club members and non-members, then, generally, a profit-maximising monopoly owner of a congestible impure public good (or club good) charges a price which is "too high" or "too low", depending on circumstances. In particular, the monopolistic pricing rule for a congested facility with a given capacity is the sum of the marginal external congestion costs *plus* a monopolistic mark-up, implying tolls *exceeding* the short-run optimal prices (Verhoef, 1996). However, if there is a noncongested alternative mode, then the owner charges a socially optimal toll again (Edelson, 1972; Glazer and Niskanen, 1995; Cornes and Sandler, 1996).

Section 6.1.1 considers first-best pricing in the conventional single-link static model; section 6.1.2 extends the discussion to dynamic and network models.

6.1.1 First-best pricing in conventional static model

Figure 6.1.1 depicts the static single-link model of congestion, originally discussed in section 5.1.1. Here also the marginal social cost curve is introduced (section 5.1.1 considered only the average cost curve). Furthermore, besides marginal congestion costs, also other categories of marginal external costs are included in the analysis. The implicit assumption, that road users are homogeneous and differ only in their willingness to pay to use the road, continues to hold. (For further details see e.g. Verhoef, 1996.)

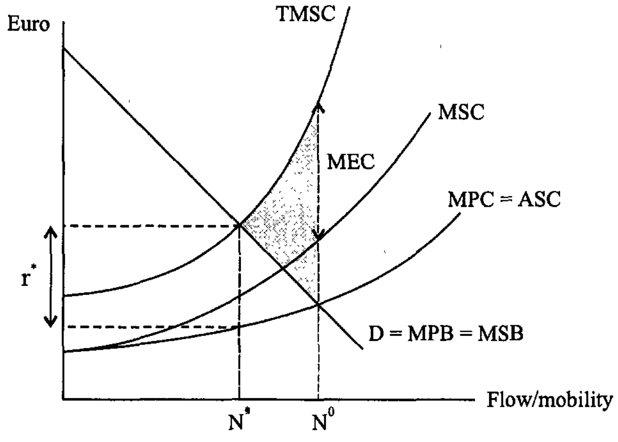

Figure 6.1.1 First-best optimum and first-best pricing in road transport.

In Figure 6.1.1, the *marginal private cost* curve (MPC) is equated to the *average social cost* (ASC); this reflects the assumption that individual road users do not consider their own impact on average speed (congestion) and safety (accidents) when deciding whether to use the road (i.e. they take the congestion and safety levels as given). This MPC=ASC curve corresponds to the average cost (AC) curve of Figure 5.1.1. The curve is positively sloped because of congestion and other intra-sectoral externalities; this reflects the increase in the private costs of road usage faced by the marginal user as the total level of road usage rises. Curve MSC represents *marginal social costs*. This curve is necessarily higher than ASC: total costs are, by definition, equal to $N \cdot ASC$, and thus $MSC = ASC + N \cdot \partial ASC/\partial N > ASC$. Along the same line of reasoning, MSC is steeper than ASC: $\partial MSC/\partial N = N \cdot \partial^2 ASC/\partial N^2 + 2 \cdot \partial ASC/\partial N > \partial ASC/\partial N$. MEC, the vertical distance between curves MSC and TMSC, represents inter-sectoral costs (accident costs and marginal environmental costs (cf. section 2.3). Curve TMSC then gives the *total marginal social cost*.

Furthermore, in Figure 6.1.1, the demand curve for trips (D) fully reflects *marginal private benefits* (MPB). These are, in turn, equal to *marginal social benefits* (MSB): D=MPB=MSB. The marginal private and marginal social benefits are identical, consistent with the view that there are no significant external benefits from road transport (the benefits are normally internal or pecuniary in nature; see Verhoef, 1996, for a further discussion).

Now, the optimality criterion can be based on the concept of net economic or social benefit as the criterion. This is defined as the sum of consumers' and producers' surpluses, i.e. the area between the curves D=MPB=MSB and TMSC. This concept reflects the more general and more fundamental theoretical concept of Pareto-efficiency.

Based on this criterion, Figure 6.1.1 shows how, due to the existence of intra-modal externalities (congestion, accidents) and inter-sectoral externalities (noise, emissions), the free market outcome will exceed the socially optimal level of road mobility. The market equilibrium is at N^0, the intersection of the curves D = MPB and MPC. Socially optimal road usage is at N^*, where net social benefits will be maximised. The welfare gain in the social optimum, as compared to the free market equilibrium, equals the shaded area.

Figure 6.1.1 also demonstrates that, after imposition of a road user charge equal to r^*, drivers between N^* and N^0 will not use the road anymore. This is socially optimal, since their benefits of road usage (MPB), which also reflect social benefits, fall short of their (common) total marginal private cost ($MPC+r^*$), which, in turn, by construction equals the social cost. In this case, the optimal marginal cost price equals the level of the marginal external cost in the optimum.

Optimality of first-best pricing such as characterised here includes marginal external costs, but does not account for any induced adjustments in the factors behind the supply and demand curves of the relevant actors in the transport market. It is widely agreed that the notion of long-run marginal cost, allowing for such adjustments, is not relevant for marginal social cost pricing. For this widely discussed topic, see e.g. Calthrop and Proost (1998) and Nash et al (1999).

This conclusion, of course, does not mean that the users of roads and other transport infrastructure should not be responsible for covering also long-run costs. However, the rationale for pricing policies aiming to such a situation, has to be different from the arguments behind the marginal social cost pricing principle.

6.1.2 First-best pricing in dynamic and network models

In conceptual dynamic models (cf. section 5.1.2) and in conceptual network models (section 5.1.3), the concept of marginal external costs is more complicated than in the single-link static model discussed in section 6.1.1 above. Also, the use of the net economic benefit or surplus to evaluate welfare impacts of marginal social cost pricing, illustrated in Figure 6.1.1 in the one-link static case, easily extends to these models. Of course, these more complicated models do not easily lend themselves to simple geometric presentation (as in Figure 6.1.1), due to interactions over time and between different spatial links.

In dynamic models, which allow for the dynamics of congestion and/or of travel demand in time and space, conclusions concerning first-best tolls can radically differ from the conclusions obtained from the static model above. The dynamic bottleneck model typically produces an optimal toll which is time-dependent, reaching its maximum for the drivers arriving at the desired arrival time. The implied kinked congestion function, which is specific to this approach, implies that, at the optimum, all travel delays are eliminated (time spent in the queue is a dead-weight loss). This means that the optimal congestion level is zero! In contrast, however, in the flow congestion approach, similarly to the static approach discussed in section 6.1.1, travel delays are not completely eliminated at the social optimum.

In network models, the concept of marginal congestion cost may be viewed both locally and globally. *Local marginal cost* is link specific. For any individual vehicle using a given link, the *local marginal social cost* is the sum of the *link travel cost* for that particular vehicle and the additional delay which it causes on that link to all other vehicles behind it in the queue. Therefore, marginal social cost will always be greater than personal cost (i.e. *link travel cost*) and vehicles at the head of the queue will expect to incur social costs which are greater than those of vehicles behind them. In contrast, *global marginal cost* bases marginal congestion cost calculations on a network of links (Ghali and Smith, 1995). Most work to date in this area has been confined to a very simple example of a fixed road travel demand and departure time profile using a single commodity network (i.e. where all road users are assumed to be travelling to a single destination).

6.2 Long-run implications of first-best pricing

The goal of first-best pricing (as discussed in section 6.1) is to provide incentives to economic agents to reduce externalities optimally. The simple static one-link model provides an illustrative and convincing argument for optimality of such pricing. However, the analysis assumes that various long-run factors determining the shapes and positions of the relevant market demand and cost curves remain stable. This means that essentially a short run-view is taken. An important question is: are also long-run impacts optimal, or are additional measures needed to steer the (induced) long-run developments into desirable directions?

The long-run optimality of marginal social cost pricing (marginal external cost pricing) in general has been shown in the context of entry-exit behaviour of firms (see e.g. Spulber, 1985). This section provides a number of applications of this general result for road transport. These examples, using simple comparative static analysis, demonstrate that marginal social cost pricing provides first-best incentives also in the long run. That is, first-best prices not only optimise transport *given* the shape and position of the demand and supply curves, but also create optimal incentives to change those aspects of behaviour that affect the actual shape and position of these curves in the long run. Of course, a fundamental theoretical idea behind these results is that the long-run marginal cost is the envelope of the short-run marginal cost, and, therefore, the two are equal in a long-run optimum. Or put differently, the envelope theorem generates the identity between short-run and long-run marginal social costs.

The relevant long-run incentives may be related to:

(i) individuals' long-run decisions behind their spatial behaviour, determining the shape and position of the demand curve;
(ii) technology choice decisions, determining the marginal external environmental cost and marginal private cost curves; and
(iii) optimal investments in road infrastructure under conditions of congestion.

In cases (i) and (ii), the relevant demand and supply (cost) curves are affected through induced adjustments in the behaviour of the users of infrastructure. In case (iii), the long-run adjustment occurs through the behaviour of the infrastructure operator.

These cases (i)-(iii) will be discussed in sections 6.2.1-6.2.3 respectively. Section 6.2.4 will then discuss implications on sustainability. Although these discussions will be cast in terms of road transport, the same principles apply to other transport modes too.

6.2.1 Long-run implications for spatial behaviour

Factors determining the shape and position of the demand curve for transport (D in Figure 6.1.1) may include in the first place locational choices of firms and households. The demand for transport, being a derived demand, depends on differences in spatial distributions of supply and demand of goods, production factors (e.g. labour in the context of commuting) and services. Considering peak demand, the flexibility in working and shopping hours can also be an important factor determining the shape of the demand function for peak traffic (in particular its elasticity).

Example 6.2.1 illustrates, by means of a simple model, the optimal incentives that first-best pricing in transport gives in terms of the locational choices.

Example 6.2.1 Implications for locational choices.

Suppose that there are N individuals, who can select a residence in either area A or area B. All individuals have identical individual (inverse) demand functions $D^{TR}(n)$ for making trips to a third area, say the city centre (CBD), where n is the number of trips made by that person. The distance between A and CBD is F times as large as between B and CBD, so both the private costs C^P and the environmental costs C^E are F times as large (there is no congestion). However, area A is generally considered to be more attractive to live in for other reasons (otherwise, area A would of course be an irrelevant alternative). Because we consider the long run, it is assumed that dwellings are offered in both areas R according to a not-perfectly-inelastic local supply curve S_R. There are no externalities other than the environmental effects of transport. Dwellings are, apart from their location, homogeneous, and it is assumed that dwellings are supplied efficiently; that is, the supply S_R coincides with the marginal social costs.

The locational benefits of living in B are normalised to zero, and $D^{LOC}(X)$ is subsequently used to give the 'excess benefits' of living in A. Hence, $D^{LOC}(X)$ represents the inverse demand for location in area A rather than B, and hence the marginal willingness to pay to live in A (rather than B) for the X'th individual (with $0<X\leq N$). We can then derive that in any long run equilibrium, with X individuals living in A and N–X in B, the 'generalised cost difference' between living in A and B must be $D^{LOC}(X)$. If this generalised cost difference is smaller than $D^{LOC}(X)$, more people would be attracted to A; otherwise, the opposite occurs. This generalised cost difference between living in A and B is in the present model given by the price difference between dwellings, *plus* the difference in net private benefits due to individually optimised mobility to the CBD, *given* the locational choice and *given* the prevailing type of transport regulation.

The total social welfare in the system can then be written out as the sum of the net benefits of locational behaviour, and the net benefits of transport (given the location chosen):

$$W=\int_0^X D^{LOC}(x)dx-\int_0^X S_A(x)dx-\int_0^{N-X} S_B(x)dx+ \tag{1}$$

$$X\cdot\left(\int_0^{n_A} D^{TR}(x)dx-n_A\cdot F\cdot\left(C^P+C^E\right)\right)+(N-X)\cdot\left(\int_0^{n_B} D^{TR}(x)dx-n_B\cdot\left(C^P+C^E\right)\right)$$

where n_R gives the number of trips made by an inhabitant of area R. We can find the overall optimum by taking the first derivatives of (1) with respect to X, n_A and n_B. This yields:

$$\frac{\partial W}{\partial X}=D^{LOC}(X)-S_A(X)+S_B(N-X)+$$

$$\left(\int_0^{n_A} D^{TR}(x)dx-n_A\cdot F\cdot\left(C^P+C^E\right)\right)-\left(\int_0^{n_B} D^{TR}(x)dx-n_B\cdot\left(C^P+C^E\right)\right)=0 \tag{2a}$$

$$\frac{\partial W}{\partial n_A}=X\cdot\left(D^{TR}(n_A)-F\cdot\left(C^P+C^E\right)\right)=0 \tag{2b}$$

$$\frac{\partial W}{\partial n_B}=(N-X)\cdot\left(D^{TR}(n_B)-\left(C^P+C^E\right)\right)=0 \tag{2c}$$

However, for given locally differentiated transportation taxes r_A and r_B, individuals will act according to the following equations:

$$D^{LOC}(X)-\left(S_A(X)-S_B(N-X)\right)+$$

$$\left(\int_0^{n_A} D^{TR}(x)dx-n_A\cdot F\cdot C^P-r_A\right)-\left(\int_0^{n_B} D^{TR}(x)dx-n_B\cdot C^P-r_B\right)=0 \tag{3a}$$

$$D^{TR}(n_A)-F\cdot C^P-r_A=0 \tag{3b}$$

$$D^{TR}(n_B)-C^P-r_B=0 \tag{3c}$$

Equation (3a) describes individually optimising locational choice, taking into account that an individual will act so as to maximise the net private benefits of transport given the location chosen, and equations (3b) and (3c) show the selection of the individually optimising number of trips, given the choice of a location, and given the prevailing transportation taxes.

Comparing (2b) and (3b), and (2c) and (3c), we first find that the optimal transport taxes should for both types of trips be equal to the marginal external costs, exactly as depicted in Figure 5.1.1:

$$r_A = F \cdot C^E \tag{4a}$$

$$r_B = C^E \tag{4b}$$

If we then substitute these taxes into (3a), it is easy to see that the incentives to locate in either area A or B are then exactly according to the social optimality condition (2a). Hence, the long-run decision of where to reside will then be made in line with overall economic efficiency, and no further regulation regarding locational decisions is necessary. At the same time, it can be seen that when the optimal transportation taxes r_R are not used and are set equal to zero, we will not only have inefficiently high mobility levels for both types of transport given the location of people, but in addition also an inefficiently high number of residents in area A, which 'boosts' also the demand for the relatively polluting type of mobility in the sense that the demand curve for trips of type A is 'too much outward rotated', and for type B '*too much inward rotated*'.

6.2.2 Implications for technology choice

In Figure 6.1.1, the vertical distance between curves TMSC and MSC, denoted by MEC, represents the marginal inter-sectoral external costs (accident costs and environmental costs) per vehicle kilometre. These externalities may often be related to the vehicle technology used. Similarly, vehicle technology is an important factor determining the shape and position of the marginal private cost curve (MPC).

Example 6.2.2 uses a simple model to demonstrate the mechanism of optimal user charges affecting the choice of vehicle technology in an optimal manner. Recall that we assume an otherwise first-best world. Possible market failures in vehicle supply as well as in the development of technologies are, therefore, ignored and we focus solely on the incentives given to the purchasers of vehicles.

Example 6.2.2 Implications for technology choice

Assume that, by obtaining more expensive cars, road users can improve the energy-efficiency, ε, *above* some default level ε_0, and hence have lower private costs (through a lower fuel input per kilometre travelled) as well as lower emissions and hence external environmental costs per kilometre travelled. The marginal cost of such improvements is given by the function $C^\varepsilon(\varepsilon)$. Assume, again, that there is no congestion. Denoting *the environmental* costs per trip as $C^E(\varepsilon)$ and the private costs as $C^P(\varepsilon)$, and the demand curve for trips as D(N) where N gives the number of trips, we can write out the following social welfare function:

$$W = \int_0^N D(x)dx - N \cdot \left(C^P(\varepsilon) + C^E(\varepsilon)\right) - \int_{\varepsilon_0}^{\varepsilon} C^\varepsilon(x)dx \tag{1}$$

The social optimum requires the selection of an optimal ε and N according to:

$$\frac{\partial W}{\partial \varepsilon} = -N \cdot \left(\frac{dC^P(\varepsilon)}{d\varepsilon} + \frac{dC^E(\varepsilon)}{d\varepsilon}\right) - C^\varepsilon(\varepsilon) = 0 \tag{2a}$$

$$\frac{\partial W}{\partial N} = D(N) - \left(C^P(\varepsilon) + C^E(\varepsilon)\right) = 0 \tag{2b}$$

Road users, on the other hand, when being informed on the nature of the user charge being made optimally dependent on the technology chosen so that it can be written as $r(\varepsilon)$, will invest in energy-efficiency improvements up to the point where the marginal private costs of doing so become equal to the marginal private benefits in terms of reduced private costs and reduced charges of road usage. Hence, road users act so as to set:

$$-N \cdot \left(\frac{dC^P(\varepsilon)}{d\varepsilon} + \frac{dr(\varepsilon)}{d\varepsilon}\right) - C^\varepsilon(\varepsilon) = 0 \tag{3a}$$

Given the choice of a technology, they will choose a level of mobility according to:

$$D(N) - C^P(\varepsilon) - r(\varepsilon) = 0 \tag{3b}$$

Again, comparing (3a) and (2a), and (3b) and (2b), it turns out that the first-best pricing rule:

$$r(\varepsilon) = C^E(\varepsilon) \tag{4}$$

simultaneously optimises the choice of technology, as well as the level of mobility *given* the technology chosen. Since the technology in this example affects, simultaneously, the marginal private costs and the marginal external costs, the example illustrates how the long-run decisions – now in terms of technology choice – are 'automatically' optimised using the marginal external cost pricing rule.

6.2.3 Implications for capacity provision

Factors determining the shape and position of the marginal intra-sectoral external cost curve (implied in Figure 6.1.1 by the vertical distance between MSC and MPC), in particular the congestion externality, are partly related to the capacity and quality of the infrastructure. The problem of the choice of optimal road capacity in the presence of congestion differs from the previous problems (in sections 6.2.1 and 6.2.2) in an important sense: in this case, the long-run decision (i.e. the choice of infrastructure capacity) is made by a different actor (i.e. the infrastructure operator) than the short-run decisions (i.e. the choice of using the infrastructure, made by the potential road users).

Example 6.2.3 demonstrates that marginal social cost pricing leads to long-run optimality result: the optimal level of infrastructure capacity would result from investing the revenues from optimal pricing.

Example 6.2.3 Implications for capacity provision.

Consider a single road, and denote its capacity as K. The average user costs of making a trip are denoted as $C^P(N,K)$, where N gives the number of users. Let $\partial C^P/\partial N \geq 0$ represent congestion (at a given capacity), and $\partial C^P/\partial K \leq 0$ mitigation of congestion through capacity expansion. The marginal social cost of capacity expansion is given by the function $C^K(K)$. Let D(N) denote the inverse demand function for using this road, so that the area under D gives the Marshallian benefit measure. We can then write out the following social welfare function:

$$W = \int_0^N D(x)dx - N \cdot C^P(N,K) - \int_0^K C^K(x)dx \tag{1}$$

The social optimum requires the selection of an optimal N and K according to:

$$\frac{\partial W}{\partial K} = -N \cdot \frac{\partial C^P(\cdot)}{\partial K} - C^K(K) = 0 \tag{2a}$$

$$\frac{\partial W}{\partial N} = D(N) - C^P(\cdot) - N \cdot \frac{\partial C^P(\cdot)}{\partial N} = 0 \tag{2b}$$

For a given capacity and with road pricing at a rate r, road users will choose a level of mobility according to:

$$D(N) - C^P(\cdot) - r = 0 \tag{3}$$

Comparing (3) and (2b) gives us the standard first-best congestion charge:

$$r = N \cdot \frac{\partial C^P(\cdot)}{\partial N} \tag{4}$$

On the basis of (2a) and (4), it is possible to show that the revenues from optimal road pricing are, under certain conditions, just sufficient to cover the cost of the optimal supply of road infrastructure capacity (Mohring and Harwitz, 1962). These conditions involve constant returns to scale in congestion technology, so that $C^P(N,K)$ can be written as $C^P(N/K)$, and constant returns to scale in capacity extension ($C^K(K)$ is constant). Writing N/K=R (ratio), we find:

$$\frac{\partial C^P(\cdot)}{\partial N} = \frac{\partial C^P(\cdot)}{\partial R} \cdot \frac{\partial R}{\partial N} = \frac{\partial C^P(\cdot)}{\partial R} \cdot \frac{1}{K} \tag{5a}$$

$$\frac{\partial C^P(\cdot)}{\partial K} = \frac{\partial C^P(\cdot)}{\partial R} \cdot \frac{\partial R}{\partial K} = \frac{\partial C^P(\cdot)}{\partial R} \cdot \frac{-N}{K^2} \tag{5b}$$

Using (5a) and (4), we can then write the total revenues from congestion pricing as:

$$T = N \cdot N \cdot \frac{\partial C^P(\cdot)}{\partial N} = \frac{N^2}{K} \cdot \frac{\partial C^P(\cdot)}{\partial R} \tag{5c}$$

Substituting (5b) into (2a) and multiplying both sides by K finally yields:

$$\frac{N^2}{K} \cdot \frac{\partial C^P(\cdot)}{\partial R} - K \cdot C^K(K) = 0 \tag{5d}$$

The first term again gives the total revenues from congestion pricing, and the second term the total costs of infrastructure capacity supply in case $C^K(K)$ is constant. In other words: the government will then have a balanced budget, and outlays on infrastructure capacity expansion can exactly be covered by the revenues from optimal pricing. The fact that (5d) implies a surplus (deficit) for the government in case of decreasing (increasing) returns to scale in capacity expansion is of course not something specific to transport – this property holds for optimal pricing in any market (Varian, 1992).

6.2.4 Implications for sustainability

The notion of *sustainable development* extends the standard concepts of environmental externalities and resource depletion to include also those negative impacts, which today's activities may have to the welfare of future generations. An important question is to what extent the standard efficiency analysis is capable of incorporating concerns about sustainability.

Regarding sustainability, important interactions exist between transportation, infrastructure, spatial behaviour, the ecological environment, and many other areas which would ideally demand a more general approach (see Verhoef, 1996). Omissions of these issues need to be balanced against the benefits of simpler analysis, which is less reliant on uncertain assumptions about the future and does not carry the risk of becoming unmanageable, due to the inclusion of too many interactions. Despite these difficulties, the notion of sustainability has to be considered if one wants to evaluate the long term impact of any transport pricing policy.

An important question is whether the results such as presented in sections 6.2.1-6.2.3 above would carry over to a broader setting, where environmental issues are considered from the perspective of sustainable development rather than using the conventional and evidently more narrow concept of marginal external environmental cost. Posing this question is like opening Pandora's box, since it inevitably raises the question of exactly what is meant by sustainability. This discussion has been lingering on for decades now, and it is fair to say that as yet no conclusive definition of sustainability has been given.

Things become even more complicated when the sustainability of transport, an economic sector among many other sectors, is studied. Sustainable development would naturally refer to the long-run behaviour of an entire economic system. It is highly debatable whether sustainability should, or even could, be defined for a sub-system, because of the likely interdependencies with other sub-systems. Clearly, the virtue and stability of sectoral sustainability is questionable as long as the overall system to which this region or sector belongs is not guaranteed to behave in a sustainable manner. At least, therefore, interactions with the entire spatio-economic system should be endogenised when studying the sustainability of transport. An attempt along those lines was recently done by Verhoef, Van den Bergh and Button (1997).

Nevertheless, for the sake of argument, we could assume that sustainable development is defined by an unambiguous set of environmental targets. These targets can either be specific to the transport sector in a single sector approach, or can refer to the entire spatio-economic system, including an endogenous transport sector under a general spatial equilibrium approach. Then, one could follow an environmental target approach, where the objective is to maximise social welfare given the environmental constraints/targets reflecting the prerequisite of environmental sustainability.

Given that such targets have been specified, it is immaterial to the welfare evaluation of transport pricing policies whether one uses the marginal external cost approach (as followed in this report) or the environmental target approach. Mathematically speaking, these two social maximisation problems are duals: one would find exactly the same tax-rules in both ways, the only difference being that the marginal external cost terms used in the standard models would be replaced by Lagrangian multipliers reflecting the social shadow price of achieving sustainability (i.e. respecting the environmental constraints) in the relevant optimum. For further details, see Verhoef, Van den Bergh and Button (1997).

Hence, as long as the general modelling approaches used are the same (partial equilibrium vs. general equilibrium, spatial vs. non-spatial, static vs. dynamic), the policy conclusions and specific tax rules found are not critically dependent on whether the marginal external environmental cost approach or the environmental target approach is used. An important difference, however, is that the explicit consideration of environmental sustainability would lead to an endogenisation of (the time-path of) the instantaneous marginal external costs, which then becomes dependent on utility levels obtained by future generations.

6.3 First-best pricing within real-world applications

Sections 6.1 and 6.2 demonstrated the efficiency properties of first-best pricing (focusing on road transport), such as they are typically presented in the context of static and simple conceptual economic model analyses. Also, we referred to the complications that may arise in dynamic models and in complicated network models.

Typically, while economic research studies have focused on deriving social welfare maximising pricing rules at a conceptual, often macroscopic level and have strongly simplified the practical contexts within which they would be applied, many mathematical simulation modelling studies of pricing, while employing complex optimisation procedures and microscopic representations of real-world conditions, have regularly simplified the economic problem and sometimes totally ignored the social welfare maximisation aspect. This also holds to many practical policy-based studies which are conducted at a mesoscopic level, combining broad economic principles for the urban transport system as a whole with extensive model representations of real-world conditions appropriate for presentation to local politicians, planners and engineers.

In addition, the consideration of marginal social cost pricing within real-world model applications requires that a number of more practical type of questions be addressed. Such questions typically relate to:

(i) the determination of the level of modelling detail required to calculate marginal costs and the corresponding practical marginal cost based prices;
(ii) the ability of (and requirement for) models to incorporate disaggregations of demand data and explicit representations of behavioural processes in order to provide useful predictions of user responses to marginal cost based charging measures; and
(iii) the ability of models to provide useful predictions of long-run travel and locational impacts of marginal cost pricing for transport and of economic effects and welfare impacts, and the extent to which this requires the inclusion of processes beyond the normal scope of a transport model.

Section 6.3.1 discusses efficiency criteria as an objective function, section 6.3.2 the representation of first-best prices within the different real-world model applications.

6.3.1 Efficiency criteria within real-world applications

In conceptual economic analyses, as discussed above (sections 4 and 6.1), the notions of first-best and first-best pricing are based on the criterion of economic efficiency, in principle covering all generated resource costs and external costs. The concept of the net economic or social benefit, defined as the sum of consumers' and producers' surpluses (cf. Figure 6.1.1), is a widely accepted and used operational approach for representing this basic welfare criterion. It is based on the consumers' willingness to pay, and, thus, on their demand functions.

In the existing real-world model applications, instead, there is a much lesser emphasis on economic efficiency as a criterion. This is the case for a number of reasons.

First, although most real-world model applications are based on conceptual modelling approaches which use economic generalised cost, their focus is primarily on providing faithful and sophisticated representations of the current real-world situation and for testing the practical policy options able to be considered by professional planners and politicians. Set against a historical background where charges and taxes for transport have traditionally had little connection to costs (cf. section 1) and where the technology for charging road users has, until recently, been considered infeasible on a large scale, the lack of emphasis on economic efficiency as a criterion within the models is hardly surprising.

Second, the range of objectives which transport policy measures are required to meet in practice is likely to be broader than a pure economic efficiency criterion. For example, there may be a desire among politicians to address equity concerns, by pursuing policies which benefit particular disadvantaged sub-sets of the population, or to aid urban planning, by improving accessibility to particular areas of the city to promote urban regeneration. This may be viewed as a way in which real-world models focus on second-best issues, as the transport policy objectives have implicitly been modified in an attempt to cater for distortions in other economic sectors which are beyond the scope of a transport model.

Third, the types of schemes to which real-world model applications have traditionally been applied include (from microscopic to macroscopic) junction improvements, public transport priorities, traffic signal policies, road building, road closures, new public transport infrastructure and new land use developments. These measures, unlike pricing approaches (and with the exception of traffic signal settings, depending on the model being applied), generally require a specific scheme design before they can be implemented in the model and do not lend themselves easily to unconstrained optimisation internally. Therefore, it is inevitable that many uses of modelling techniques in practice will be to answer '*what if...?*' style questions and achieving an optimum solution is largely dependent on the planners and politicians defining the best scenarios.

Fourth, most real-world modelling software packages have been designed primarily to provide information on the operational performance of detailed policy proposals for presentation to politicians and the public. Economic evaluation has generally been handled separately, using operational data from the models as input to a government approved Cost Benefit Analysis framework. In the UK, the resulting economic benefit measures are used to identify priorities for investment, comparing schemes from different settings, cities, regions and urban / inter-urban contexts in absolute terms, rather than through reference to a theoretical first-best optimum.

Finally, the more disaggregate *detailed* and *tactical* modelling approaches have traditionally focused mainly on supply side issues, concentrating on achieving more faithful simulations of congestion within spatial networks under the assumption of fixed demand. Therefore, the definition of economic efficiency has often been constrained to a route choice optimisation and has been applied more frequently to the investigation of network control policies (e.g. traffic signals) than to transport pricing.

There is an obvious problem that, in their current form, those (*detailed* and *tactical*) models which are best equipped to provide insights into the appropriate first-best prices which would be charged to travellers in time and space may be least able to address overall economic efficiency. Certainly, there are examples of real-world model applications being used in practice with the efficiency criterion playing no major role, so that the connection between the objective function used in social optimisation and the behavioural assumptions of the model are, at best, left implicit or, potentially, nonexistant. A minimum requirement for analysis should be to ensure that the specific goals adopted are in contradiction neither with the behavioural hypotheses nor with the social welfare maximisation (often left implicit) reflecting these hypotheses. Where necessary, outputs from *detailed* and *tactical* models may be used as inputs to external economic evaluation procedures or to *strategic* and *geographic* models which carry out a comprehensive economic welfare and economic efficiency analysis, provided that sufficient consideration is given to the difficulties (and potential inconsistencies) of transferring data between different model platforms.

More work is undoubtably required to achieve better integration of real-world model applications and theoretical economic analysis. The desire to achieve more detailed predictions for the impacts of more ambitious policies, combined with increasing computing power, is resulting in the gradual expansion of the more disaggregate transport modelling approaches to cover both supply and demand issues. This provides an opportunity to pursue improved integration. However, at the same time, research in disaggregate modelling is moving towards improved representations of behaviour, which, in some cases, has led to the abandonment of traditional equilibrium-based assignment in favour of more psychological evolutionary approaches which take account of imperfect knowledge and user learning (Liu, Van Vliet and Watling, 1995). Such initiatives could make the representation of economic efficiency within real-world transport models significantly more difficult in future.

6.3.2 Representing first-best pricing

As indicated above, any application of the concept of first-best pricing (as well as second-best pricing) within real-world model applications has to presuppose use of an objective function which is consistent with and reflects the notion of economic efficiency. This is required because the very concept of marginal cost pricing is based on it. Important variations exist between different real-world model applications in taking this issue into account.

While various aspects of pricing policy have previously been investigated within the existing real-world transport model applications, not much is said specifically about marginal cost pricing issues. However, the different types of real-world model applications have the potential to shed light on different aspects of the marginal social cost pricing problem.

Detailed simulation model applications can provide representations of first-best pricing in a continuous time context for realistic spatial networks, within the road transport setting. In the case of *micro-simulation,* this can be done separately for each road user. An important distinction exists between the concepts of *local* and *global* marginal cost. The latter is more consistent with the spirit of marginal social cost pricing theory, but is much more computationally intensive and rather less likely ever to be achievable in practice (Djikstra, 1959; Ghali and Smith, 1992a, 1992b, 1995). In particular, investigation of the uniqueness and differentiability of the marginal cost problem has shown that, unlike the steady state modelling formulation, models with a dynamic treatment of time may not produce a unique solution and may provide a non-convex and non-differentiable problem (Ghali and Smith, 1993). This may underpin difficulties in finding ways to achieve a general solution for the marginal cost problem (in a dynamic network context) in road transport in reality.

Tactical network model applications can represent first-best pricing for realistic spatial networks within the road transport sector (and, in some cases, beyond), under the steady state assumption. First-best is achieved by switching from the conventional Wardrop user equilibrium assignment approach to a system optimal assignment, which minimises total travel cost for the road transport system as a whole, rather than for individual drivers. Optimal first-best prices are defined as those charges which would be required to move from the user optimal to the system optimal equilibrium in the assignment model, either with a fixed travel demand or with demand variability under elastic assignment. There are some difficulties moving from user optimum to system optimum, related to the treatment of delays above capacity in static models. In the user optimum, there may be a change in gradient of the cost-flow curves at capacity which causes a discontinuity or *kink* in the system optimal curves, giving the potential for convergence problems and sub-optimal solutions. This may be resolved by assuming that the trend in gradient below capacity may be projected into the above capacity situation, although there is then a danger that costs above capacity may be unrealistically high (the reason behind the change in gradient in the user optimum case) and, where junction simulation models are employed, there are possibilities of mathematical problems within the turning movement specific cost-flow calculations. Some tactical network model applications are able to optimise assignment and network control in combination, allowing traffic signals to respond to changes in network flows. Such models may also include facilities to predict other impacts of road travel, such as injury accident rates, fuel consumption and atmospheric emissions, allowing more complex objective functions to be defined.

Strategic transport model applications can represent first-best pricing over a broader range of settings (i.e. including public transport), can deal better with *generally applied prices* (e.g. fuel taxes, CO2 taxes) and can usually address issues related to simple differentiation by household type, income, journey purpose and time of day. However, their relatively coarse spatial and temporal representations of travel demand patterns and marginal travel costs means that they can only provide very rough approximations for origin to destination, route, time, mode and individual traveller specific first-best prices. These limitations may be addressed to some extent by interfacing with appropriate *detailed* and *tactical* model applications.

Geographic model applications can usually represent first-best pricing across the broadest range of settings, including road transport, public transport and land use / other related markets (taxes/subsidies on income, housing or other consumption, regulation of land-use and availability of land for construction). Issues of coarseness apply in a similar manner to *strategic transport* models and to at least the same extent. In some cases, representations of the transport sector may be sufficiently limited to make interactions with a more disaggregate (typically, *tactical*) model for this purpose essential.

6.4 First-best as a benchmark

Marginal social cost pricing considers economic *efficiency* as the sole criterion of performance. A widely accepted way to represent economic efficiency as a welfare criterion in conceptual economic model approaches is through the idea of the net economic or social benefit. This criterion ultimately reflects the consumers' willingness to pay and is defined as the sum of consumers' and producers' surpluses. However, this efficiency concept is based on the standard static demand-supply framework and, thus, represents short-run or static efficiency only.

Whether also long-run efficiency is true has to be verified separately. This requires that the implications of various factors behind the supply and demand curves of the relevant actors in the transport market be investigated, in addition to the long-run effects of possible adjustments in such factors. It turns out that, under first-best conditions, short-run marginal cost pricing provides optimal incentives for long-run optimality too. That is, first-best prices not only optimise transport *given* the shape and position of the demand and supply curves, but also create optimal incentives to change those aspects of behaviour that affect the actual shape and position of these curves in the long-run.

An important property of conceptual economic model analyses is that the objective function and behavioural assumptions within the model are consistent.

This is due to the fact that the efficiency criterion reflecting individuals' behaviour has such a central role. It is self-evident that net economic or social benefit as the social welfare criterion reflects the behavioural decisions governed by the demand and supply functions.

In most existing real-world model applications, in contrast, the economic efficiency criterion has, typically, not been given such a major role (cf. section 6.3). However, any consideration of marginal social cost pricing within real-world model applications has to presuppose an objective function which reflects the notion of economic efficiency and is consistent with it. As long as other goals dominate and economic efficiency has practically no role as a goal, a real-world model application is not suited to analysing marginal cost pricing issues. As a result, while various other aspects of pricing policy have widely been investigated within the existing real-world model applications in urban transport, not much is said specifically about marginal cost pricing issues. However, notable exceptions in this respect exist, for example, the EC DGVII projects OPTIMA, FATIMA and TRENEN (see e.g. TRENEN II STRAN Final Report, 1999). In the forthcoming Deliverable 2 of the AFFORD project we will be using the EEFP objective function developed in the OPTIMA project.

The provision of a *benchmark* is the ultimate purpose of the notion of first-best. A first-best optimum typically provides an idealised benchmark only: the assumptions behind it are rather unrealistic and, typically, cannot be satisfied in real life. However, in providing such a reference state, a first-best optimum enables the analyst and the policy maker to place practical solutions and their internal differences into perspective in relation to an idealistic optimum and, thus, to better evaluate the realistic options.

The determination of a first-best benchmark is extremely straightforward in conceptual economic model analysis. The simplicity derives from the fact that the conceptual models and their assumptions, as well as the social welfare criterion, are typically highly stylised and unambiguously defined. In contrast, in practical real-world model applications, the identification of a benchmark can be less obvious. Indeed, there may be no single unambiguously correct way to define it: the analyst typically faces a great number of open choices. Which institutional constraints should be accounted for and which not? Which variables should be optimised, and which should be treated as given? And so on.

For example, it is necessary to choose whether the costs of implementing different policies should be taken into account in the optimisation or not; this ambiguity leads us to distinguish between the notions *broadly first-best* and *narrowly first-best* in section 8 below, the former including such costs and the latter not including. Implementation costs for sure exist in real life; the issue here is whether they should be accounted in the notion of the first-best benchmark, or should they rather be introduced in the context of second-best analysis only. In situations like this, one has to be practical, and simply try to figure out what would be the best strategy to serve the overall goals of the analysis.

Making such choices and, hence, deciding on a first-best benchmark , is to a large extent a matter of judgement: the first-best optimum should represent an ideal situation which, however, should not be too unrealistic. Indeed, it would make little sense to compare potential practical outcomes to something that everyone knows can never ever be achieved.

A closely related question concerns the choice of the appropriate level of detail in the characterisation of a first-best benchmark. An important feature of first-best analysis should be the power of the pricing measures to differentiate between different links and different types of users of the transport infrastructure, as well as between the externalities they are generating. However, the operational definition of first-best can still be very different depending on the setting and real-world model application adopted. The definition of first-best in this respect has to be consistent with the level of (dis)aggregation adopted in the analysis at large.

How detailed the definition of a first-best benchmark in any given practical situation should be, also depends on the technical feasibility and costs of generating the data required to calculate marginal cost prices. If data can only be made available at a coarse level, any marginal cost pricing scheme based on such data necessarily has to be at the same level of detail. Correspondingly, the first-best benchmark , which the pricing scheme is designed to support, has to be defined at the same coarse level too. Equally, the adoption of a coarser approach may be encouraged by the costs of implementing and administering detailed (in relation to the differentiative power over time and vehicles and spatially) marginal cost based pricing schemes. In some cases, when the costs are extremely prohibitive, this may result in the implementation of a pure average cost based scheme.

Another assumption or feature of first-best analysis, in particular when it is (explicitly or implicitly) partial equilibrium by nature, is that all other relevant markets operate efficiently: the spatio-economic system within which the transport system operates is otherwise in a first-best optimum. That is, there exist in other routes, modes or sectors (whichever is relevant depends on the adopted setting) no uncorrected market failures due to external effects, market power etc. In particular, the *other sectors* can include the housing market and the automobile market.

A third feature is the lack of consideration of explicit revenue constraints of the government. However, as a more general matter, the shadow price of government revenue, reflecting the costs of collecting tax revenues in other sectors by means of distortionary fiscal taxes as compared to the costs of collecting the same revenues by (supposedly) non-distortionary externality taxes in the transport sector, can be allowed in the first-best. This is another example reflecting the relativity of the notion of the first-best: a feature which in a truly general equilibrium type of analysis could be explicitly treated as a second-best constraint may be accepted as a fact of life in a partial equilibrium analysis focusing on the transport sector alone.

7 Second-best pricing

The distinction between the notions of first-best and second-best is a central feature in much of the literature on conceptual economic models. The wide range of second-best situations addressed reflects the variety of complications and imperfections brought by different technological, institutional, legal and political constraints affecting urban transportation that must be taken into account in practical policy making.

This section considers such an analysis. The following five major types of second-best situations will be considered (cf. section 4.3):

(i) insufficient power of pricing measures to differentiate;
(ii) distortions in other routes;
(iii) distortions in other modes;
(iv) distortions in other sectors; and
(v) shadow price of public funds.

The reasons for second-best in cases (iv) and (v) of course are beyond the scope of transportation policy. Therefore, typically, when comparing first-best and second-best solutions on partial equilibrium approaches focusing on the transport sector (rather than being general equilibrium and covering all sectors), they are accepted as *facts of life*. This reflects the degree of freedom that the analyst has and the need for judgement the (s)he has to exercise when defining a first-best benchmark.

The economics literature dealing with second-best issues in terms of conceptual economic models is large. In contrast, the distinction between first-best and second-best, as standard and well-established as it is within conceptual economic analysis, has not received much attention in the existing real-world model applications.

Section 7.1 discusses second-best pricing principles within conceptual economic models. The discussion is confined to short-run issues only (cf. the rather broad coverage of long-run impacts in relation to first-best pricing in section 7). Section 7.2 shortly summarises the treatment of second-best complications within currently available real-world model applications.

7.1 Second-best pricing within conceptual models

The second-best pricing (and investment) rules have received ample attention in the economic literature on transport-related issues (on road pricing in particular). For instance, Wilson (1983), and d'Ouville and McDonald (1990) study optimal road capacity with suboptimal congestion pricing; Braid (1989), Arnott, De Palma and Lindsey (1990) and Laih (1994) consider uniform or step-wise pricing of a bottleneck. Arnott (1979) and Sullivan (1983) look at congestion policies through urban land use strategies.

Two classic examples on second-best regulation in road transport are Lévy-Lambert (1968) and Marchand (1968), studying optimal congestion pricing with an untolled alternative, an issue that is discussed also by Braid (1996) and Verhoef, Nijkamp and Rietveld (1996). Glazer and Niskanen (1992) as well as Verhoef, Nijkamp and Rietveld (1995) study second-best aspects related to parking policies, and Mohring (1989) considers fuel taxation.

A joint conclusion from these studies is that, in second-best situations, economic efficiency requires pricing instruments to be applied according to different *rules* than those that apply for the first-best policy. A general feature is that the second-best optimal price no longer equals *direct* marginal externality cost, i.e. marginal congestion cost and/or marginal environmental cost, but instead allows for a correction term reflecting the relevant second-best distortion.

Sections 7.1.1-7.1.5 discuss second-best cases (i)-(v) as categorised above. The models used are simple, but are sufficient to make the point. Models could be made more realistic, but at the price of increasing complexity: they may no longer have analytically tractable solutions, and may be solvable only using numerical procedures for models with explicit functions. However, the general conclusions given would not be affected. Also, similarly, as in section 6 above, the examples and models are cast in terms of road transport.

7.1.1 Insufficient power of pricing measures to differentiate

First-best prices have to vary in many dimensions. In particular, they should vary along with variations in marginal external costs. Considering road transport as an example (cf. section 2.3), optimal user charges should, typically, vary according to the following dimensions: the kilometrage, the time of driving, the place of driving, the specific route chosen, the vehicle technology used, the actual operating condition of the vehicle (age, level of maintenance etc.) and the driving style. Strictly taken, only then could the ability of road user charging to provide optimal incentives to change behaviour in both the short-run and the long-run carry over to real-life situations.

Clearly, such charges are not realistic in practice, due to the limited power of practical pricing measures to differentiate. Detailed approximations of such charges would only be possible if the government/regulator were to adopt complex electronic systems, using very sophisticated technologies which can monitor actual emissions, place and time of driving, driving style and the prevailing traffic conditions and could adjust the charge accordingly. Even then, there are issues relating to the calculation of first-best prices in real time on real networks and the need for differentiation to cater for the real heterogeneity of users (income, journey purpose etc.) which would be beyond the scope of any currently foreseeable technology. Also, collecting this type of information may not be socially acceptable, as it could be considered an excessive intrusion of drivers' privacy. Therefore, first-best systems are not likely to be introduced on any significant scale in practice in the foreseeable future.

As a result, when implementing road pricing in reality, policy makers must typically rely on imperfect substitutes to the first-best scheme: second-best pricing measures with imperfect power to differentiate between different types of road users and the different types of externalities they generate. This can be the case, for example, when vehicles have different emission coefficients, which the regulator cannot observe, or when, for other practical and political reasons, (s)he is not capable or permitted to differentiate taxes accordingly.

The literature on second-best pricing/taxation presents increasingly complex policy rules for increasingly imperfect instruments. Examples 7.1.1a and 7.1.1b illustrate the situation. The first example considers an environmental externality, the latter a congestion externality. These examples are taken from Verhoef, Nijkamp and Rietveld (1995).

Example 7.1.1a Second-best optimal common environmental taxes.

Suppose that M different groups of car drivers jointly use a certain road network. Each group (denoted m=1,2,...,M) has its own specific marginal private cost of driving c_m and its own specific marginal external (environmental) cost $\partial E(N_1,N_2,...,N_M)/\partial N_m$ (conveniently denoted $\varepsilon_m(N_m)$ hereafter). N_m gives the number of trips made by group m. Now suppose that the regulatory body wishes to reduce emissions by means of some second-best undifferentiated regulatory tax policy. More ambitiously, it wishes to find the optimal common regulatory fee. This fee maximises social welfare under the inherent limitation of the policy, being the impossibility of first-best tax differentiation. Such an optimal common fee can be found by solving the following Lagrangian:

$$\Lambda = \sum_{m=1}^{M} \int_{0}^{N_m} D(x)dx - \sum_{m=1}^{M} N_m \cdot c_m - E(N_1,\ldots,N_M) + \sum_{m=1}^{M} \lambda_m \cdot \left(c_m + f - D_m(N_m)\right) \qquad (1)$$

yielding the following first-order conditions:

$$\frac{\partial \Lambda}{\partial N_m} = D_m - c_m - \varepsilon_m - \lambda_m \cdot D'_m = 0 \quad m = 1..M \qquad (2)$$

$$\frac{\partial \Lambda}{\partial f} = \sum_{m=1}^{M} \lambda_m = 0 \qquad (3)$$

$$\frac{\partial \Lambda}{\partial \lambda_m} = c_m + f - D_m = 0 \quad m = 1..M \qquad (4)$$

Using (2) and (4), it follows that:

$$\lambda_m = \frac{\varepsilon_m - f}{-D'_m} \quad m = 1..M \qquad (5)$$

Using (3) and (5), the following second-best toll can then be derived:

$$f = \sum_{m=1}^{M} \frac{\frac{\varepsilon_m}{-D'_m}}{\sum_{n=1}^{M} \frac{1}{-D'_m}} \qquad (6)$$

This fee is a weighted average of the marginal external environmental costs in the second-best optimum, where the weights are inversely related to the slope of the groups' demand curves. This shows that a group should more strongly affect the second-best optimal fee, the more responsive it is to price changes. The Lagrangian multipliers λ_m represent 'shadow prices of non-optimal pricing'. These multipliers would all be equal to zero when s_m is equal for all m and hence equal to f by (6). In all other (non-trivial) cases, the (absolute value of) λ_m increases with the difference between s_m and f, and with the groups responsiveness ($-D_m'$) to this differential.

Example 7.1.1b The optimal common congestion fee.

Along the same lines as in example 7.1.1a, an optimal common congestion fee can be derived. An analytical solution for the general case with M groups only exists when congestion is strictly group specific. The Lagrangian then reads:

$$\Lambda = \sum_{m=1}^{M} \int_0^{N_m} D(x)dx - \sum_{m=1}^{M} N_m \cdot c_m(N_m) + \sum_{m=1}^{M} \lambda_m \cdot \left(c_m(N_m) + f - D_m(N_m)\right) \qquad (1)$$

yielding the following first-order conditions:

$$\frac{\partial \Lambda}{\partial N_m} = D_m - c_m - N_m \cdot c_m' - \lambda_m \cdot \left(c_m' - D_m'\right) = 0 \quad m = 1..M \qquad (2)$$

$$\frac{\partial \Lambda}{\partial f} = \sum_{m=1}^{M} \lambda_m = 0 \qquad (3)$$

$$\frac{\partial \Lambda}{\partial \lambda_m} = c_m + f - D_m = 0 \quad m = 1..M \qquad (4)$$

Using (2) and (4), it follows that:

$$\lambda_m = \frac{N_m \cdot c_m' - f}{c_m' - D_m'} \quad m = 1..M \qquad (5)$$

Using (3) and (5), the following second-best toll can then be derived:

$$f = \sum_{m=1}^{M} \frac{\dfrac{N_m \cdot c_m'}{c_m' - D_m'}}{\displaystyle\sum_{n=1}^{M} \frac{1}{c_m' - D_m'}} \qquad (6)$$

Equation (6) gives a weighted average of the marginal external congestion costs in the second-best optimum. However, the weights are somewhat different. A group now receives a larger weight, not only the flatter the demand curve, but also the flatter the average cost function in the second-best optimum. The reason is that the weights should be such that the welfare losses due to deviations of the common fee from the optimal individual fees are minimised. With congestion, such welfare losses not only depend on the slopes of the demand curves, but on those of the average cost functions as well. The steeper both curves, the less responsive the group size to deviations of the optimal common fee from the optimal individual fee, and therefore the smaller the group's weight.

7.1.2 Distortions in other routes

The second-best optimal road tolls given distortions on other routes in a road network can be illustrated by considering the classic two-route problem (Lévy-Lambert, 1968). This type of second-best problem may actually often be self-imposed by the regulator. In particular, when electronic charging mechanisms are used, it may be considered inefficient to apply charges on all links, due to the high fixed costs of installing the necessary equipment. Hence, the regulator may choose to have toll-points installed only on a few key-links in the network. More generally, it is likely, for institutional reasons, that most urban charging systems will operate within a geographic boundary which causes route distortions around the periphery of the city. The second-best tolling problem resulting from such situations has recently been studied for general networks by Verhoef (1998c). Example 7.1.2 illustrates the situation in terms of the two-roads model.

Example 7.1.2 The two-route problem.

Consider a two-link network, connecting a joint origin-destination pair. Road users distribute themselves over both routes according to the rule that marginal private costs, including tolls if there are any, are the same on the two routes. Now if there is congestion, it is easy to show that if the regulator can levy a charge on both routes, optimal congestion charges for these routes are:

$$r_i = N_i \cdot \frac{dC_i^P(N_i)}{dN_i} \tag{1a}$$

where the subscript i denotes the particular route considered. However, if the regulator is for some reason only capable of putting tolls *into effect on one route only* (say route T), and has to leave the other route (U) untolled, it would be incorrect to apply the first-best tax rule (1a) as if first-best conditions apply throughout the network. Instead, for this particular problem, the following second-best tax rule for route T can be derived:

$$r_T = N_T \cdot \frac{dC_T^P(N_T)}{dN_T} - N_U \cdot \frac{dC_U^P(N_U)}{dN_U} \cdot \frac{-\frac{dD(N)}{dN}}{\frac{dC_U^P(N_U)}{dN_U} - \frac{dD(N)}{dN}} \tag{1b}$$

This rule reflects the principle that, for the specification of the second-best *one-route* toll, one has to take account of the specific situation on the other, untolled route, as well as of the prevailing demand structure (in particular the demand elasticity, or more precisely: the slope of the demand curve). *This has to do with* the spill-overs that regulation on route T implies for the driving conditions on route U, and with the fact that one single tax aims to control two variables affecting the overall efficiency: the overall level of demand, and the route split. Note that the expression is composed of a term reflecting the marginal external cost on the tolled route, and a second (negative) term representing those on the untolled route, weighted with a fraction that may vary between 0 and 1. Also observe that the second-best toll may be negative. (For more details see Verhoef, Nijkamp and Rietveld, 1996.)

7.1.3 Distortions in other modes

Another important assumption underlying the standard first-best prices/taxes is that alternative transport modes are efficiently priced. The validity of this assumption is often questionable. For example, public transport services may be inefficiently priced from an overall social welfare perspective, due to subsidies. The implications for second-best tax rules in private transport can be derived in a manner which is comparable to the two-route problem considered above. Example 7.1.3 demonstrates how in a second-best situation, where alternative modes are not efficiently priced, the standard first-best tax rule is no longer optimal. The second-best tax rule to be used reflects the distortions occurring in the other modes.

Example 7.1.3 Distortions in other modes.

Consider a simple model, with the following assumptions. Consider the short-run, so that only variable costs matter. The generalised variable costs for public transport, as experienced by its users, are made up of two components: the price of the ticket P^T, and a term C^T reflecting the valuation of the average (per passenger) travel time in public transport. The *total short-run* social costs of public transport are given by the total variable costs made by the operator, TVC^{TO}, plus the travel time costs C^T times the number of users N^T. There is neither congestion nor a Mohring-effect present *in* public transport. (The Mohring-effect is the reverse of congestion, reflecting the positive externality that public transport users create for each other through the increased frequency that is (in the long run) associated with *increased usage.*) We use, as before, $C^P(N^P)$ to denote the average generalised costs for road usage, where N^P gives road usage. Finally, there is one shared demand for transport, D(N), where $N=N^T+N^P$.

Mode choice in this model results from generalised private cost differences. One could also model this choice using finite cross-elasticities of demand. However, that would be an unnecessary complication for the present purpose, but also quite restrictive in the sense that this cross-elasticity assumes much homogeneity of users, and little substitutability between private and public transport. Taking the view that the eventual good demanded is the move from A to B, a trip by private or by public transport would be perfect substitutes. (The same argument could also be applied to multimodal trips, by allowing consumers to make modal split decisions at multiple points (under appropriate constraints of 'car-availability' at each of these points), and thus by treating the entire trip as a sequence of uni-modal trips.)

Finally, we wish to take account of the fact that the public transport operator may have some market power. In particular, he is not a price-taker and can, for instance, change the price depending on average costs. This will be reflected below by the very general formulation that the ticket price P^T may depend on the level of usage N^T. For the present purpose, the exact pricing rule used need not be made explicit.

Under these assumptions, the second-best congestion toll for road transport r can be found by solving the following Lagrangian, showing that the objective is to maximise the difference between total benefits and total costs, under the restrictions caused by individually optimising behaviour, equating marginal benefits to marginal private generalised costs for both modes:

$$\Lambda = \int_0^{N^P+N^T} D(x)dx - N^P \cdot C^P(N^P) - N^T \cdot C^T - TVC^{TO}(N^T)$$
$$+\lambda_P\left(C^P(N^P)+r-D(N^P+N^T)\right)+\lambda_T\left(C^T+P^T(N^T)-D(N^P+N^T)\right) \tag{1}$$

yielding the following first-order conditions:

$$\frac{\partial\Lambda}{\partial N^P} = D - C^P - N^P \cdot \frac{dC^P}{dN^P} + \lambda_P \cdot \frac{dC^P}{dN^P} - \left(\lambda_P + \lambda_T\right) \cdot \frac{dD}{dN} = 0 \tag{2}$$

$$\frac{\partial\Lambda}{\partial N^T} = D - C^T - \frac{dTVC^{TO}}{dN^T} + \lambda_T \cdot \frac{dP^T}{dN^T} - \left(\lambda_P + \lambda_T\right) \cdot \frac{dD}{dN} = 0 \tag{3}$$

$$\frac{\partial\Lambda}{\partial r} = \lambda_P = 0 \tag{4}$$

$$\frac{\partial\Lambda}{\partial\lambda_P} = C^P + r - D = 0 \tag{5}$$

$$\frac{\partial\Lambda}{\partial\lambda_T} = C^T + P^T - D = 0 \tag{6}$$

Using (3), (4), and (6), it can be shown that:

$$\lambda_T = \frac{\dfrac{dTVC^{TO}}{dN^T} - P^T}{\dfrac{dP^T}{dN^T} - \dfrac{dD}{dN}} \tag{7}$$

Using (2), (4), (5), and (7), the following second-best toll can then be derived:

$$r = N^P \cdot \frac{dC^P}{dN^P} - \left(\frac{dTVC^{TO}}{dN^T} - P^T\right) \cdot \frac{-\dfrac{dD}{dN}}{\dfrac{dP^T}{dN^T} - \dfrac{dD}{dN}} \tag{8}$$

The sign of (8) is ambiguous. The first term shows the direct impact of the toll on congestion on the road itself. The second term reflects that in the second-best optimum, also account should be taken of a possibly non-optimal price in public transport. The term between the large brackets represents the difference between the marginal social costs of using public transport and the ticket price. Evidently, if public transport is efficiently priced, this term vanishes, showing that the standard Pigouvian toll suffices for the regulation of road use if the alternative mode is managed according to first-best standards. Note that the correction factor increases in the extent to which public transport prices are distorted.

As is shown by the fraction behind the term between the large brackets, the extent to which this distortion affects the second-best road price r depends also on the elasticity of the total demand for transport, and the sensitivity of public transport prices to its usage, both evaluated in the second-best optimum.

At one extreme, where the demand is perfectly elastic, the second-term vanishes, reflecting that the usage of public transport cannot be affected with the road price. The same extreme results if the public transport system is operating at its capacity, and the price P^T is used by the operator to keep out excessive demand exceeding the capacity. Then, the use of public transport is given and determined by its capacity. *In both cases*, the road price can be set according to the first-best rule, since the use of public transport is not affected.

At the other extreme, where either the demand is perfectly inelastic or the public transport price is insensitive to it usage, the second-best road price becomes equal to the difference between the marginal external congestion costs on the road, and the extent to which the marginal social costs of public transport exceed the ticket price. With inelastic demand, this reflects that the total usage of both modes together is given, and the road price should be used so as to equate the *marginal social costs for both modes in the second*-best optimum. With insensitive public transport prices, it also reflects that the overall level of transport demand is given, but now by the intersection of the price-line P^T and the demand curve D. Also *then, the distribution* of this given number of users over both modes in the second-best optimum should of course be such that the marginal social costs are equalised.

For intermediate cases, the interpretation of the correction term can be given by considering the joint impacts of the effects just discussed for extreme situations.

Finally, it can be noted that the Lagrangian multipliers λ again reflect the 'shadow prices of non-optimal pricing'. Such multipliers are typical for second-best analyses. These multipliers cause the second-best optimum to differ from the *first-best situation, where* also the alternative mode is optimally priced. In particular, note that if P^T could be chosen freely by the regulator in the optimisation procedure, we would find $\partial\Lambda/\partial P^T=\lambda_T=0$.

7.1.4 Distortions in other sectors

A further important assumption underlying the standard first-best tax rule is that all other economic sectors, somehow connected to the transport sector, operate under first-best conditions. In particular, given the fact that most economic sectors require transportation for their operation, the assumption actually requires the absence of market power and unpriced environmental pollution throughout the economy. Of course, this normally will not be the case.

Example 7.1.4 illustrates that if these conditions do not exist, or if taxes are not used optimally, the second-best tax rule for transport will be affected accordingly. Simply ignoring the distortions elsewhere in the economy would then be non-optimal and would lead to regulatory taxes for transport that can be improved upon. As is illustrated in Verhoef, Van den Bergh and Button (1997), in a spatial analysis of the above problem, the naive use of standard first-best taxes may in some cases even be counter-productive, in the sense that positive taxes for transport could lead to a reduction in social welfare. An important issue here is also the scale of different economic sectors. Where the size of an external sector may be much larger than that for transport (e.g. business), a relatively small distortion in the larger market may have a very major impact on optimal transport prices.

Example 7.1.4 Distortions in other sectors.

Consider the case of freight transport. Assume two polluting economic sectors (A and B), and assume that their production processes are polluting, causing constant average external costs C_A^E and C_B^E. Observe that the demand for freight transport is a derived demand, *that is* closely connected to the demand (and supply) structure for the transported good itself. In particular, the transportation of a good is (normally) a necessary step in the process of bringing the demand and supply *physically together*, and accomplishing a transaction. The model therefore assumes that every unit of good traded requires a transport movement. Defining the units of both goods such that the transport effort for one unit requires the same *unit transport service*, the equilibrium demand for transport is simply equal to the sum of equilibrium quantities traded, Q_A and Q_B (note that in this non-spatial model all trips have equal length).

Assume that no congestion occurs, and denote the constant average private and external costs of transport as C^P and C^E, respectively. Denote the demand and supply curves for both goods (i) as D_i and S_i, and assume that apart from the externality, both markets operate efficiently, with prices reflecting true marginal social costs. The private average transportation costs C^P *will* thus drive a wedge *between* the marginal benefits D and the marginal production costs S (see also the restrictions in the Lagrangian below). Finally, assume that only regulatory transport taxes r are available (otherwise, we would not have a second-best problem). The following Lagrangian then represents the second-best optimisation problem:

$$\begin{aligned}\Lambda = &\int_0^{Q_A} D_A(x)dx - \int_0^{Q_A} S_A(x)dx - Q_A \cdot C_A{}^E + \int_0^{Q_B} D_B(x)dx - \int_0^{Q_B} S_B(x)dx - Q_B \cdot C_B{}^E \\ &-\left(C^P + C^E\right)\cdot\left(Q_A + Q_B\right) \\ &+\lambda_A\left(S_A(Q_A) + C^P + r - D_A(Q_A)\right) + \lambda_B\left(S_B(Q_B) + C^P + r - D_B(Q_B)\right)\end{aligned} \tag{1}$$

which has the following first-order conditions (where primes denote derivatives):

$$\frac{\partial\Lambda}{\partial Q_A} = D_A - S_A - C_A{}^E - C^P - C^E + \lambda_A\cdot\left(S_A' - D_A'\right) = 0 \tag{2}$$

$$\frac{\partial\Lambda}{\partial Q_B} = D_B - S_B - C_B{}^E - C^P - C^E + \lambda_B\cdot\left(S_B' - D_B'\right) = 0 \tag{3}$$

$$\frac{\partial\Lambda}{\partial r} = \lambda_A + \lambda_B = 0 \tag{4}$$

$$\frac{\partial\Lambda}{\partial \lambda_A} = S_A + C^P + r - D_A = 0 \tag{5}$$

$$\frac{\partial\Lambda}{\partial \lambda_B} = S_B + C^P + r - D_B = 0 \tag{6}$$

Substitution of (5) in (2), and (6) in (5) yields:

$$\lambda_A = \frac{C_A{}^E + C^E - r}{S_A' - D_A'} \tag{7}$$

$$\lambda_B = \frac{C_B{}^E + C^E - r}{S_B' - D_B'} \tag{8}$$

As in the previous model, these multipliers reflect the shadow price of non-optimal pricing, now in the two goods markets. These multipliers are for both goods increasing in the difference between the marginal external costs, of production and transportation together, and the regulatory tax. Equations (4), (7) and (8) finally imply the following second-best transportation tax:

$$r = C^E + \frac{\dfrac{C_A{}^E}{S_A' - D_A'} + \dfrac{C_B{}^E}{S_B' - D_B'}}{\dfrac{1}{S_A' - D_A'} + \dfrac{1}{S_B' - D_B'}} \tag{9}$$

The second-best tax rule shows that, in addition to the 'first-best component' reflecting the marginal external costs of transport itself, a term is added which reflects the marginal external costs caused by production in the two sectors. More precisely, a weighted average of these marginal external costs is included in r, where the weight reflects the sensitivity of the equilibrium output to price distortions: if either the demand or the supply for a sector is fully inelastic, the associated term vanishes. This is of course rather intuitive: due to the inelasticity, the emissions cannot be affected, and the best thing to do for the regulator is to set the tax such that emissions from transport and from the other sector are optimised. Note also that if we happen to find $C_A{}^E = C_B{}^E$, the first-best outcome can be reproduced, since the road tax then simply includes also the external costs of production. Because every good produced is also transported, and all shipments are assumed to be equally long, a tax on transport alone is then in fact indistinguishable from the set of first-best taxes on both production and transport.

The tax rule in equation (9) shows that distortions in other economic sectors will generally affect optimal transportation taxes. Clearly, the economically optimal way of dealing with such distortions would be to use regulatory taxes directly targeted at the sectors involved, and to apply first-best tax rules throughout the economy.

7.1.5 Shadow price of public funds

The final and somewhat different type of second-best distortion concerns the case where the government's budget constraint is binding and the government uses the tax revenues from transport pricing to reduce distortive taxes elsewhere, on, for instance, labour. It has been argued, for instance by Ochelen, Proost and Van Dender (1998), that a double dividend can be obtained. Such a higher social value of tax revenues is often modelled using a shadow price of public funds (SPPF), which denotes the additional reward that is given to each unit of tax revenues. Example 7.1.5 derives a second-best rule reflecting these issues.

Example 7.1.5 Shadow price of public funds.

Consider a simple transport model with an environmental externality only, and maintain the notation used in the previous examples. The following Lagrangian can be set up:

$$\Lambda = \int_0^N D(x)dx - N \cdot \left(C^P + C^E\right) + \lambda_P \cdot r \cdot N + \lambda \cdot \left(C^P + r - D(N)\right) \qquad (1)$$

where $\lambda_P>0$ denotes the case where toll revenues are used by the government in a way that enhances economic efficiency; $\lambda_P<0$ denotes the situation where the government spends the money in a less efficient way than consumers would. The following first-order conditions apply:

$$\frac{\partial \Lambda}{\partial N} = D - C^P - C^E + \lambda_P \cdot r - \lambda \cdot D' = 0 \qquad (2)$$

$$\frac{\partial \Lambda}{\partial r} = \lambda_P \cdot N + \lambda = 0 \qquad (3)$$

$$\frac{\partial \Lambda}{\partial \lambda} = C^P + r - D = 0 \qquad (4)$$

Substitution of (3) and (4) into (2) yields the following tax rule:

$$r = \frac{C^E - \lambda_P \cdot N \cdot D'}{1+\lambda_P} = \frac{C^E \cdot \left(1 + \lambda_P \cdot \frac{N \cdot -D'}{C^E}\right)}{1+\lambda_P} \qquad (5)$$

where the second formulation merely facilitates interpretation. Equation (5) shows how, even if $\lambda_P>0$, the implied tax certainly needs not exceed the standard first-best rule $r=C^E$ applying with constant *marginal* external cost. The reason is that the sub-goal of revenue maximisation may require an upward or a downward adjustment, depending on the elasticity of demand. This is caused by the fact that a marginally higher tax rate on the *one hand increases* the tax revenue per road user, but on the other hand decreases the number of road users. In the extreme of a perfectly inelastic demand, additional taxes revenues can be generated without affecting demand, and the sub-goal of *externality* regulation then becomes completely unimportant (observe, however, that the assumed constancy of λ_P will of course become less realistic as total tax revenues approach infinity). With a relatively elastic demand, however, (when $-D'$ approaches zero from above) a downward adjustment on the standard Pigouvian tax rule is called for, since in that case a lower tax rate is associated with higher revenues. In particular, we find:

$$sign\left(r - C^E\right) = sign\left(\frac{N \cdot -D'}{C^E} - 1\right) \qquad (6)$$

Hence, if either the demand is relatively inelastic or the marginal external cost relatively low, a tax rate exceeding C^E will be found.

The shadow price of public funds carries a close relation to the so-called Double-Dividend literature as it has emerged in the environmental economics literature (e.g. De Mooij, 1999). This literature is basically concerned with questions of whether a shift from distortionary labour taxes to environmental taxes would lead to social benefits in addition to the direct Pigouvian benefit (the 'first dividend') of reducing – or, ideally, optimising – external costs.

The concept of a shadow price of public funds is highly appealing and relevant from the economic viewpoint (see e.g. Snow and Warren Jr, 1996). As illustrated in example 7.1.5, it would directly affect tax rules to be set. There are, however, many pitfalls in its actual use. In the first place, it is evident that the SPPF would depend directly on the specific allocation of tax revenues considered. The SPPF may well be negative (or smaller than unity, depending on the exact definition of the SPPF) if revenues are used inefficiently. For example, Mayeres and Proost (1999) find a net welfare reduction from road congestion taxes when the revenues are used to subsidise (already inefficiently low priced) public transport. A consequence of the SPPF depending directly on the allocation of revenues is that, when using the same tax instruments (e.g. fuel and cordon charges) and different recycling schemes, different second-best optimal tax levels will be found. Moreover, the SPPF for a given type of revenue use (e.g. lowering sales taxes) will, of course, also generally vary between sites.

It is important to emphasise that the use of an exogenously determined SPPF is a modelling approach that can only be justified on the basis of pragmatic arguments. Ideally one should evaluate the efficiency of government expenditures in the same structured manner as the efficiency of private behaviour. This in particular holds true for uses of tax revenues within the transport sector. Marginal social net benefits of, for instance, investments in public transport and/or road capacity are typically declining (implying a decreasing SPPF if revenues were strictly ear-marked for such purposes), and *endogenous* with respect to the type of pricing prevailing on this new (as well as the initial) infrastructure capacity. But also for other purposes, SPPF is typically non-constant. This would become relevant in the application of a positive (or above-unity) SPPF in transport network models with inelastic overall demand. This would invariably drive optimal second-best taxes, allowing for SPPF, to infinity. This observation is not primarily meant to criticise these models for the assumption of inelastic demand, but more as a reminder that modelling results become less meaningful when relating to (price) regions for which the model was never intended to be used.

Related to the endogeneity of SPPF is the observation that all other taxes in the economy are normally justified by some policy objectives. These include the financing of the provision of public goods, but normally also include a general purpose of wealth redistribution (e.g. from higher to lower incomes, or from employed to unemployed). Strictly speaking, the primary incidence of, for example, road charges considered ought to be evaluated partly also according to this criterion. To be concrete, an externality reducing road tax would, for a given type of redistribution, be more attractive when it is the groups with a relatively low, rather than the groups with a relatively high 'welfare weight', that are relatively strongly present on the road(s) taxed. In the former case, the road tax plus recycling better fits society's preferences for income redistribution. As a consequence, the second-best optimal charges will be higher.

The said complexities pose enormous problems for a responsible use of the SPPF in concrete modelling exercises. In particular, while the assumed value of SPPF may strongly affect the second-best optimal level of taxes, due to relative inelasticity of demand, no consensus on empirical values of SPPF exist, in particular not on the detailed level required for many analyses. Furthermore, since the use of an exogenous SPPF in a partial equilibrium model is methodologically problematic anyway, a rather pragmatic solution necessarily will often have to be adopted in practical modelling exercises. In the AFFORD project, these issues will be further addressed in Deliverable 2.

7.2 Second-best pricing within real-world applications

The distinction between the notions of first-best and second-best has been a central feature in much of the literature on conceptual economic models. The usefulness of this distinction is in that the first-best provides a benchmark against which more realistic second-best cases representing different real-life situations can be contrasted.

A general result applying to all the second-best cases analysed above in section 7.1 is that, while first-best prices fully internalise marginal external costs and are equal to *direct* marginal externality cost, the second-best optimal pricing rules include correction terms reflecting the relevant and explicitly assumed second-best distortions.

Even in stylised conceptual models, the second-best rules can easily become complicated as models are added more detail. Therefore, only the simplest model settings possible were considered. These models allow us to concentrate on the basic economic issues at hand in analytically solvable ways. The conclusions that could be made from more detailed model extensions may, possibly, not be very different.

Within real-world applications, in contrast, the distinction between first-best and second-best (in the economic sense as it has been discussed here) has achieved much less attention thus far. Also, separate treatment of different second-best situations (e.g. as above in section 7.1) has not often been carried out explicitly. Therefore, there has not been a clear focus on the economics of second-best pricing issues either. All this reflects the difference in perspective which has traditionally existed between conceptual economic modelling approaches and the real-world model applications which are subjected to a much greater range of complexities to represent and demands to be met.

However, the generation of real-world model applications which currently exist have a great potential for representing different second-best cases and for contrasting the second-best optima with the corresponding relevant first-best benchmark. Given the wide range of different second-best situations, and the wide range of differences in coverage and detail of the alternative real-world model applications, it is obvious that different applications are best suited to analysing different situations. For instance, it would clearly not be sensible to introduce government revenue constraints explicitly within a *detailed simulation* *model* application. At the other extreme, a *geographic model* application has very limited ability to represent road transport congestion and may be completely unable to distinguish between different technological approaches for levying road user charges and, therefore, for representing the very different impacts that they may have.

Table 7.2 provides a summary of the applicability of different types of real-world model applications for addressing the different second-best pricing issues, based on typical examples of such models (cf. Table 5.2.1a).

Table 7.2 *Relevance of different real-world applications to analysing different second-best situations.*

	First-best	Second-best due to				
		inability to differentiate	distortions in other routes	modes	sectors	shadow price of public funds
Detailed	***	***	***	-	-	-
Tactical	***	***	***	*	-	*
Strategic	***	**	*	***	*	***
Geographic	***	*	-	-	***	***

***	highly relevant
**	moderately relevant
*	hardly relevant
-	no relevance

Detailed simulation model applications and *tactical network model* applications, due to their detailed treatment of network structure and route choice, are well suited to the representation of second-best pricing issues under distortions in other routes and for assessing the performance of second-best pricing due to limited capability to differentiate. In particular, they can provide faithful representations of real-world pricing technologies and distinguish between the impacts they may have (Milne, 1997). In addition, *detailed simulation models* represent variability which occurs in the transport system over time explicitly, allowing temporal elements of the route and differentiation distortions to be addressed, while *tactical network models* may be used, to some extent, for representing the shadow price of public funds.

The primary strength of *strategic* modelling approaches is their ability to represent comprehensive (packages of) practical second-best pricing strategies across the full multimodal urban transport system. Therefore, this type of model is, in principle, the only category which is able to represent all cases (i)-(v). However, the primary focus is on the inter-modal issues. Differentiation may be strong in terms of distinguishing different types of user, but the limited representation of spatial networks results in a much lesser ability to distinguish between different methods of levying charges and to address distortions between routes. Representation of the shadow price of public funds is possible, but the distortions within other sectors may only be covered to a limited extent, if at all. *Geographic model* applications focus on the economics of land use and, generally, only represent the transport sector in so far as it affects long-run spatial decisions. Therefore, these models are particularly suited to analysing inter-sectoral pricing distortions (and associated shadow price issues), but are rarely able to provide inputs in the other areas.

An obvious strategy to proceed is to combine different types of model applications to provide a sufficiently comprehensive overall view. Important questions which arise in this context are: (i) the need for addressing compatibility issues when significantly different real-world modelling approaches and levels of detail are incorporated within a common framework; and (ii) the need for securing that marginal cost pricing for road use properly interact with charging approaches for other modes.

In the first-best case, as was discussed in section 6 (technically shown in section 6.2), the optimal incentives created by the first-best pricing imply that decisions regarding long-run issues can be left to the market. In the second-best case, there is no reason to expect that the implied long-run incentives would be perfect. Instead, it is evident that complementary policies directly aiming to take care of the relevant long-run issues are needed. However, no explicit analysis of this issue is included in this study.

The informational and organisational burden for the regulator trying to implement such complementary long-run measures, along with the short-run second-best pricing measures themselves, can become an enormous problem. One can ask whether the regulator would ever be able to collect and process efficiently all the information which is required regarding the relevant markets. As a result, inevitable additional welfare losses due to government failures will further reduce the social benefits of second-best regulation, over and above the welfare losses that result from the use of imperfect second-best instruments rather than the perfect first-best instruments.

However, it should be emphasised that the informational and organisational problems related to the second-best situations are not restricted to pricing/taxation instruments only. Evidently, careful implementation of alternative administrative means can easily involve the regulator in similar considerations. Therefore, in many cases, an obvious alternative to complicated second-best policies and complementary long-run policies, with their heavy informational and organisational burden, should be to try to directly remove the very reasons for the second-best under considerations, by direct intervention at the source of the distortion.

8 Marginal cost based policy packaging

The task of marginal social cost pricing, in providing optimal incentives to transport infrastructure users to change their behaviour, can be extremely complex. This is due to the great number of behavioural dimensions and categories of external costs to be accounted for. A further complication is that the different behavioural dimensions can simultaneously affect several cost categories. Rather than considering overall solutions, the literature typically has addressed individual measures in taking care of isolated problems. This reflects the general feature of conceptual economic model approaches in focusing on individual issues at a time.

The AFFORD study is investigating the concept of *policy packaging*, that is the joint use of different practical measures to achieve the marginal social cost pricing goal. An important feature of the policy packages considered is that they are specifically *marginal cost based*, so that first-best and second-best pricing rules are at their core, rather than the much broader range of issues which have often been used as a justification for policy packages in practice.

Section 8.1 discusses economic rationale for designing marginal cost based policy packaging. Section 8.2 lays down general categories of policy packages. Section 8.3 gives examples of actual packages.

8.1 Economic rationale of policy packaging

If first-best pricing measures were possible, the various behavioural dimensions determining marginal external costs in urban transport would, by definition, be affected optimally. However, in real life, the available pricing measures typically have only limited power to differentiate between different types of infrastructure users and the different categories of external costs they are generating. Therefore, such second-best measures can only reproduce partially the behavioural responses that first-best pricing would have. In the construction of marginal cost based policy packages, the full set of first-best incentives, with respect to all behavioural dimensions, should ideally be covered.

Table 2.4, considering road transport as an example, provided an overview of the dependence of various external costs of road transport on different behavioural dimensions. As indicated, similar illustrations could be given regarding freight transport and public transport. Another fundamental link behind marginal cost based policy packaging concerns the question to what extent various possible policy instruments are capable to discriminating according to these behavioural dimensions, and therefore to affecting the different externalities.

Continuing the road transport example, Table 8.1 considers, for a selection of possible second-best pricing instruments, the extent to which they are capable of discriminating according to the behavioural dimensions identified in Table 2.4.

Table 8.1 Impact of various second-best pricing instruments on different behavioural dimensions.

	Car use					Car ownership		Spatial behaviour
	Vehicle kilometres	Number of trips	Time of driving (peak or off-peak)	Place of driving (area or route)	Driving style	Fleet size	Vehicle technology	
First-best benchmark pricing	***	***	***	***	***	***	***	***
Direct demand management								
ERP per km	**	**	**	*	-	**	-	**
Toll booths	*	**	*	*	-	**	-	*
ERP Cordon	*	**	**	*	-	**	-	*
Peak permits	*	*	*	*	-	**	-	*/-
Area licences	*	*	*	*	-	**	*	**
Parking fees	*	**	*	*/-	-	*	-	*
Non-differentiated fuel taxes	**	**	-	-	-	**	**	**
Indirect demand management								
Non-differentiated vehicle taxes	*	*	-	-	-	**	-	-
Subsidising public transport	*	*	-	*/-	-	*	-	*
Subsidising tele-working	*	**	*	*/-	-	*	-	-
Location subsidies/taxes	*	-	-	*/-	-	-	-	*
Supply-side oriented measures								
Differentiated vehicle taxes	-	-	-	-	-	-	**	-
Differentiated fuel taxes	-	-	-	-	-	-	**	-

*** optimal (first-best) impact
** likely direct impact, possibly approaching first-best standards
* possible direct impact
– no particularly strong direct impact, or at least unlikely in practice

The qualifications made in Table 8.1 are rough and the scores assigned assume that the particular instrument is used optimally, given its inherent practical restrictions. Item "fleet size" allows for car size, and "spatial behaviour" allows for location of residence vs. work and leisure activities. The table distinguishes between non-differentiated fuel and vehicle taxes as demand side instruments and corresponding differentiated taxes as supply side instruments (cf. section 4.1). A non-differentiated vehicle tax will not directly (if at all) affect the specific type of vehicle (in terms of emission coefficients) selected. A differentiated vehicle tax will: the socially (more) optimal choice is decentralized, in that consumers are rewarded for buying cleaner vehicles by paying lower taxes.

Second-best pricing measures, with their imperfect power to differentiate, cannot provide perfect incentives for infrastructure users in relation to the external costs they are generating. Also, there is no reason to believe that such measures would be sufficient to secure optimality of long-run developments in mobility behaviour either. Therefore, besides second-best pricing, there may be a need for complementary policies to affect long-run decisions concerning the factors behind the long-run position and shape of the relevant demand and cost curves (cf. Figure 6.1.1) in a socially desirable manner.

Besides the second-best issues related to external effects, as discussed above, a comprehensive policy packaging should also allow for short-run efficiency in relation to resource costs, long-run efficiency, sustainability, equity concerns, fiscal needs and wider social goals (cf. section 4.1). In this situation, when individual measures are used to take care of a number of different goals at the same time, it is extremely important for policy making to be transparent: that is, for policy makers to indicate explicitly to what extent each measure carriers what task.

In particular, on many occasions, it is important to consider equity issues in an explicit manner. Implicitly equity impacts were included in the efficiency analysis of first-best and second-best pricing in sections 6 and 7. First, when summing up the benefits and losses of the different groups, in the concept of net social benefit, we give the same weight to the benefits (and losses) of all groups. Second, when using the willingness-to-pay approach in the valuation of externalities and other impacts, the incomes of consumers of course are a very relevant factor. These observations reflect the fact that the Pareto-criterion favours the prevailing income distribution.

Distributional or equity considerations can play an important role for various reasons. First, there of course is fairness in economic terms (based on changes in welfare, as calculated by an economic model) and a perceived equity perspective, covering a range of noneconomic impacts. In turn, these more or less objectively measurable distributional impacts can, along with various subjective considerations, be extremely important for the acceptability of policies, as there typically are clear winner's and clear losers. For these reasons, separate measures as an integrated part of policy packages, may be needed to allow for equity aspects.

Finally, an economically sound policy packaging should reflect and provide an explicit treatment of various relevant trade-offs. First, such trade-offs arise when individual pricing measures are required to take care of different behavioural dimensions and different marginal external costs at the same time (cf. section 2.4). Second, important trade-offs exist between efficiency in the short-run and long-run, and in particularly between efficiency and equity: increasing equity may require compromising in efficiency, and vice versa. A particularly important trade-off exists between equity impacts in the short-run and efficiency impacts in the long-run. For example, compensations given to the receptors of externalities can have an impact on their incentives, possibly defeating the (primary) purpose of the policy. That is why it is important to specify how much weight should be given to the equity criterion, if we cannot rule out this trade-off by appropriate compensation schemes.

8.2 Categories of policy packages

A useful framework for the design of practical policy packages contains a categorisation of different package types. These types should be sufficiently comprehensive so that they together reflect the main aspects of pricing in urban transport. Typically, packages should address the following dimensions:

(i) the reference case;
(ii) the first-best benchmark ;
(iii) the second-best cases; and
(iv) measures to cater for acceptability concerns.

In practice, there may be more than a single package within each dimension. Table 8.2a illustrates the range of packages which need to be considered.

Table 8.2a Policy packages.

Reference packages: *Business as usual* *Foreseen*
First-best packages: *Narrowly first-best* *Broadly first-best*
Best practice second-best packages: *Second-best under current institutions* *Second-best after institutional reform*
Acceptable package

Regarding the *reference package* or *scenario package*, reflecting local current situations and policy plans, it is useful to distinguish between two versions: the package *Business as usual*, which assumes that all instruments are set on the current levels, and the package *Foreseen*, which assumes instruments which are designed to reflect current policy plans. Which one is more appropriate in each case depends on local circumstances.

Table 8.2b Packages and settings.

	Settings Focusing on road	Covering multimodal	Covering interaction with inter-urban transport	Covering interaction with land use
Reference packages:				
Business as usual	X	X	X	X
Foreseen	X	X	X	X
First-best packages:				
Narrowly first-best	X	X	X	X
Broadly first-best	X	X	X	X
Best-practice second-best packages:				
Second-best under current institutions	X	X	X	X
Second-best after institutional reform	X	X	X	X
Acceptable package	X	X	X	X

Regarding *first-best benchmark packages*, again, two different packages can be defined. The package *narrowly first-best* assumes optimal unconstrained use of all possible instruments. All pricing instruments reflect the principle of marginal social cost pricing. All prices are fully flexible over time and place, can be differentiated perfectly, and charging itself is costless. The package *broadly first-best* represents optimisation of social welfare while taking account of the costs of implementing the policy measures. If the equipment necessary for charging fully flexible and differentiated prices is too costly, it is broadly first-best to apply second-best taxes.

As to *second-best packages*, or *best-practice second-best packages* as they are labelled in Table 8.2a, it is useful to distinguish between two packages: *second-best under current institutions*, which reflects optimal use of instruments that are available under current institutional arrangements, and *second-best after institutional reform*, which reflects optimal use of all instruments that are available after relevant institutional reform, including establishment of appropriate co-operation between local, national and EU levels. The former package can be referred to as a *best practice short-run solution*, the latter as a *best-practice long-run solution*. In some cases, the latter package may be very near to the package Broadly first-best, in particular if other governments set their instruments in a way that happens to be optimal from the local perspective.

Finally an *acceptable package* should assume instruments which are designed in regard to particular local (country or city specific) acceptability concerns as well as possibly conflicting interests of different agencies.

An important issue is the practical representation of different packages types in the different settings as introduced in section 4.2. In principle, all package types should be relevant in all settings. The situation is illustrated in Table 8.2b.

The different types of packages in the context of the different settings will be analysed in the forthcoming Deliverables 2 and 3 of the AFFORD project.

8.3 Examples of practical policy packages

The mapping of Tables 2.4 and 8.1 allows one to design concrete policy packages that satisfy the general principles laid down in sections 8.1 and 8.2. Example 8.3 specifies simple illustrations of such packages. These particular examples will be developed further and analysed in Deliverable 2 of the AFFORD project.

Example 8.3 Practical policy packages.

Package 1: "Best practice second-best"

Charge motorists
(i) toll cordon with charges of 2 Euro during the morning peak (7.00 - 9.00 a.m.) and 0.5 Euro thereafter
(ii) parking charges increased with 0.5 Euro/hr
(iii) fuel taxes increased with 0.5 Euro/litre
and use the revenues:
(iv) two thirds to lower labour taxes
(v) one third to invest in capacity expansion of known road traffic bottlenecks

In any practical case, it is important to specify whether this package should be interpreted as that of before or after institutional reform. Note, however, that the relatively large increase in fuel taxes can only be accomplished in at least a national, and probably only a European context. Hence, in most cases this package is probably closer to after institutional reform.

Package 2: "Acceptable"

For cities with an existing toll cordon (i.e. Oslo):

Charge motorists
(i) toll cordon charges of 1 Euro at all times (including nights and weekends)
(ii) parking charges increased with 0.25 Euro/hr
(iii) fuel taxes increased with 0.125 Euro/litre
and use the revenues:
(iv) one third to direct tax recycling (lowering of fixed vehicle taxes)
(v) one third to invest in capacity expansion of known road traffic bottlenecks and/or to improve parking facilities
(vi) one third to lower fares in public transport and/or to improve facilities for cyclists and pedestrians

For cities with no existing toll cordon:

Charge motorists
(i) *parking charges increased with 0.25 Euro/hr, and with a fixed fee of 1 Euro*
(ii) *fuel taxes increased with 0.125 Euro/litre*
and use the revenues:
(iii) *one third to direct tax recycling (lowering of fixed vehicle taxes)*
(iv) *one third to invest in capacity expansion of known road traffic bottlenecks and/or to improve parking facilities*
(v) *one third lower fares in public transport and/or to improve facilities for cyclists and pedestrians*

The difference with the situation where a toll ring already exists is that we do not want a bias here due to the introduction of an extra tax as such. The overall increase in taxes, though, is about the same. The package reflects a relatively mild use of pricing instruments; that is: a modest move towards marginal cost prices for road transport.

The packages defined in Example 8.3 are highly stylised. However, the types of charges are selected to reflect the main behavioural dimensions (time of day (peak-cordon) / spatial concentration (parking) / total trip length (fuel)), but the levels are illustrative only.[1]

Realistic optimal first-best and second-best levels should result from the AFFORD modelling exercises. When analysing the prices in individual cities, it is important to allow for local conditions. Local differences may give rise to deviations, and such deviations should be acknowledged, motivated and considered explicitly in the analysis. However, at the same time, it is important to ensure comparability between cities. These issues will be addressed in Deliverable 2 of the AFFORD project.

More comprehensive packaging procedures, still focusing on road transport, could also cover other policy goals, such as revenue raising and cost recovery. A further extension concerns the possibility that different measures in one package can in practice be the responsibility of different agencies at different levels: local,

---

[1] The prices mentioned are based on earlier work with the TRENEN model (Van Dender, 1998). For Brussels, the TRENEN model concluded that for a small petrol car "The optimal tax is 0.487 ECU per kilometre (off-peak: 0.208 ECU/km)". Taking as the weighted average car trip length 10 km's, this means that the optimal taxes per trip are 4.87 and 2.08 Euro. Assuming that, in urban traffic, the average car uses 1 litre per 10 kilometers in the off-peak and 1 litre per 8 kilometers in the peak, and that paid parking duration is on average 4 hours for peak travellers (not everybody has to pay for parking, not everybody stays 8 hours) and 2 hours for off-peak travellers, our package produces the following taxes per trip: *Peak*: Cordon toll 2.0 + Fuel (10*0.125*0.5) 0.6 + Parking (4*0.5) 2.0 = Total 4.6; *Off-peak*: Cordon toll 0.5 + Fuel (10*0.1*0.5) 0.5 + Parking (2*0.5) 1.0 = Total 2.0. This seems reasonably close to computed optimal prices for Brussels. Surely, these assumptions are open to discussion. The direct translation of 'optimal prices' into 'tax increases' reflect the implicit assumption that current taxes reflect other considerations than those considered here; in particular externality regulation. The validity of this assumption may also vary between sites.

national and EU. An example of a package illustrating the situation is the following:

(i) parking charges are set locally or even at the neighbourhood level;
(ii) cordon tolls are set regionally or locally;
(iii) fuel taxes are set by the EU or nationally;
(iv) differentiated vehicle taxes are set nationally; and
(v) vehicle technology standards are set at the EU-level.

In this example, no fewer than five different agencies would be involved in the implementation of the package. Clearly, this may yield numerous problems in the field of institutional co-operation.

It is rather straightforward to extend the policy packages so as to include explicitly also other transport modes and interactions with inter-urban transport and land use. In particular, specific policies covering interactions with land-use would include spatial planning and locational subsidies and taxes. In fact, the pricing principles considered here can be directly applied to land use as well. A typical approach for accounting the impacts of land use changes are regulatory planning constraints, while developers are seldom required to contribute towards the transport costs of their activities other than through minor local infrastructure improvements. However, recently, in the United States, concern about urban sprawl has resulted in the charging of *impact fees* for new land use developments to ensure that the full costs of infrastructure provision are paid and, thus, to encourage optimal development decisions. Another feature characterising land use is that the overall costs of major land developments are often artificially low due to competition between localities for investments that are expected to bring employment and economic growth. In such cases, the application of regulatory planning constraints may be even less rigorous. See e.g. Brueckner (1999). An example of a package allowing for such interactions could be to use:

(i) parking charges to deal with largely time-invariant local externalities, such as parking and noise annoyance;
(ii) ERP-cordon charges to deal with congestion and other time-varying local externalities;
(iii) location taxes and subsidies to affect the choice of residence outside the toll cordon;
(iv) fuel taxes to deal with non-localised environmental externalities; and
(v) differentiated vehicle taxes to affect environmental characteristics of cars other than directly related to fuel use.

Depending on the availability and feasibility of each of these measures, alternative packages could be constructed in order to intervene as efficiently as possible. For instance, if fuel taxes are not available, one could construct a similar package where, instead, highly differentiated vehicle taxes are used to promote fuel efficiency. If also an overall mobility reduction would be important for achieving the optimum, this new package may involve considerable welfare losses compared to the original one, unless other distance based charges can be introduced. In this way, one could use Table 8.1 in specific cases to construct check-lists for relevant externalities and behavioural dimensions and to assess tentatively the efficiency of the suggested policy packages and, ultimately, to construct an optimal package.

9 Theory vs. current approaches to pricing in urban transport

The previous sections have discussed the principles of marginal cost pricing when applied to urban transport, in relation to both conceptual or theoretical models and real-world model applications. The theory strongly supports marginal social cost pricing overall, and presents a number of clear policy suggestions.

However, the discussion reflected the view that transferring the marginal social cost pricing principle from economic theory to reality involves many practical complications. In particular, while the general goal of promoting efficiency is the sole criterion behind the notion of marginal social cost pricing, politicians and planners of practical pricing systems are constrained by the need to take account of a broad range of other objectives (e.g. equity, revenue generation, sustainability etc.), the prevailing technological mechanisms available for levying charges and the human issues relating to whether users perceive the system and react to it as intended. As a result, somewhat paradoxically, in those few cases where road pricing measures have been implemented in practice, the objectives which have motivated the move have not always been in line with the efficiency objectives.

This section contrasts the marginal cost pricing principles discussed previously with the current approaches to pricing in urban transport on a more practical level. Section 9.1 discusses policy objectives such as they have been interpreted in theory and modelling work alongside the way they are applied in practice. Section 9.2 discusses practical pricing measures both in relation to economic theory and wider objectives.

9.1 Objectives of pricing in practice

Despite its central role in theory, the notion of economic efficiency as a general level objective has had a rather less visible role in practical policy making and in deciding the practical pricing measures in urban transport. This is, in a way understandable, for this concept can easily seem too abstract and macroscopic to be adopted as a criterion at the operational level of the real-world transport system. Therefore, a more pragmatic set of policy objectives has tended to be adopted along the following lines (May, 1992):

(i) relieving road congestion;
(ii) protecting the local urban environment;
(iii) aiding more general urban planning goals;
(iv) promoting equity, both amongst travellers and between travellers and residents; and
(v) raising revenue.

This does not necessarily mean that a major contradiction exists between economic theory and practice: in many cases, the objectives adopted are practical proxies for those considered by the theory. Relieving road congestion is similar to economic efficiency, but with more pragmatic goals appropriate to the particular urban context, such as inducing modal shift. Protecting the local urban environment is again more pragmatic than covering external environmental costs, usually related to the desire to remove traffic from a sensitive area, such as a historic city centre. Aiding urban planning may mean, for example, improving accessibility and revitalising run-down parts of urban areas. Promoting equity, of course, means redistributing transport costs and benefits in a manner that is perceived to be more just. In particular, it may be used to improve the welfare of disadvantaged members of urban society and to address issues of *social exclusion*. It is also likely to be a major consideration to ensure the public and political acceptability of transport policy. Raising revenue usually means providing capital for investment in transport schemes which cannot easily be funded by other means. For example, a major issue which underlies the current moves towards urban road pricing in the UK is the widely held view that all forms of demand management designed to reduce private car use are likely to be ineffective and, potentially, very damaging to welfare and the economy in the short-run, unless very large sums of money are spent on alternatives.

For any individual urban transport pricing scheme, the relative priority placed on these practical objectives may vary. Congestion relief has always been the principal motive in Singapore and is the driving force behind the ongoing consideration of road pricing in Hong Kong. By contrast, modelling studies in Edinburgh and York have tended to focus on protecting the historic core. Work in both Edinburgh and Cambridge has envisaged pricing as a revenue raising tool to provide funds for investment in public transport as part of a package of transport measures. In Oslo, cordon tolls have been implemented to raise revenue for road building projects, which have been occurring alongside the regeneration of dockland areas.

In addition to general policy objectives, practical urban transport pricing systems also need to take account of more detailed design objectives, to ensure that they are socially acceptable, effective for producing desired modifications in travel behaviour and reasonably efficient to operate. One particular set of detailed design objectives was proposed by the Smeed Report in the UK (Ministry of Transport, 1964) for aiding the design of successful urban road user charging systems. They have subsequently been extended by Milne (1997) and, although their focus is on road traffic, the principles they embody could be applied more widely to any transport pricing system. They state that:

(i) charges should be closely related to the amount of use made of the transport system;

(ii) it should be possible to vary prices for different areas, times of day, week or year and types of user;

(iii) prices should be stable and readily ascertainable by users before they embark upon a journey;
(iv) the incidence of the system upon individual users should be accepted as fair;
(v) the method of charging should be simple for users to understand;
(vi) the system should be capable of being applied nationally and even internationally, to the whole traveller population;
(vii) payment in advance should be possible although credit facilities may also be permissible;
(viii) any equipment should possess a high degree of reliability;
(ix) the system should be reasonably free from the possibility of fraud and evasion, both deliberate and unintentional;
(x) the system should allow occasional users and visitors to be included rapidly at low cost;
(xi) any charging records generated should be designed both to allow accountability of the system and to protect privacy; and
(xii) the system should facilitate integration with other technologies, particularly those relating to traveller information and data collection.

Criteria (i) and (ii) are economic by nature, reflecting the basic principles of marginal cost pricing, while criteria (iii) and (v) suggest practical constraints on the manner in which charges are derived in order to have the desired behavioural impact and to be socially acceptable. The focus of criteria (iv), (ix) and (xi) is primarily on equity and resulting acceptability, addressing issues of accountability and confidentiality, while criterion (viii), about system reliability, can be related to both technical and acceptability issues. In contrast, criteria (vi), (x) and (xii) have implications mainly for technical and institutional arrangements.

These two sets of practical objectives and criteria cover a wide range of different aspects of charging systems and there is no need to attempt to prioritise them. Instead, it is useful to see that there can be trade-offs between the multitude of issues involved.

One point worthy of specific mention, related to criteria (i) and (ii), is the ineffectiveness of any approach which does not vary charges related to use of the transport system in time and space. An obvious example of this occurs allows where users pay a flat fee periodically for unlimited access. Arrangements of this nature are the accepted norm for much existing transport pricing because they require minimum cost and effort to be expended on administration and enforcement, because they are simple and convenient for users and because, as a result, they are most socially and politically acceptable. Unfortunately, they also weaken seriously the behavioural incentives of the charges by detaching payment from the day-to-day context of travel decisions and tending to encourage the notion that maximum usage equals maximum benefit (or *value for money*) once the payment has been made. This is not only a problem for traditional simple pricing systems based on low technology. The desired impacts of any new innovative pricing scheme may also be undermined to a considerable extent by offering periodic payments at discounted rates for frequent users.

9.2 Practical pricing measures in relation to theory and wider objectives

Various pricing measures are currently being applied in urban transport in Europe (and beyond). The different types of measures were categorised earlier (section 4.2.1) to cover: charges for using particular stretches of road; charges for parking; public transport fares and subsidies; taxes upon the purchase and licensing of vehicles and upon associated commercial services; and taxes on the purchase of fuel.

The following sections 9.2.1-9.2.5 discuss how these pricing measures are currently used.

9.2.1 Charges for using particular stretches of road

Direct *charges for using particular stretches of road* have so far been restricted to bridges and tunnels with particularly high investment and maintenance costs, inter-urban motorway tolls in five member states (France, Italy, Austria, Spain, Greece and Portugal), urban toll rings in three Norwegian cities (Oslo, Bergen and Trondheim) and travel time related charges for HGVs in Germany, Denmark and the BENELUX countries. Therefore, a strong precedent currently exists that roads are free at the point of use. Furthermore, as stated in section 9.1, the Norwegian toll systems have been designed solely for the collection of revenue. One-off journeys in Oslo face a single inbound charge rate throughout the day, while frequent users may pay a periodic fee which represents a discounted rate for daily commuters. This provides little incentive to modify travel behaviour towards a more efficient outcome in an urban context.

The different road user charging approaches that have been considered in road transport are, in increasing order of complexity:

(i) supplementary licensing, an additional periodic licence required to drive in particular areas of the network and/or at particular times;
(ii) point or cordon charging, where a fixed charge is levied to pass particular locations on the road network;
(iii) distance related charging, where a fixed charge is levied for the precise distance travelled;
(iv) time related charging, where a variable charge is levied dependent on the time spent travelling; and
(v) delay related charging, where a variable charge is levied dependent on the amount of time spent travelling in congestion.

Supplementary licences can discriminate in time and space, unlike existing annual vehicle licences, but are still tied to the concept of periodic payment, meaning that there is only a very coarse relationship between the fee paid and actual road usage. Point or cordon charges are sufficiently flexible to allow fees to vary by location, time of day and number of journeys, but are unable to discriminate by trip length and are not easily located to cater appropriately for complex urban travel patterns, resulting in particularly sharp impacts on the boundary. Distance related charges combine a fair measure of precision with ease of understanding and reasonable levels of predictability, but tend to encourage traffic to choose shortest routes, which is often undesirable in a congested urban situation.

Time and delay related charges may both provide closer proxies to the economic theory because the fees paid vary automatically with the travel conditions encountered. In the case of delay related charges, a fee would only be paid when a particular threshold level of congestion is exceeded. However, both systems have been criticised for their potential to encourage dangerous driving, a fear supported by recent research (Bonsall and Palmer, 1997), and there is evidence that fees could be difficult to predict, which, in turn, leads to concerns over equity and ease of understanding (Montecino, 1998). One response to these problems is the development of more complex and innovative cordon and distance related charging systems, designed to combine day-to-day predictability with longer term flexibility in time and space, to reflect more dynamic charging concepts.

These alternative approaches also differ in their technological requirements. Only supplementary licensing and cordon charging may be operated manually or through paper based systems, an option which is usually required to cater for infrequent users and visitors in current applications. These systems are also the only options which may easily be operated electronically through readable tag technology. Distance, time and delay related charges all require microwave technology with an in-vehicle smartcard payment system. Although these approaches have been shown to be reliable in technology trials (Hills, 1992), they were rejected as being beyond immediate feasibility for implementation during the London Congestion Charging Research Programme (The MVA Consultancy, 1995).

A recent UK Government White Paper has proposed enabling legislation to allow urban authorities to charge drivers directly for using the local road network (Department of the Environment, Transport and the Regions (DETR), 1998) and progress technology trials and pilot studies is ongoing. A number of Commission funded studies have focused on in-vehicle technology and electronic payment systems. These include PAMELA (Thorpe and Hills, 1991), ADEPT (Blythe and Hills, 1994) and GAUDI (Meland, 1995). Tag based electronic payment systems operate successfully in practice for the three Norwegian toll systems and for the

Dartford Tunnel in London. The long-standing urban charging system in Singapore has recently converted successfully to electronic operation.

9.2.2 Charges for parking

Parking charges are applied in most urban centres, but their scope is currently limited by difficulties in extending charging systems beyond small, clearly defined geographical areas and by inabilities to apply charges to private property. Regulation is frequently used instead of, and in addition to, pricing through parking restrictions. On-street parking may contribute significantly to congestion problems in urban road networks, but is more difficult to regulate and charge effectively than designated off-street parking lots. Periodic payment at preferential rates for frequent users is common.

Parking charges are already commonly used in urban centres to control overall demand for parking capacity and it is not unusual for charging structures to discriminate by location, duration and arrival/departure time. However, there are three main drawbacks which are often encountered:

(i) it is difficult to extend effective and comprehensive parking charging systems beyond limited geographical areas, due to the normal requirement of manual operation and enforcement;
(ii) it may be difficult to control capacity and fee levels where parking is provided in part by private operators, a common situation (e.g. in the UK); and
(iii) it may not be possible to levy effective charges for privately owned parking capacity, in particular workplace spaces in city centres.

The first issue may be overcome ultimately by the extension of electronic technology to cater for on-street parking and the second and third issues could clearly be remedied by urban local authorities acquiring greater legal powers. For example, in the UK there are currently proposals to allow urban local authorities to charge employers a *Workplace Parking Levy*, related to the volume of free parking capacity they operate. However, in the meantime the greatest potential for effective parking controls may lie in innovative approaches such as the *parking cordon* system proposed in Leeds under the TRANSPRICE project (Ghali et al, 1997). This is effectively a hybrid system, falling somewhere between a conventional parking charge and a cordon toll. It operates by charging a duration and/or entry and exit time related fee across a cordon around the city centre and would enable fees to vary by time of day, number of journeys and (to some extent) type of trip, while providing a neat solution to problems (ii) and (iii) identified above. However, it would not be able to discriminate by trip length and would require either a major manual/paper based ticketing system or fairly complex electronic payment technology.

Overall, the relationship between parking charges and marginal external costs is hindered by their limited geographical coverage. Also, applying an appropriate price to on-street parking is complicated by the fact that the activity of parking itself may generate external costs on the road network through capacity reduction.

9.2.3 Public transport fares and subsidies

Public transport services in Europe are currently provided by a mixture of public and private funding. Deregulation policies followed during the last decade have brought about significant increases in commercially operated services in a number of member states. However, the social benefits of maintaining comprehensive and coherent public transport networks and schedules have justified the retention of basic regulatory structures and the provision of public financial subsidies in many instances.

The Commission's Green Paper "*The citizens' network: fulfilling the potential of public passenger transport in Europe*" (1996) seeks new EU policies to promote passenger transport. The report states that, in addition to promotion of better connections between different modes and networks, public transportation should be a service open to all citizens. This is taken to mean: accessibility to vehicles and infrastructure; affordability in terms of fare levels; and availability in terms of coverage and frequency of services.

Public transport users pay fares which, typically, relate to the number of journeys made but which, in urban areas, often do not vary significantly by time or location of travel. Periodic payment at discounted rates is the norm for frequent users. Peak fares in urban areas may fall significantly below the marginal costs of service provision due to crossover effects from other policies designed to benefit lower income travellers and to make public transport a more attractive alternative to private car users.

Currently, most public transport fare payment systems are manual or paper based and subsidies are usually provided directly from government agencies to service providers. Progress towards electronic ticketing systems is allowing more discriminatory fare structures to become feasible and could even be used to provide subsidies directly to those users intended to benefit. Also, the trend towards integrated multimodal payment systems is allowing much greater transferability of tickets between modes and may even be used for modal credit transfers, to provide direct incentives designed to influence mode choice where the individual pricing mechanisms are considered unbalanced. In general, advanced electronic payment systems should make it possible for the user costs of public transport to be equated approximately with the marginal costs of service provision.

The use of new technologies for public transport information and integrated payment systems across different modes is gathering momentum, aided by research in Commission funded projects such as BUS PROJECT and TRANSPRICE. These developments are providing an important range of new policy tools for refining existing fare payment structures and for encouraging travellers to make more efficient decisions.

9.2.4 Taxes upon the purchase and licensing of vehicles and upon associated commercial services

All member states operate annual licence schemes for vehicle use and charge excise duty on vehicle purchases. The levels of these taxes vary significantly between states. Licences do not discriminate by the amount, timing or location of travel, although variations in charges related to vehicle type (e.g. to distinguish between motorcycles, cars and lorries) are the norm and those for cars are sometimes varied by engine size (e.g. a lower licence fee for the smallest engines, introduced recently in the UK). Some states operate lower tax levels for diesel vehicles, causing significant variations in the sums paid across the national vehicle fleet. Taxes may also be imposed on associated commercial services, such as on vehicle insurance premiums.

9.2.5 Taxes on the purchase of fuel

All member states impose taxes on the purchase of fuel. Levels are almost always uniform within national boundaries but vary significantly between states. Community law specifies a minimum level. Some states vary fuel tax levels between fuel types. In particular, lower taxes for the purchase of diesel are common, resulting in significant variations in the sums paid across the national vehicle fleet. An interesting exception to uniform national fuel excise duty levels has been tried in the Norwegian city of Tromso. In order to finance the development of the road infrastructure, including an underwater tunnel, an extra excise tax of (at present) NOK 0.65 (plus 23% value added tax) is levied per litre of fuel sold within the municipality of Tromso. The practicality of this taxing system relies crucially on the fact that the city limits are quite wide in relation to the population centre, most inhabitants having to drive at least 70 kilometres in order to buy fuel in a neighbouring municipality.

9.2.6 Applicability of different practical pricing measures

While parking charges and direct charging for road use are most naturally applied to suit the local context, vehicle taxes and fuel duty are typically only feasible at a uniform national level. Therefore, in addition to the ways in which different pricing measures may affect different dimensions of behaviour (Table 8.1), it is important to recognise that the spatial scale of measures may affect their applicability. For example, if significant differences exist in the levels of external costs generated by road travel in the urban, inter-urban and rural environments, then any attempts to raise uniform national charges to cater for the imbalance in urban areas may cause serious disbenefits elsewhere. In some cases, the excess of taxes collected from rural traffic over expenditure on infrastructure have been calculated to cover almost exactly one hundred percent of external costs. Therefore, the use of annual vehicle licences and fuel taxes as proxies for marginal cost pricing may be difficult to justify.

However, a number of innovative approaches to taxation which may have more potential are:

(i) relating the cost of vehicle licences to distance travelled;
(ii) creating more discriminating vehicle licensing systems by area, time of day or day of week;
(iii) relating general vehicle taxes more closely to characteristics of the vehicle, such as engine size; and
(iv) using taxes to redistribute funds between road users based on travel behaviour, through revenue neutral *feebate* systems.

In general, there are many different ways in which practical policies (in some cases existing ones) can be modified to become closer to marginal cost pricing.

10 Concluding comments

For application in real life in urban transport, the theoretical idea of marginal cost pricing needs to be operationalised. The term operationalisation means taking actions which aim to narrow the gap between academic theory and the practice of real-world policy making. Such a gap may exist in urban transport due to insufficient consideration given to:

(i) the definition of the relevant marginal cost concepts and items, including the correct level of (dis)aggregation;
(ii) the transferability of insights and results derived in highly simplified graphical and analytical settings to practical cases;
(iii) the broader socio-economic context in which marginal cost pricing would be implemented; and
(iv) the integration of the two major strands of research in the field, conceptual economic approaches and real-world modelling approaches.

This report has addressed these sets of issues, however, focusing most on issues (ii), (iii) and (iv).

Marginal cost pricing in urban transport may be applied by transport service operators, transport infrastructure operators and the government/regulator, with the aim to internalise marginal external costs by means of corrective charges and taxes. This report has focused on the role of the government/regulator and has referred to *marginal social cost pricing* in this context.

In relation to the broader socio-economic context of marginal social cost pricing (cf. case iii above), determined by various technological, institutional, legal and political constraints, this study has highlighted three important aspects or distinctions: the distinction between policy situations with different coverage (referred to as *settings*), the distinction between first-best and second-best situations and the need for policy packaging (rather than considering individual measures at a time).

All these aspects are extremely important for conceptual clarity and theoretical structure of the analysis. In particular, actual pricing rules suggested by the marginal social cost pricing principle can be very different depending on circumstances, such as the level of coverage and (dis)aggregation chosen and relating to the explicit inclusion of various constraints and policy instruments in the analysis.

Furthermore, making explicit distinctions between different settings and the first-best and second-best situations provides a natural framework for the integration of the two main approaches (cf. case iv above). In particular, the different settings and the different real-world model applications identified in this study correspond to each other. Evidently, this is no coincidence. Rather, the policy relevance of the settings has played a role in directing the development of the types of real-world model applications available.

The integration of real-world model applications and conceptual economic analyses has not previously received the attention it deserves. In particular, greater interaction of the two approaches is needed to allow real-world model applications to cater better for the goal of economic efficiency, side by side with the traditionally well represented broader objectives. This is a precondition for the use of a model application to analyse marginal social cost pricing.

This report has focused on operationalisation problems of marginal cost pricing. Deliverables 2a, 2b, 2c and 3 of the AFFORD project will investigate the implementation issues and policy implications, respectively. This means, specifically, analysis of institutional and legal barriers, estimation of the magnitudes of efficiency and equity impacts (and any trade-offs between them) and investigation of the acceptability issues. Besides providing a theoretical framework for such analyses, this report has also presented a structure for defining practical policy packages and an example set of concrete policy packages to be considered.

References

Agnew, C.E. (1973), The dynamic control of congestion – prone systems through pricing. Report No. 6, Stanford University Center for Interdisciplinary Research.

Algers, S., Bernauer, E, Boero, M., Breheret, L., Di Taranto, C., Dougherty, M., Fox, K. and Gabard, J-F. (1997)Review Of Micro-Simulation Models. Deliverable 3 of SMARTEST Project.

Anas, A., R. Arnott and K. Small (September 1998), Urban Spatial Structure, *Journal of Economic Literature*, Vol. XXXVI, pp. 1426-1464.

Arnott, R.A. (1979), Unpriced transport congestion, *Journal of Economic Theory* 21,294-316.

Arnott, R., A. de Palma and R. Lindsey (1990), Departure time and route choice for the morning commute, *Transportation Research,* 24B (3), 209-228.

Arnott, R., A. de Palma and R. Lindsey (1992), Route choice with heterogeneous drivers and group-specific congestion costs, *Regional Science and Urban Economics*, 22, 71-102.

Arnott, R., A. de Palma and R. Lindsey (1993), A structural model of peak-period congestion: a traffic bottleneck with elastic demand, *American Economic Review,* 83 (1), 161-179.

Arnott, R., A. de Palma and R. Lindsey (1995), Information and usage of free-access congestible facilities, *International Economic Review* (forthcoming).

Arnott, R., A. de Palma and R. Lindsey (1998), Recent developments in the bottleneck model. In: K.J. Button and E.T. Verhoef (1998) *Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility* Edward Elgar, Cheltenham.

Bates, J. (1991) Building A Strategic Model For Edinburgh. PTRC Summer Annual Meeting.

Battye, A., M.J. Smith and Y. Xiang (1998), The Cone Projection Method of Designing Controls for Transportation Networks, Paper presented at the IMA Conference on Mathematics in Transportation Planning, Cardiff and to appear in the Proceedings.

Baumol, W.J. and W.E. Oates (1988), *The Theory of Environmental Policy,* second edition. Cambridge University Press, Cambridge.

Blythe, P.T. and P.J. Hills (1994), Automatic debiting and electronic payment for transport - the ADEPT project - 1: overview, *Traffic Engineering and Control,* 35(2).

Braid, R.M. (1989), Uniform versus peak-load pricing of a bottleneck with elastic demand, *Journal of Urban Economics,* 26, 320-327.

Braid, R.M. (1996), Peak-load pricing of a transportation route with an unpriced substitute, *Journal of Urban Economics,* 40, 179-197.

Brueckner, J.K. (1999), Elements of an ideal urban policy. In: E. Niskanen and I.B. Voipio (eds) *Seminar on Urban Economic Policy 13.8.1998*, VATT Memorandum 41.

Calthrop, E. and S. Proost (September 1998), Deliverable 2: General Economic Principles of Pricing Transport Services, CAPRI report, Project Coordinator ITS University of Leeds. European Commission.

Chu, X. (1995), Endogenous trip scheduling: the Henderson approach reformulated and compared with the Vickrey approach, *Journal of Urban Economics,* 37, 324-343.

Chu, X. and K.A. Small (1996), Hypercongestion, Paper prepared for the meeting of the American Real Estate and Urban Economics Association, New Orleans, Jan. 1997.

Commission of the European Communities (1995), *Towards fair and efficient pricing in transport: policy options for internalising the external costs of transport in the European Union.* Green Paper, Brussels.

Commission of the European Communities (1996), *The citizens' network: fulfilling the potential of public passenger transport in Europe.* Green Paper, Brussels.

Commission of the European Communities (1998), *Fair payment for infrastructure use: a phased approach to a common transport infrastructure charging framework in the EU.* White Paper, Brussels.

Cornes, R. and T. Sandler (1996), *The Theory of Externalities, Public Goods and Club Goods*, Cambridge, USA.

De Borger, B. and S. Proost (1999), *Reforming Transport Pricing in the European Union: A Modelling Approach.* Edward Elgar, Cheltenham (forthcoming).

De Meza, D. and J.R. Gould (1987), Free access versus private property in a resource: income distributions compared, *Journal of Political Economy,* 95 (6), 1317-1325.

De Mooij, R.A. (1999) "The double dividend of an environmental tax reform" in J.C.J.M. van den Bergh (1999) *Handbook of Environmental and Resource Economics* Edward Elgar, Cheltenham, 293-306.

De Palma, A., F. Marchal and Y. Nestorov (1995), METROPOLIS: A Modular Architecture for Dynamic Traffic Simulation, Manuscript.

Department of the Environment, Transport and the Regions (DETR) (1998), HMSO, London.

Dijkstra, E.W. (1959), A note on two problems in connection with graphs, *Numerische Mathematik*, 1, 269-271.

DIW, INFRAS, Herry, NERA (1998), Infrastructure capital, maintenance and road damage costs for different heavy goods vehicles in the EU.

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

Edelson, N.M. (1972), Congestion tolls under monopoly, American Economic Review, 61,873-82.

Else, P.K. (1981), A reformulation of the theory of optimal congestion taxes, *Journal of Transport Economics and Policy*, 15, 217-232.

Else, P.K. (1982), A reformulation of the theory of optimal congestion taxes: a rejoinder, *Journal of Transport Economics and Policy,* 16, 299-304.

European Conference of Ministers of Transport ECMT (1998), *Efficient Transport for Europe. Policies for Internalisation of External Costs.*

External Costs of Transport (1998), R. Friedrich, P. Bickl and W. Kretwitt (eds), Research funded in part by The European Commission in the framework of the Non Nuclear Energy Programme JOULE III.

Evans, A.W. (1992), Road congestion pricing: when is it a good policy?, *Journal of Transport Economics and Policy*, 26, 213-243.

Evans, A.W. (1992), Road congestion: the diagrammatic analysis, *Journal of Political Economy*, 100, (1) 211-217.

Evans, A.W. (1993), Road congestion pricing: when is it a good policy?: a rejoinder, *Journal of Transport Economics and Policy*, 27, 99-105.

FATIMA (1998), Project FATIMA: Report on work package 20. Project funded by the European Commission under the Transport RTD Programme of the 4th Framework Programme.

Ghali, M.O. (1992), ASCOT: overview and user guide, Department of Mathematics, University of York.

Ghali, M.O. and M.J. Smith (1992a), Road pricing: a new model for assessing the many options, *Traffic Engineering and Control*, 33(3).

Ghali, M.O. and M.J. Smith (1992b), Approximation to optimal dynamic traffic assignment of peak period traffic to a congested city network. Proceedings of the Second International Capri Seminar on Urban Traffic Networks, Capri, Italy, July 5-8, 1992.

Ghali, M.O. and M.J. Smith (1993), Traffic Assignment, Traffic Control and Road Pricing. Proceedings of the 12th International Symposium on Transportation and Traffic Theory, Berkeley, California, (ed. C. Daganzo), Elsevier, New York, 147-170.

Ghali, M.O. and M.J. Smith (1995), A model for the dynamic system optimum traffic assignment problem, *Transportation Research*, 29B, 155-170.

Ghali, M.O., M. Pursula, D.S. Milne, M. Keranen, M. Daleno and M. Vougioukas (1997), Assessing the impact of integrated trans-modal urban transport pricing on modal split. Proceedings of *25th European Transport Forum*, PTRC, University of Brunel.

Glazer, A. and E. Niskanen (1992), Parking fees and congestion, *Regional Science and Urban Economics*, 22, 123-132.

Glazer, A. and E. Niskanen (1995), Imperfect competition, clubs, and two-part tariffs, VATT Discussion Papers 95.

Glazer, A., E. Niskanen and S. Scotchmer (1997), On the uses of club theory: Preface to the club theory symposium. In A. Glazer, E. Niskanen and S. Scotchmer (eds.), Special Issue: Club Theory Symposium, *Journal of Public Economics*, Vol. 65, No. 1, July 1997.

Harris, B. (1996), Land use models in transportation planning, *Review of Land Use Models and Recommended Model for DVRPC*, Delaware Valley Regional Planning Commission Publication No. 96008, Appendix B, B1-B14.

Henderson, J.V. (1974), Road congestion: a reconsideration of pricing theory, *Journal of Urban Economics*, 1, 346-365.

Henderson, J.V. (1981), The economics of staggered work hours, *Journal of Urban Economics*, 9, 349-364.

High Level Group on Infrastructure Charging (forthcoming), Recommended methodologies for the estimation of infrastructure, congestion, environmental and accident costs.

Hills, P.J. (1992). Information technology and the road user. Paper presented to *Lex Symposium on Motoring*, London.

Hills, P.J. (1993), Road congestion pricing: when is it a good policy?: a comment, *Journal of Transport Economics and Policy*, 27, 91-99.

Kimber, R.M. (1976), The capacity of some major/minor priority junctions, Department of the Environment, TRRL report 735, Transport and Road Research Laboratory, Crowthorne.

Kimber, R.M. (1980), The traffic capacity of roundabouts, Department of the Environment, Department of Transport TRRL Report LR942, Transport and Road Research Laboratory, Crowthorne.

Kimber, R.M. and E.M. Hollis (1979), Traffic queues and delays at road junctions, Department of Transport Report LR909, Transport and Road Research Laboratory, Crowthorne.

Laih, C.-H. (1994), Queuing at a bottleneck with single- and multi-step tolls, *Transportation Research*, 28A, 197-208.

Lévy-Lambert, H. (1968), Tarification des services à qualité variable: application aux péages de circulation, *Econometrica*, 36 (3-4), 564-574.

Link, H. and M. Maibach (1999), Calculating transport infrastructure costs. April 28, 1998.

Liu, R., Van Vliet, D. and Watling, D. (1995) DRACULA: Dynamic Route Assignment Combining User Learning and Microsimulation. PTRC Annual Summer Meeting, Seminar E.

Maddison, D., D. Pearce, O. Johansson, E. Calthrop, T. Litman and E. Verhoef (1996), *Blueprint 5: The True Costs of Road Transport* Earthscan, London.

Marchand, M. (1968), A note on optimal tolls in an imperfect environment, *Econometrica*, 36 (3-4), 575-581.

May, A.D. (1992), Road pricing: an international perspective, *Transportation* 19(4).

May, A.D., L. Rand, P.M. Timms and S. Toffolo (1997), OPTIMA: Optimisation of Policies for Transport Integration in Metropolitan Areas : a review of the results applied to nine European cities. Paper presented to the 25th European Transport Forum, 1-5 September 1997.

May, A.D. and Milne, D.S. (1999) The Effects Of Alternative Road Pricing Systems On Network Performance. Paper accepted for publication by Transportation Research A.

May, A.D., Bonsall, P.W. and Hills, P.J. (1998) The Impacts Of different Road User Charging Systems. Proceedings of 'Transportation Into The Next Millennium', Nanyang

Technical University, Singapore.

Mayeres, I. and S. Proost (1999) "Marginal tax reform, externalities and income distribution" Discussion Paper, CES – Katholieke Universiteit, Leuven.

McFadden, D.L. (1978), Modelling the Choice of Residential Location, *Spatial Interaction Theory and Planning Models*, Karlqvist et al (eds.), North-Holland, Amsterdam, 1987, 75-96.

Meland, S. (1995), Generalized and advanced urban debiting innovations: the GAUDI project (part 3), Trondheim toll ring, *Traffic Engineering and Control*, 36(3).

Milne, D.S. (1997), Modelling the network effects of urban road user charging. PhD. thesis, Institute for Transport Studies, University of Leeds.

Milne, D.S. and D. van Vliet (1993), Implementing Road User Charging in SATURN. Working Paper 410, Institute for Transport Studies, University of Leeds.

Ministry of Transport (1964), *Road pricing - the technical and economic possibilities*. HMSO, London.

Mohring, H. (1989), The role of fuel taxes in controlling congestion. In: *Transport Policy, Management and Technology Towards 2001: Proceedings of the Fifth World Conference on Transport Research* (Yokohama), 1, 243-257.

Montecino, R. (1998), An Analysis Of Charges Levied By Time-Variable Road Pricing Systems Using The DRACULA Microscopic Traffic Simulation Model. MSc dissertation, Institute for Transport Studies, University of Leeds.

MVA Consultancy (1995), The London Congestion Charging Research Programme - principal findings and final report. Prepared for the Government Office for London, HMSO, London.

MVA Systematica (1991) Introduction to TRIPS. The MVA Consultancy, London.

Nash, C.A. (1982), "A reformulation of the theory of optimal congestion taxes: a comment", *Journal of Transport Economics and Policy* 26, 295-299.

Nash, C.A., B. Mathews, P. Granero, W. Rothengatter (1999), Rail, other public transport and air transport pricing. CAPRI (Concerted Action on Transport Pricing Research Integration). Deliverable 6. Funded by the 4rth Framework RTD Programme. Institute for Transport Studies, University of Leeds, March 1999.

Newell, G.F. (1988), Traffic flow for the morning commute, *Transportation Science*, 22, 47-58.

Ochelen, S., S. Proost and K. van Dender (1998), Optimal pricing for urban road transport externalities, Mimeo, Centre for Economic Studies, KULeuven, Leuven.

Oi, W.Y. (1971), A Disneyland Dilemma: Two-part Tariffs for a Mickey Mouse Monopoly, *Quarterly Journal of Economics*, February, 77-96.

Ramjerdi, F. (1996), Road Pricing and Toll Financing, Helsfyr Hurtigtrykk, Oslo.

Ramjerdi, F. and Rand, L. (1992) The National Model System for Private Travel: report on phase II of model development. TOI Report 150, Oslo.

Rand, L and Rekdal, J. (1996) PC-Versjonen Av Den Nasjonale Transportmodellen: fase IV. Arbeidsdokument, TOI, Oslo

Rennings, K., A. Ricci, C. Cessa and S. Weinreich (1999), Valuation of Transport Externalities. CAPRI (Concerted Action on Transport Pricing Research Integration). Deliverable 3. Funded by the 4rth Framework RTD Programme. Institute for Transport Studies. University of Leeds, February 1999.

Rosen, H.S. (1995), *Public Finance*, 4th ed, Irwin, Chicago.

Scotchmer, S. (1985a), Profit-maximizing clubs, *Journal of Public Economics*, 27, 25-45.

Scotchmer, S. (1985b), Two-tier pricing of shared facilities in a free-entry equilibrium, *Rand Journal of Economics*, 16(4), 456-472.

Scotchmer, S. (1994), Public goods and the invisible hand, in *Modern Public Finance*, J.M. Quigley and E. Smolensky (eds), Harvard University Press.

Shepherd, S.P., G. Emberger, K. Johansen and T. Jarvi-Nykanen (1997), OPTIMA: Optimisation of Policies for Transport Integration in Metropolitan Areas: a review of the method applied to nine European cities. Paper presented to the 25th European Transport Forum, 1-5 September 1997.

Small, K.A. (1992), Urban Transportation Economics, Harwood Academic Publishers, Chur.

Smith, M.J., M.O. Ghali and D.S. Milne (1998), Real-world model applications within the AFFORD project. ITS Institute for Transport Studies, University of Leeds, unpublished.

Snow, A. and R.S. Warren Jr. (1996), The marginal welfare cost of public funds: Theory and estimates, *Journal of Public Economics*, 61, 289-305.

Spulber, D.F. (1985), Effluent regulation and long-run optimality, *Journal of Environmental Economics and Management*, 12, 103-116.

Stiglitz, J.E. (1988), Economics of the Public Sector, 2nd edition, Norton, New York.

Sullivan, A.M. (1983), Second-best policies for congestion externalities, *Journal of Urban Economics*, 14, 105-123.

Taylor, N.B. (1990) CONTRAM 5: An Enhanced Traffic Assignment Model. Transport and Road Research Laboratory, Research Report 241.

Thorpe, N. and P.J. Hills (1991), The DRIVE project PAMELA - 2: the scope of automated pricing systems, *Traffic Engineering and Control*, 32(7/8).

Tirole J. (1990), *The Theory of Industrial Organization*, The MIT Press, Cambridge, Massachusetts; London, England.

TRENEN II STRAN (1999), Final Report. Project funded by the European Commission under the Transport RTD Programme of the 4th Framework Programme, March 1999.

Van Dender, K. (1998), Urban Case Study Report: Brussels 2005, TRENEN Deliverable D13. Centre for Economic Studies, KULeuven, Leuven.

Van Vliet (1982), SATURN - a modern assignment model, *Traffic Engineering and Control*, 23(12).

Varian, H.R. (1992), *Microeconomic Analysis*, 3rd edition. Norton, New York, London.

Watling, D. (1994) Urban Traffic Network Models and Dynamic Driver Information Systems. Transport Reviews 14(3), 219-246.

Wegener, M. (1998), *Applied Models of Urban Land Use, Transport and Environment*, in Kim, T.J. and Lundqvist, L. (eds) "Network Infrastructure and the Urban Environment - Recent Advances in Land-Use/Transportation Modelling", Springer Verlag 1998.

Verhoef, E.T. (1996), *The Economics of Regulating Road Transport.* Edward Elgar, Cheltenham.

Verhoef, E.T. (1998a), Time, speeds flows and densities in static models of road traffic congestion and congestion pricing, *Regional Science and Urban Economics* (forthcoming).

Verhoef, E.T. (1998b), An integrated dynamic model of road traffic congestion based on simple car-following theory, Discussion paper TI 98-030/3, Tinbergen Institute, Amsterdam-Rotterdam.

Verhoef, E.T. (1998c), Second-best congestion pricing in general static transportation networks with elastic demands Discussion paper TI 98-***/3, Tinbergen Institute, Amsterdam-Rotterdam.

Verhoef, E.T., J.C.J.M. van den Bergh and K.J. Button (1997), Transport, spatial economy and the global environment, *Environment and Planning*, 29B, 1195-1213.

Verhoef, E.T., P. Nijkamp and P. Rietveld (1995), Second-best regulation of road transport externalities, *Journal of Transport Economics and Policy*, 29 (2), 147-167.

Verhoef, E.T., P. Nijkamp and P. Rietveld (1996), Second-best congestion pricing: the case of an untolled alternative, *Journal of Urban Economics*, 40 (3), 279-302.

Verhoef, E.T., R.H.M. Emmerink, P. Nijkamp and P. Rietveld (1996), Information provision, flat- and fine congestion tolling and the efficiency of road usage, *Regional Science and Urban Economics*, 26, 505-529.

Vickrey, W.S. (1969), Congestion theory and transport investment, *American Economic Review*, 59, (Papers and Proceedings) 251-260.

Williams, H.C.W.L. (1977), On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit, *Environment and Planning*, A9, 258-344.

Wilson, J.D. (1983), Optimal road capacity in the presence of unpriced congestion, *Journal of Urban Economics*, 13, 337-357.

Yang, H. and H.J. Huang (1998), Principle of marginal-cost pricing: how does it work in a general road network?, *Transportation Research*, 32A (1), 45-54.

VATT-TUTKIMUKSIA -SARJASSA ILMESTYNEITÄ

PUBLISHED VATT-RESEARCH REPORTS

43. Lehtinen Teemu: The Distribution and Redistribution of Income in Finland 1990-1993. Helsinki 1998.
44. Rantala Juha: Työvoimapolitiikan rooli ja työttömien työllistyminen. Helsinki 1998.
45. Laurila Hannu: Suomalaisen kaupunkipolitiikan taloudelliset lähtökohdat. Helsinki 1998.
46. Tuomala Juha: Pitkäaikaistyöttömyys ja työttömien riski syrjäytyä avoimilta työmarkkinoilta. Helsinki 1998.
47. Tossavainen Pekka: Panosverot ja toimialoittainen työllisyys. Helsinki 1998.
48. Holm Pasi – Kiander Jaakko – Tuomala Juha – Valppu Pirkko: Työttömyysvakuutusmaksujen työttömyysriskin mukainen porrastus ja omavastuu. Helsinki 1998.
49. Kari Seppo – Kröger Outi – Rauhanen Timo: Henkilöyhtiöiden verotuksen investointi- ja työllistämiskannustimet. Helsinki 1998.
50. Kajanoja Jouko: Lasten päivähoito investointina. Helsinki 1999.
51. Kari Seppo: Dynamic Behaviour of the Firm Under Dual Income Taxation. Helsinki 1999.
52. Holm Pasi – Sinko Pekka – Tossavainen Pekka: Työpaikkojen syntyminen ja päättyminen ja rakenteellinen työttömyys. Helsinki 1999.
53. Mäkelä Pekka (toim.): EU:n kauppapolitiikkaa itälaajenemisen kynnyksellä. Helsinki 1999.
54. Sinko Pekka: Taxation, Employment and the Environment – General Equilibrium Analysis with Unionised Labour Markets. Helsinki 1999.
55. Rantala Anssi: Finanssikriisit, yritysten nettovarallisuus ja makrotaloudellinen vakaus. Helsinki 1999.
56. Kyyrä Tomi: Post-Unemployment Wages and Economic Incentives to Exit from Unemployment. Helsinki 1999.
57. Korkeamäki Ossi: Yksityisen ja julkisen sektorin palkkoihin vaikuttavat tekijät. Ekonometrinen tutkimus 1987 - 1994. Helsinki 1999.
58. Venetoklis Takis: Process Evaluation of Business Subsidies in Finland. A Quantitative Approach. Helsinki 1999.
59. Kuusi Osmo: Expertise in the Future Use of Generic Technologies – Epistemic and Methodological Considerations Concerning Delphi Studies. Helsinki 1999.
60. Hakola Tuulia: Race for Retirement. Helsinki 1999.

61. Korkeamäki Ossi: Valtion palkat yleisiin työmarkkinoihin verrattuna: vuodet 1989 - 1997. Helsinki 2000.

62. Uusitalo Roope: Paikallinen sopiminen ja yritysten työvoiman kysyntä. Helsinki 2000.