Harald M. Hjelle

A Foundation of
Road User Charges

Thesis submitted to

The Norwegian University of Science and Technology (NTNU),

Faculty of Engineering, Science and Technology,

Department of Civil and Transport Engineering

for the Dr. Ing. Degree

Thesis no. 2003-49

ISBN 82-471-5600-8

CONTENTS

IV

LIST OF FIGURES

X

PREFACE

ACKNOWLEDGEMENTS

Although writing a thesis like this is to a large extent a solo project, there are always some persons that have contributed to the final result in different ways.

Let me start with my appointed advisors at the Norwegian University of Science and Technology (NTNU). Professor Tore Sager and professor Ivar Horvli, now both representing the Department of Civil and Transport Engineering, have been my supervisors. Although I have not had my daily residence in Trondheim, and thus have not knocked at their doors every week, it is always reassuring to have their experience and skills to rely on when comments and second opinions are needed.

Apart from the courses taken at NTNU, and the guidance received from my mentors there, I want to thank my colleagues at Molde University College. In addition to the apparent usefulness of discussing matters directly related to my doctoral work, good colleagues are also important for more general support when things look awfully complicated. My deepest thanks to all of you for providing a challenging and supportive environment! Special thanks go to Odd I. Larsen, Svein Bråthen and Øyvind Sunde, whom I have troubled most with my frustrations and problems over the last years. I would also like to give special thanks to our very skilful librarians. Sometimes I have just left a list of articles and reports that I would like to read before leaving for my holidays, and when I come back it is all there in my mailbox! Having access to such professional help is invaluable in this world of "information overflow". Among all the librarians helping me over these years, I have bothered Vera Høstmark most. Thank you for always being so service-minded!

BACKGROUND AND MOTIVATION

The idea of pricing transport according to marginal social costs is not a new one. The basic principles have been known for years, and the policy propositions have been many, especially over the last three or four decades. However, we are still waiting for a full-scale real implementation of such a pricing policy. Recent policy papers indicate that the need for regulating traffic by pricing properly is now becoming so importunate that politicians actually are ready to put the theories into practice. The policy papers of the European Union (EU) have set a time-table that could bring about marginal cost pricing in many transport sectors within a few years. In other parts of the world there is not such a high focus on implementing marginal cost pricing in general, but congestion pricing is certainly an option being considered in many major cities of the world.

A necessary prerequisite for implementing marginal cost pricing is well founded estimates of marginal costs. The EU has therefore commissioned a number of research programmes trying to establish the necessary empirical background for implementing the new pricing policy. These programmes have also explored related subject areas such as public attitudes to different pricing alternatives, theoretical models for second best regimes, possible behavioural responses to pricing, technical solutions for collecting fees etc.

In Norway the Government has not given any binding commitment to implementing marginal cost pricing schemes. However, following a recent modification of the legislation, local authorities may now implement marginal cost pricing if they so wish.

Policy-documents from the 1970s onwards have always had some focus on the marginal cost pricing principle, and marginal road user costs have been calculated since the mid 1970s. These calculations have been updated on several occasions, but have always been done with application of the same basic framework, the

TØI[1]-model.

Norway is not a member of the EU, still our membership in the European Economic Area (EEA) links our policy to EU policies in many fields. If the EU carries out its intentions with respect to promoting marginal cost pricing principles, Norway will probably follow[2]. Norway has to some degree been active in the EU-programmes, seeking to establish the necessary foundation for implementing such a pricing policy, but has *not* been present in many of the major research areas. If Norway is to move towards marginal cost pricing there will be a need for new evidence on the magnitude of marginal road user costs under the prevailing conditions on the Norwegian road network. This thesis critically reviews the foundation of the established Norwegian calculation procedure, and presents a revised model, which is based on recent international and national research results. A first attempt to establish a model that could be used for estimating marginal road wear from a bottom-up approach is also made here.

OUTLINE

In Chapter 1 I give a short introduction to the basic theory behind the principle of marginal cost pricing. I also try to establish some more practical and political considerations that, along with the theoretical models, form the ideal foundation for road user charges.

In Chapter 2 the principles, methods, and empirical evidence related to the estimation of marginal road user costs are presented and critically discussed. Most of the tax-relevant marginal costs are intangibles without a proper market value. This means that we have to apply some form of shadow prices or alternative

---

[1]   The model does not actually carry a specific name, but it was developed by the Institute of Transport Economics (TØI) in Oslo, and TØI has also been responsible for the subsequent updates of the model in the 1980s and 1990s.

[2]   Actually tax and charging regimes are not a part of the EEA-agreement, but the responsibility for creating a fair competitive environment is. This means that Norway probably will have to adjust to EU regulations with respect to the taxation of commercial vehicles at least.

values when estimating these costs.

The major focus in this thesis with respect to the different cost components relevant to taxation, is put on the marginal road wear. The purpose of Chapter 3 is to review some different approaches to estimating road wear, and to give concrete examples of such studies, in order to establish the background for the empirical study presented in Chapter 4. Here the so-called FAMAROW-model is presented, which is a first effort to estimate marginal road wear on in-service roads in Norway. The data is partly based on Weigh-In-Motion data collected from the Automatic Traffic Control units on the Norwegian road network. This information about factual traffic loads is combined with measurements of the development of rutting and roughness recorded in the Norwegian Road Data Bank.

Chapter 5 comprises an effort to establish a complete calculation model for external marginal costs (CATERU). The major part of the Norwegian road network is located in areas that are not densely populated, and outside the major cities there is very little congestion on the network. Therefore the focus in this thesis is put on the inter-urban case, and not on congestion costs, thus forming the foundation for a basic km-based charge working along with additional congestion charges in special areas. Most of the estimates of marginal external costs are flawed by a high degree of uncertainty. Therefore I have put some effort into illustrating the magnitude of this uncertainty by presenting alternative scenarios and different calculation prices (a sensitivity analysis).

The foundation of a new pricing regime is laid in Chapter 5. In Chapter 6 conclusions are drawn in the form of a proposed outline of a new pricing regime for the Norwegian road network.

Molde University College, February 2003

Harald M. Hjelle

1 THEORETICAL FOUNDATION OF USER CHARGES

1.1 THE POINT OF DEPARTURE: ECONOMIC WELFARE THEORY

1.1.1 Social efficiency

The point of departure when establishing the theoretical rationale for user charges is Economic Welfare Theory. This branch of economics focuses on the maximation of societies' welfare. The normative welfare theory has mainly developed from Arthur Cecil Pigou's *Economics of Welfare* published in 1919. There the basic principles related to achieving social efficiency from optimal allocation and distribution of resources, were introduced.

When defining social efficiency, Pigou based his work on notions established by Vilfredo Pareto a few decades earlier, later known as the Pareto Optimum and Pareto Improvement. Social efficiency relates to achieving both *distributional efficiency, allocational efficiency* and *production efficiency*. Prices are important instruments in achieving these efficiencies, as they are the main carrier of information about costs and benefits between actors in the economy.

If we assume that welfare is well represented by the consumers' willingness to pay, the best way to make sure a limited supply of a commodity is made accessible to those who benefit most from it, is to raise prices until the total demand equals total available quantity. This way the consumers with the highest willingness to pay receive the good. This is what is called *distributional efficiency*.

If prices equal the marginal social costs of producing a good, consumers face the correct signals about the cost to society of producing an extra unit of the good.

1

Consumers will then only buy goods that are worthwhile producing from societies' point of view. This is what is meant by achieving *allocational efficiency*.

Production efficiency is achieved when producers have an incentive to produce goods at the lowest possible costs, and the competitive environment ensures that only the most efficient producers stay in business.

These prerequisites for social efficiency can be expressed much more stringently and completely by the use of a mathematical model formulation. Such derivations of the necessary conditions for achieving social efficiency can be found in most standard textbooks dealing with economic welfare theory (e.g. (Bohm 1987)

1.1.2   External effects

One of the prerequisites for achieving social efficiency is that prices should reflect social marginal costs, thereby passing on the correct information about the true costs of providing the commodity in question. In the presence of *market distortions* (or *imperfections*), market prices will *not* reflect the true social costs of providing the good. These distortions arise when there are imperfections in the markets, which ideally should be characterised by *op. cit.*:

    a. An economy with perfect competition in all markets, and these markets being in equilibrium (i.e. situations where demand equals supply in every market)

    b. Every equilibrium position is socially efficient, or Pareto-optimal[3]

    c. Every conceivable Pareto-optimal situation (i.e. income distribution) corresponds to such an equilibrium situation

    d. There is perfect information

---

[3]   Pareto optimality is defined as a situation where it is not possible to reallocate resources in such a way that at least one person is better off, without anyone becoming worse off.

2

Examples of market *imperfections* are

- The existence of *external effects*
- The existence of *public goods*
- The existence of *decreasing production costs* (economies of scale)
- *Imperfect competition* / Existence of market power
- Market imbalances due to *regulations*
- *Distortive taxes*

Although all these examples could have some relevance in my setting, the main focus will be on the case of *external effects* in this thesis[4]. Whenever the actions of the individual producer or consumer affects a third part as well, then the decisions made by this individual may no longer be optimal from society's point of view. The individual actor will generally not include in his welfare function the costs or benefits imposed on other persons.

A more precise definition of external effects is (Verhoef 1994):

> *An external effect exists when an actor's (the receptor's) utility (or profit) function contains a real variable whose actual value depends on the behaviour of another actor (the supplier), who does not take these effects of his behaviour into account in his decision making process.*

Such side-effects are very prominent in transportation, not least related to road use. In Figure 1-1 a typology of the different external costs of road transport can be found. My focus will be on the effects resulting from actual transport activities,

---

[4]   Indeed, the literature about road user charges carry examples of *all* the mentioned market imperfections, e.g. uncongested roads could be regarded as public goods, there are probably economies of scale in the provision of roads, hence roads could also become natural monopolies. Moreover, some roads have regulated access, and tolling roads for financial purposes is certainly an example of distortive taxes being applied.

i.e. leaving out the externalities related to the mere existence of vehicles and road infrastructure.
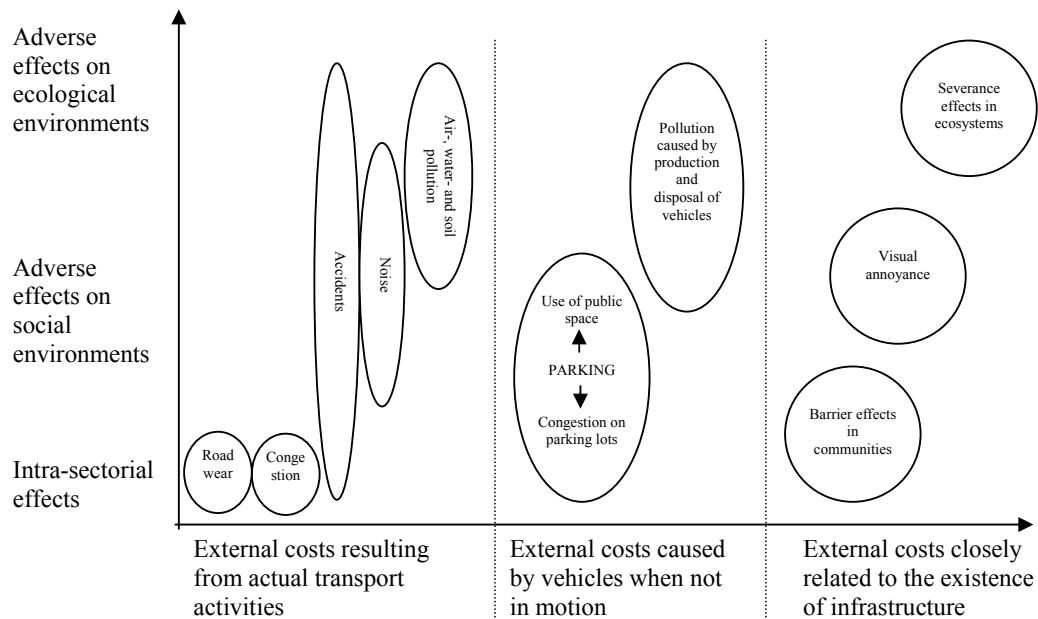
Figure 1-1   A typology of external costs of road transport (mainly based on Verhoef 1994)

Based on this limitation, the four main categories of externalities related to road *use* are:

- *Congestion costs* (i.e. external time-costs). Whenever a road user enters a congested road network, this also delays the other road users, and the sum of the delays for all the other road users constitute the external congestion costs.

- *Accident costs.* If an additional kilometre driven by a road user also increases the accident risk for other road users, then this represents an externality.

4

- *Environmental costs.* Environmental hazards are imposed on other road users, residents, and society in general. These costs are external to the individual road user. Such marginal environmental costs arise from air emissions (e.g. CO_2, NO_x, CO, SO_2 and particles) and noise. The costs may be both related to health effects of humans and animals, and to more direct economic impacts (e.g. agricultural productivity, need for cleaning streets, need for noise insulation etc.)

- *Infrastructure costs.* These costs mainly arise from two "sources". Firstly, increased maintenance costs related to traffic volume. These costs are external to the individual road user, as they are generally borne by the road authorities. The second category of external infrastructure costs arises from the increased road wear stemming from the individual's extra road use. This extra road wear may cause increased vehicle operating costs (and possibly comfort and safety costs) for other road users.

These external effects represent social inefficiencies if no market intervention takes place. Generally, market imperfections can be corrected by technical measures (regulations) or economic policy instruments. In the following section I focus on the use of Pigouvian taxes for internalising external effects.

1.1.3   Pigouvian taxes

Pigouvian taxes represent an economic policy instrument that may be used for correcting the market imperfection of external effects. The basic idea is to make the actors in the economy aware of the true social costs of providing a good by adding a tax that reflects the magnitude of the externality.

This is illustrated in Figure 1-2. Here the external effect is represented by the difference between the marginal social cost and the corresponding marginal private cost. In the case of road use, this might e.g. be external environmental costs, or external time costs. In the figure, the marginal externality is assumed to

increase with higher traffic volumes. The social optimum is where the marginal willingness to pay (represented by the demand function) equals the marginal social cost (including the externality). This is true at traffic level X_s. Where this situation is left to the market, the individual road user would only consider the private marginal costs (petrol, own time etc.), and the realised traffic volume would be too high (X_p). This would incur a social efficiency loss as illustrated in the figure (area C). Introducing a Pigouvian tax, i.e. a tax equal to the marginal externality at the optimal traffic level, would yield the wanted level of traffic, thus forcing the road users to consider not only their private costs, but also the external costs related to their road use, when deciding on their demand for road services. The revenue from the tax is the sum of areas A and B in the figure.

Pigouvian taxes have been advocated as a very attractive form of taxation because they have been said to produce a *double dividend*. Firstly, there is an efficiency gain from correcting the prices, as illustrated in the figure. Secondly, the tax income generated from these taxes may enable a reduction in other, distortive, taxes (e.g. income taxes), and thereby reduce the efficiency loss related to non-optimal prices for these taxed commodities (e.g. labour).

One of the central problem areas related to Pigouvian taxes is that in order to dimension the tax optimally, one would need very detailed information about the magnitude of the externality in question. Once again referring to Figure 1-2, in this situation it is not only sufficient to know the size of the externality at the current traffic level X_p (before taxation); one needs to be able to predict the magnitude of the externality at the optimal traffic level. Thus, to achieve this, one must have information about both the demand function and true marginal social costs. An additional problem is, as will be seen later in this thesis, that many of the commodities in question do not have readily available prices that could be used for calculating the optimal tax.

6
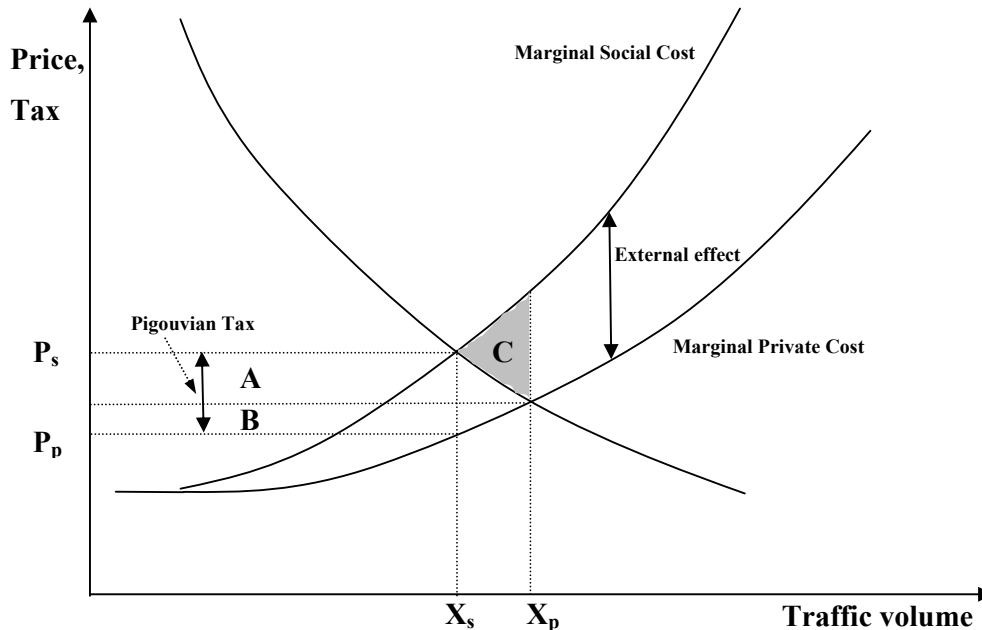
Figure 1-2   Illustration of the principle of Pigouvian taxes

One important practical issue with respect to the application of Pigouvian taxes is to consider the *transaction costs* related to making use of this instrument. High transaction costs may jeopardize the allocative efficiency gains (Coase 1960). Such transaction costs may arise from different sources. Firstly, I have already commented upon the major challenges related to merely estimating the level of the appropriate tax. This is not only costly, but may in many situations be impossible to do with the required precision level. Secondly, there may be practical problems related to finding a suitable medium for taxation. E.g., in the case of external environmental costs, one should ideally tax the emissions themselves, but as this is technically very difficult, one is forced to choose another medium which is highly correlated, but still less efficient (e.g. petrol). Thirdly, there are always administrative costs connected to the administration and enforcement of any taxing regime. Consequently, the sum of these costs may very well exceed the efficiency gains from internalising the externality, and then the rationale for introducing the tax is not present.

The implementation of a Pigovian tax regime may also be found to have undesirable distributional consequences. However, one might argue that the policy goals with respect to income distribution should be solved trough other more direct instruments.

Although the possibilities of obtaining the "double dividend" may seem attractive, these are factors that may moderate the attractiveness of an extensive use of Pigouvian taxes. I will revisit some of these items briefly reviewed here at the end of the thesis, when assessing the feasibility of charging regimes and discussing issues like political and public acceptability of proposed schemes.

1.2 ECONOMIC AND POLITICAL CONSIDERATIONS FOR THE SHAPING OF ROAD USER CHARGES

1.2.1 The basic principle: Price equal to short term marginal cost

The basic pricing principle is that optimal prices should equal marginal social cost. The term "marginal social cost" needs a closer look. A frequently used distinction when discussing transport costs is between *short run* marginal costs, and *long run* marginal costs. The difference between these notions is that one assumes capacity to be fixed in the short run, but adjustable in the long run.

When capacity is fixed, the costs of providing capacity are not relevant for pricing, and prices should equal short run marginal costs. This means that if peak demand is high enough to yield a willingness to pay that exceeds the marginal social cost at the capacity limit, prices should increase until supply and demand balance (*peak load pricing)*.

In the long run it is also possible to expand or reduce capacity, hence the long run marginal costs should also comprise capacity costs. One should invest in extra

8

capacity as long as the marginal willingness to pay exceeds the long run marginal costs (also referred to as the *optimal investment rule*). When capacity is optimally adjusted like this, short run and long run marginal costs are equal.

When dealing with transport infrastructure it is most often impossible to adjust capacity unit by unit, there is a lumpiness[5] that makes capacity increase stepwise. In addition to this, transport infrastructure capacity cannot be altered to match short time fluctuations in demand (e.g. rush-hours vs. non-rush hours). This means that the typical situation tends to be that capacity is *not* optimally adjusted at any given time and place, hence long term and short term marginal costs will also be different in most cases. This is why pricing at long term marginal cost will generally be wrong. The pricing policy should ensure we make the best use of existing resources, and this is done by charging according to the short run marginal social cost (Walters 1968).

Some writers have advocated pricing at long term marginal cost to achieve full cost recovery, but it is important to understand that neither is there a guarantee that such a pricing policy will result in full cost recovery (due to the lumpiness and stickiness of road capacity), nor does such a pricing policy necessarily represent the most efficient way of financing roads (Walters 1968). Indeed, a lot of current political pricing doctrines refer to "development cost", "average cost" or "full cost" rather than short run marginal social cost (Mayeres et al. 2001). These are all concepts that are closer to the notion of long term marginal costs than the short run marginal social cost. In cases where economies of scale are not important, when externalities are not important, and when the general economic environment are not too far from the optimum, all these pricing options are close to each other, hence the efficiency loss from applying such rules would be limited in such cases. However, in most cases there will be a significant efficiency loss from applying these alternative pricing rules. The political hesitation related to

---

5   E.g. it is not possible to build a road with 1.5 lanes, or to increase capacity by 0.3 lanes.

implementing short run marginal social cost pricing mainly revolve around uncertainty of determination, efficiency and incentive considerations and equity concerns. All these problem areas could, however, be addressed by modifications of the pricing rule based on developments in economic theory (see e.g. *op. cit.* section 3).

Because road transport infrastructure is generally subject to increasing returns to scale[6], pricing at short run marginal cost will not cover total infrastructure costs. In most cases there will be a financial deficit in the provision of roads when applying such a pricing principle on road networks with no, or small, congestion problems (Nash and Matthews 2001). In highly congested networks the optimal congestion charge could very well be high enough to yield a financial surplus in the provision of roads. Indeed, it might be possible that congestion charging on urban networks could cross-subsidise rural networks in many European countries (Roy 2000). However, this may not be the case for Norwegian (or other Nordic) road networks where a very large proportion of the roads carries very low traffic volumes.

From a political perspective it may be desirable to make the road sector self-financed by ensuring full cost recovery. This means that one would have to charge taxes higher than the social marginal costs in most cases. There will inevitably be an efficiency loss related to such a policy, but theories are developed for how such a pricing regime should be designed in order to minimise these losses (Ramsey-pricing[7]). This is an example of *second best pricing regimes*, which is briefly discussed in the following section.

---

[6]   The economies of scale in the provision of road capacity, is generally accepted in most cases, at least in the interurban or rural case. When prices of land become very high (as might be the case in densely populated areas), diseconomies of scale may prevail (see e.g. Walters 1968).

[7]   Ramsey prices are Pareto-optimal prices which achieve a required level of profits. The rule says that the excess of price over marginal cost should be higher for the commodities that have the lowest price elasticities. (Baumol, W. J., and Bradford, D. F. 1970)

10

1.2.2   Second best pricing regimes

When approaching a practical and political viable pricing regime, it soon becomes clear that one or more of the requirements for the perfect market conditions will be violated in almost all real life situations. This seems to be the rule rather than the exception (Verhoef 2001). It is a well known result from basic welfare theory that if prices deviate from marginal costs in one sector of the economy it will be optimal to deviate from marginal costs in all other sectors as well. The importance of putting focus on this problem is of course dependent on the level of interaction between the sectors in question. If the cross price elasticities are high (in absolute value), the efficiency loss of not taking this into consideration when setting prices, becomes considerable. This means for road transport that if pricing is not optimal in related transport sectors (e.g. rail), the prices should not equal marginal costs in road transport either. Moreover, this may very well also become a problem if governments choose to implement marginal cost pricing in a step-wise manner, e.g. by only including motorways first. Then the efficiency gain from introducing marginal costs pricing in one part of the road network may be jeopardized because prices are not optimal in another part of the network.

The awareness of such problems has triggered a lot of research into how these deviations from the ideal conditions would affect the optimal pricing rules. Some of the latest contributions to this literature are provided by Dr. Erik T. Verhoef (op.cit. and Verhoef 2000). The list of possible imperfections is long, each element leading to a specific pricing rule. I will confine myself to give only a very brief example on how such considerations could affect the general pricing rule, based on Verhoef's work. Let us have a look at the case of second-best tolling with non-optimal pricing on other road routes, provided by Verhoef. The first-best congestion-pricing rule could be expressed as:

Equation 1-1 $r_i = N_i \dfrac{dC_i^p(N_i)}{dN_i}$

Here r_i is the optimal congestion charge on route i

       N_i is the number of users of route i

       C_i^p is the external average user cost of driving route i

If the authorities only want, or can, toll one of the routes, this will affect the optimal congestion charge on the tolled route, r_T in the following way:

Equation 1-2 $$r_T = N_T \frac{dC_T^p(N_T)}{dN_T} - N_U \frac{dC_U^p(N_U)}{dN_U} \cdot \frac{-\dfrac{dD(N)}{dN}}{\dfrac{dC_U^p(N_U)}{dN_U} - \dfrac{dD(N)}{dN}}$$

Here r_T is the optimal congestion charge on the tolled route

       N_T is the number of users of the tolled route

       N_U is the number of users of the un-tolled route

       C_T^p is the external average user cost of driving the tolled route

       C_U^p is the external average user cost of driving the un-tolled route

       $D(N)$ is the traffic demand function

The point here is not to give a thorough discussion of the optimal pricing rule in question, but only to point out that this single example of a second-best pricing rule puts *much greater demands on the information necessary* to implement the pricing strategy. In this case it means that one would need information about the demand structure for the whole road system instead of just the information about the external congestion costs on the route in question.

Other second best pricing problems are:

- *Price distortions on competing modes*. In this case the second-best pricing rule should reflect the distortions occurring in the other transport modes.

Again this complicates the pricing rule, and realistic specifications may very well lead to solutions that are no longer analytically tractable.

- *Price distortions elsewhere in the economy.* The first best solution in such a case would of course be to introduce corrective taxes in the sectors where the distortions occur. If this is not possible, the second-best solutions would be to deviate from the first-best pricing rule in the transport sector. The severity of this, and the previous problem is related to the level of interaction between the affected sectors or modes (e.g. expressed by the cross-price elasticities).

- *Distortions due to government budget restraints.* The optimal tax may be affected by a positive shadow price on public funds. However, *op. cit.* shows that the adjustment of the optimal tax relative to the first best solutions may both be upwards and downwards, depending on the elasticity of demand.

More examples of concrete second-best pricing rules, and their corresponding demands on information about demand and cost structures in correlated economic sectors, could be found in Verhoef (2000).

In this thesis the focus is put on a first best pricing approach. This does not mean that the author perceives the Norwegian economy to be free from distortions that may call for modified pricing rules. However, analyzing the potential presence of such market distortions is beyond the scope of this thesis.

I am aware of the fact that the development in political pricing doctrines in Europe seems to open up for a pricing policy based on "full cost recovery", allowing for significant mark-ups on top of marginal costs. Whether the Norwegian government will follow this development remains to be seen. In such a case, there will be a need for evaluating a pricing policy along Ramsey-pricing rules.

1.2.3   The problem of choosing the appropriate tax instrument

I have already mentioned the fact that there is a choice between direct regulations and economic measures when one is to internalise external effects. There is no general answer to which of these approaches should be chosen. However, economic instruments are often considered superior to the direct regulations because they require less information to work well (Coase 1960) . If a regulator should internalise an externality through regulations, he would have to know the preferences of each actor in order to allocate the scarce resources right. When applying an economic instrument, the allocated resources are in markets where the interplay between the prices and the willingness of actors to pay for the commodity determines the allocation. However, when one needs exact direct control over the consumption of a commodity, or one considers willingness to pay as an unacceptable rationing mechanism, direct regulations may be considered most efficient[8].

There is a long list of economic instruments available for pricing road use, some of which are already in use (although not necessarily for internalisation purposes). Examples of such instruments are:

- General taxation
- Fuel taxes
- Tradable permits
- Vehicle taxes
- Infrastructure user charges
- Tolls
- Access charges
- Fines
- Insurance premiums
- Emissions charges

---

[8]   E.g. the allocation of lanes to traffic going in each direction would never be considered a pricing issue.

14

When considering the appropriateness of the different pricing measures, there are many dimensions to consider. The EU High Level Group for infrastructure charging (HLG), and later EU policy reports[9], have focused on the following assessment criteria:

- *Effectiveness*. The policy instrument's effectiveness in achieving the policy goal is the major consideration. If the application of an instrument does not result in the internalisation of external effects (with a highest possible degree of precision), it may be judged unsuitable even before considering the other assessment criteria below. Ability to vary tariffs along with the highly volatile marginal costs is an important feature.

- *Transparency and simplicity*. The chosen instrument must be clear and understood by the users in order to be effective. This relates to the very basic assumption underlying the theory of user charges: The presence of full information about the options at hand. Transparency and simplicity may also have attractions when considering political acceptability and also cost effectiveness (see items below).

- *Compliance and enforcement*. Enforcement costs of the policy instrument are always an important element when choosing among the different options. These costs are generally a function of the aims regarding level of compliance. A high resultant compliance level will usually mean high enforcement costs.

- *The cost of implementation, operation, information and transactions*. Generally this is about evaluating the overall cost effectiveness of the different measures available. Highly detailed and differentiated charging schemes will normally be accompanied by high costs of implementation and operation. Other policy instruments may require big outlays in information systems in order to achieve the necessary level of transparency. Systems that require traditional manual payment (i.e. manual toll collection) will have high transaction costs.

---

9   e.g. Outputs from the TRANSPRICE and CAPRI projects.

- *Acceptability*. Alternatives that may be both cost effective and transparent, and the same time achieve the policy goals, may not be politically or morally acceptable. Typically, these problems revolve around issues of privacy or equity. A high level of compliance may be achievable through a high level of control, but this may not be in accordance with privacy rights. Other policy instruments may favour high-income groups, and therefore be less attractive due to equity considerations.

- *Interoperability*. A charging system should be combined with existing payment systems, such as systems for collecting national tolls, on-street and off-street parking charges etc.

- *Possibilities of incentives, privileges and concessions*. Maximum flexibility for treating special groups of road users differently, according to policy wants.

The choice of instruments should be made also considering the totality of the regulating regimes. This means that as long as there are differences between the EU member states with respect to other forms of taxes and regulations, the choice of policy instruments for internalising externalities in the transport sector may give different results in different countries. Still, even when the choice of instrument turns out to be the same, the resultant *level* of taxation will vary from mode to mode and region to region due to different marginal external costs.

Most EU policy papers on this subject treat all the transport sectors in general. I will only discuss matters applicable to pricing *roads* here. The same policy instrument is not necessarily suitable for covering all fractions of externalities related to roads. The relevant external cost items related to road use are marginal infrastructure costs, congestion, environmental costs, and accident costs. In the following sections I will discuss the features of potentially suitable policy instruments for each of these cost elements. I will not consider the full range of criteria listed above, because many of them will have to be evaluated in a more concrete setting. Instead, I will return to these issues when considering a

proposition for a new pricing policy for Norway towards the end of the thesis (see section 6).

Policy instruments suitable for covering marginal infrastructure costs

Here I consider *road damage costs resulting from vehicle use*. These costs mainly depend on axle weight and axle configuration (and road durability features). Most EU member states have some form of taxation that is dependent on weight and sometimes axle configurations, but they are generally not dependent on the extent of use, i.e. the number of kilometres driven. Clearly, a charge must also vary with the distance travelled. Some form of a "weight-distance-tax" (WDT) should be chosen. Since marginal road wear also depends on the location and road type, the charge should also vary according to these differences. Generally such a charging system could either be based on a continuous or a point-specific system.

A *satellite-based* road charging scheme is becoming technically feasible[10], and may provide sufficient accuracy to be used on all major road networks in combination with digital road maps developed primarily for navigation purposes. Technological development (e.g. the introduction of the European Galileo positioning system) will eventually provide a sufficient accuracy for such a charging system.

A more *point-specific system* based on road-side beacons and gantries can register passing vehicles (equipped with electronic tags) with greater precision, but it would probably be much more costly to develop such systems to cover all road networks in Europe. On the other hand such systems have been in full-scale operation for many years (e.g. the Norwegian Q-Free or Autopass systems), and have proven reliable and efficient.

---

[10]   As demonstrated in the trials conducted in Copenhagen recently (Nielsen, O. A., and Herslund, M.-B... Although this first generation experiment had some technical problems, further software development and improved quality of positioning systems will make this technology feasible for full-scale operations.

Both these systems could be developed towards an automatic debiting system with electronic cash provided by Smartcards. The major advantage of having such a payment system is the interoperability, meaning that the same card could be used for several purposes, like paying for ferries, parking, public transport etc (Vougioukas et al. 2000).

With these systems it would also be possible to price according to registered maximum weight, but neither of these systems are capable of pricing according to *factual* axle loads. However, combining these systems with a state-of-the-art WIM-system (see the Appendix for a review of such systems) should also facilitate such differentiation. It is also possible to imagine on-board units registering the actual loads.

A major concern with some of these systems, is the privacy issue. Most of these alternatives make it possible to track the movements of a car, thus possibly violating basic privacy requirements. From this point of view, an on-board unit only receiving "one-way" signals for debiting a Smartcard or similar, would generally be preferable to systems that rely on centralised accounts or invoicing systems. Still, there are many examples of centralised systems having achieved political acceptance through specially designed routines to protect privacy[11].

Policy instruments suitable for covering congestion costs
We are looking for a time-, distance- and location dependent tax when considering policy options to contain the congestion costs. The purpose of such a charging system would be partly to achieve changes in modal split, partly to achieve changes in time of travel, and finally to re-route traffic to less congested areas. Existing charging systems do to a very limited extent vary with time of day, and consequently they are do not function as congestion charges. The first step

---

[11]   E.g. by having routines for deleting non-relevant information from photos, by only keeping registered data for a limited time etc.

towards a better congestion pricing system would therefore be to change current charging systems (toll cordons etc.) in order to make the level of charges dependent on external time costs incurred. In areas where the major problem is related to traffic terminating in a city area, an extended use of time-of-day dependent non-resident parking charges may also bring about some improvement.

Apart from the need to charge for different axle loads, the same requirements as for marginal infrastructure costs, also apply here. The calculation of marginal congestion costs is extremely complicated in urban road networks compared to single roads. The external time costs vary very much with time and location, but charging differently for every road section and every minute is not feasible for several reasons. Firstly, it is a question of what is technically possible. Secondly, the primary aim of such a pricing scheme is actually to affect the behaviour of the road users. This means that the information about the prices of different alternatives (modes, routes etc.) must be present at the time when the relevant decisions are actually made. This means that any feasible congestion-charging scheme will have to be based on a lot of averaging within geographical zones, and time-spans. Indeed, it seems that quite a high proportion of the potential efficiency gains from urban road pricing schemes could be reaped through a rather crude (but differentiated) cordon toll system (Grue et al. 1997). The new London Road Pricing scheme is based on only one zone and one price at the rush hours. However, some of the new pricing schemes have fairly differentiated charges, e.g. the mentioned trials in Copenhagen with nine zones and road classes, plus time differentiation. Also the new Singapore system[12] comprises a lot of differentiation with respect to fares, but it is still based on toll gantries.

Policy instruments suitable for covering environmental costs
Environmental impacts from road use cover a wide variety of effects, ranging from contributions to *global* warming from CO_2-emissions to local noise pollution

---

[12]   Implemented in August 2002.

and water run-off. The very different nature of all these impacts calls for a differentiated set of policy tools. CO_2-emissions are very closely correlated to fuel consumption, and the fact that these emissions cannot be cleansed makes fuel taxation a fairly good alternative for this part of the problem. Such a fuel tax should be differentiated according to carbon content.

Regional air pollution (i.e. acidifying substances, ozone etc.) causes effects over a large geographic area. The actual location of the emissions is therefore not that crucial to the magnitude of the impacts. Different vehicle types with different engines running on different fuels may cause very different effects on the regional level. A charging regime to cover these effects must be capable of reflecting these variations. Most of the differences could be reflected in a vehicle type-distance based taxing system, combined with different levels of fuel-taxes (e.g. with respect to sulphur-content). The severity of these impacts varies very much with initial levels and this means the level of charges should be subject to running adjustments.

Local air pollution and noise form a group of effects whose severity is very dependent on location. Generally the severity of the problem increases in densely populated areas and on heavily congested road networks. A taxing regime should also reflect differences with respect to vehicle types, as e.g. the presence of catalytic converters and engine sizes severely influence the magnitude of the marginal costs imposed. A pure vehicle-distance tax would not be a sufficient measure though, because of the huge variations from site to site. Such a tax should therefore be combined with site-specific tolls in densely populated areas.

Policy instruments suitable for covering accident costs
Generally the current use of economic policy instruments for internalisation of external accident costs is very limited. In some countries discounts are given on purchase taxes related to vehicle safety equipment etc., but these measures are not directly related to the factors that influence accident risks and costs related to

20

accidents (injury costs, loss of earnings, delays to other traffic, human suffering etc.). The EU High Level Group advocates a greater use and variability of traffic fines, based more closely on incremental accident risks. A somewhat more advanced instrument would be to use differentiated *insurance premiums* for reflecting different levels of risk. Traditional insurance systems would not, however, be able to separate risks occurring at specific roads under specific traffic situations etc. If advanced pricing instruments that are area and time specific should be implemented, it would on principle be possible also to implement a charge for external costs related to accident risk that is responsive to the prevailing traffic conditions. However, since the main point here is to attach a price to risk, insurance systems seem to be the best way for internalising external accident risks, provided that these systems become more responsive to actual risks by differentiating the system further. In the more distant future it is possible to imagine an electronic pricing system, dependent on current traffic conditions, climatic factors etc. that could actually mirror factual risks related to each trip better.

Figure 1-3 summarises a general assessment of the applicability of the different pricing-instruments, partly based on High Level Group (1999b) and De Borger et al. (2001). The crosses with the boldfaced types indicate my recommendations based on the assessment given above for the various cost categories.

For covering *infrastructure damage*, I find a *differentiated kilometre tax* as being the most applicable instrument for the time being. In the future a more advanced system dependent on location, actual axle loads, and possibly prevailing climatic conditions (e.g. spring thaw) can be foreseen.

Instrument	Infra-structure damage	Conges-tion	CO_2	Regional emis-sions	Local emis-sions	Noise[13]	Accidents
Insurance with detailed bonus and malus system	-	-	-	-	-	-	✦✦✦*
Parking charges	-	+	-	-	+	-	-
Fuel tax	+	+	✦✦✦	✦✦*	+	-	-
Annual vehicle tax or Purchase tax	+	-	+*	+	+*	+*	-
Public transport subsidies	-	+	+	+	+	-	+
Highway vignette	+	-	-	+	-	+	-
Area access charge or Cordon pricing	-	++	-	-	++	++*	-
Kilometre-tax	✦✦✦*	++*	++*	✦✦*	++*	✦✦*	+
Full electronic road pricing	++	✦✦✦	++	++	✦✦	++	+

+++   =   Instrument is highly recommended for the given cost element

++   =   Instrument may function fairly well as a proxy instrument

+   =   Instrument is a less desirable proxy instrument

-      =   Instrument has little (or negative) effect

*   =   If differentiated (e.g. by road type, time of day, vehicle type etc.)

✦✦(✦)   =   Recommended instrument

Figure 1-3   Pricing Instrument Options for Reflecting Social Marginal Costs

For capturing *congestion* costs a full *electronic road pricing* scheme would probably be the best solution, however cordon pricing or a area-dependent kilometre-tax could function as a (intermediate) crude form of road pricing.

The widely used *fuel tax* is only considered suitable for internalising *global*, and to some extent (if differentiated) *regional emissions*. An alternative to the latter could be an area specific kilometre tax.

Local emissions and noise have to be dealt with by applying some form of an area-specific and vehicle specific tax instrument. A *kilometre-tax* dependent on

---

[13]   Direct regulations also plays an important role in this area (technical standards).

22

these factors is recommended. An alternative would be to put this into a full electronic pricing scheme.

Finally, for covering the external *accident costs* I recommend a differentiated *insurance system* to price the external marginal risks connected to road use This is also supported by Lindberg (2001b).

In the future it is most likely that a *satellite based electronic road pricing scheme* would be able to capture all the cost components very well. This is also recently proposed by the European Commission (COM (2003)132 final) to be introduced in the period from 2005 to 2012, provisionally based on the American GPS system, and eventually on the planned European Galileo satellite navigation system. However, this proposal needs political support, and it may be a controversial issue mainly due to privacy issues. Such a system will inevitably be very close to a "Big Brother" system allowing governments to have access to very detailed information about people's movements. So, even if the technology is ready, it is my assessment that the political issues will curb this development, and therefore we will need some intermediary charging instruments that are better suited than the current fuel taxes.

1.2.4   Other policy (non-price) instruments

Other policy instruments may also have an economic character, e.g. investment policies, design and dimensioning of public transport systems etc. However, for some cost categories, externalities may be dealt with more efficiently through direct regulations rather than market interventions like the Pigouvian taxes. This is probably most prominent in the fields of traffic safety and environmental problems. In the safety area we have a lot of regulations with respect to the design and maintenance of the infrastructure, vehicle specifications and not least; driver behaviour. There are also many examples of regulations with respect to environmental issues, e.g. prohibition of excessive noise emissions, standards for

maximum air emissions, etc.

We also apply many technical "demand management" systems trying to improve traffic flows, thus reducing congestion problems. Traditional instruments like banning certain vehicles from certain areas, requiring minimum occupancy of vehicles for certain lanes, imposing parking restrictions etc., are now being supplemented with more adaptive intelligent transport systems like "ramp metering", adaptive traffic lights etc. From an economic point of view, the main problem with these technical rationing devices, is that they do not allocate the scarce road capacity according to willingness to pay, thus imposing an efficiency loss. However, many of these regulations are very cheap and easy to install compared to introducing a full electronic road pricing system. The trade-off will then be between the efficiency loss imposed by a technical rationing device, and the transaction costs of a full road pricing scheme.

Generally, all instruments that affect road use levels could be included in this review of policy instruments. This means that land use considerations, development of alternative transport modes, tele-commuting etc. all are instruments that to some degree also will have an effect on transport externalities in the longer term.

I have now tried to establish the basic and most relevant theory elements that form the foundation of road user charges. I started by sketching the fundamental elements of welfare theory, then focusing on the existence of market imperfections, focusing on externalities. Then I moved on to introducing the notion of Pigouvian taxes for internalising the external effects. Finally I have, in general terms, considered the applicability of some alternative pricing instruments available for pricing road use.

24

Moving on towards a proposed pricing policy, the next natural step would be to start establishing the empirical foundation for pricing policies, i.e. the estimation of marginal external costs related to road use.

2   THE MARGINAL COSTS OF ROAD USE

2.1   *DEFINITIONS AND BASIC ASSUMPTIONS*

I define the marginal cost of road use as the avoidable costs related to the actual use of the road. The term cost is here linked to the value of the best alternative use of the resources spent on *using* the road. This notion is closely linked to the individual user choice between using the road, or not using the road in question. The costs of choosing to use the road are then mirrored in the benefits forgone in the form of other alternative consumption alternatives that could have been chosen instead of using the road.

The term "marginal" reflects that I only consider the costs related to an individual road user's decision to use the road or not. This means that I do not consider the full costs of actually providing the road, but am concerned only with the *extra* cost connected to use. In some studies (e.g. the US Federal Highway Cost Allocation Study), the major focus is on average costs, rather than marginal costs. In this case the full costs of providing the roads are also relevant.

Basically I assume that the infrastructure is owned and operated by the authorities, making them responsible for the associated costs. Lately there has been a development towards so-called Public-Private-Partnerships (PPP) in the provision of roads. As long as the contracts and incentives of these partnerships are well designed, i.e. likely to achieve socially optimal solutions with respect to investment and maintenance, this should not alter my conclusions. Accordingly I assume vehicles to be owned by a large number of road users (i.e. no monopoly in the use of the road).

2.2   THE COMPONENTS OF MARGINAL EXTERNAL ROAD USER COSTS

2.2.1   Cost items related to Road Use

Vehicles impose *four main types* of cost on the rest of society: Road damage costs, congestion costs, accident externalities and environmental costs (Newbery 1988c, High Level Group 1999b, etc.).

Road damage costs

I define road damage costs as those arising from road wear due to vehicles passing. They fall mainly into two categories: Increased costs of repairing the road, and additional vehicle operating cost related to driving on a rougher road. The first, i.e. the pavement costs, are usually borne by the highway authority, whereas the latter is borne by other road users. If road maintenance policies are condition-responsive, i.e. roads are repaired when their condition reach a predetermined state[14], it can be shown that the road damage externalities (i.e. the vehicle operating costs) are negligible (Newbery 1985), and that road damage costs equal the average cost of road repair[15]. This result, called the 'Fundamental Theorem of Road User Charges', makes the calculation of these costs much easier (see section 2.4 for a thorough presentation of Newbery's theorem).

Congestion costs

Congestion costs may take a variety of forms. The most direct form arises when external time loss is imposed on other drivers on the road network when one additional vehicle enters a congested road. More indirectly, the need of traffic regulating measures (e.g. traffic signals) is also related to a high level of traffic, so the costs related to these systems may also be regarded as congestion costs.

---

[14]   This predetermined state does not necessarily relate to an *optimal* level of maintenance.

[15]   This is based on the presumption that there is no traffic growth, and that road damage is solely caused by traffic. Climatic impacts on road deterioration may be very different from situation to situation. On a high volume weak road the costs would be very sensitive to weathering. In severe climates a relatively higher proportion of road wear costs should be allocated to climate rather than traffic.

28

Accident costs

Accident risk increases for other road users when another vehicle enters the road network. Accident costs are partly met by insurance payments, but not entirely internalised.

Environmental costs

Furthermore one usually divides the damages from air emissions into global, regional and local impacts. The main global issue is related to the so-called greenhouse effect (i.e. global warming), but also to the depletion of the ozone layer, which leads to less protection from UV radiation for life on Earth. Regional effects are effects that arise in specific geographical regions of the World as a result of a high accumulated concentration of emissions in that (or another) region. A typical example of regional effects is the problem of acidification ("acid rain"), which mainly stems from sulphur emissions. Local air emissions are fumes and particles that generally jeopardise health and welfare close to the heavily trafficked road networks.

2.2.2 *The notion of Cost Responsibilty - Average or Marginal approach*

Ronald H. Coase (Coase 1960) argues that if there were no transaction costs, and property rights were well defined, then a market for transactions between the perpetrators and victims of an externality would make Pigouvian taxes redundant. However, for a number of reasons, it seems that the market actually *does not* solve the problem of externalities in general. This is partly so because of the prohibitive transaction costs related to the internalisation of many externalities. In our setting it would be quite unrealistic having the individual motorist actually bargaining with all other motorists, residents, stakeholders etc. to compensate the externalities related to his road use.

The problem of establishing property rights is most often raised in relation to environmental problems. "Common ownership" of resources like "clean air", does not make it clear who are the perpetrators or who are the victims in many

cases. The OECD has established the so-called Polluter Pays Principle, which is a political/moral solution to the problem of ill-defined property rights. Even though the interpretation of such a principle may also be ambiguous (Pearce and Turner 1990), it makes it clarifies who should be liable for pollution. Most policy proposals on pricing environmental externalities are made subject to the Polluter Pays Principle.

The notion of *cost responsibility* is sometimes prominent in cost allocation studies related to road use. The interpretation of the term is rather different in the various studies. Allocating a portion of some total cost (sometimes referred to as the average cost approach) of providing roads to different user groups will inevitably be a rather arbitrary affair due to the fact that a lot of the cost items in question are genuinely joint costs. These costs allocated to different groups will have to be according to some principle not deductible from economic welfare theory. Instead some notion based on equity or fairness has to be applied.

2.3 NON-MARKET VALUATION

Economic valuation of time, health (accidents) and environmental impacts is central when trying to establish the foundation for road user charges. These are commodities without market prices applicable in the calculation of marginal external costs. Hence, *shadow prices* are applied instead, and this notion is presented in the following section. Then state-of-the-art estimation techniques are briefly presented for the three cost components in question, based primarily on recent European research.

2.3.1 Shadow prices

Shadow prices are imputed values of commodities based on the *opportunity costs* of producing or consuming those commodities. The shadow price should thus represent the value of the commodity in its best alternative use (consumption or production).

In the absence of real markets for such commodities we estimate the *willingness to pay*[16] (WTP), to evaluate their economic values. Basically this could be done by studying *stated preferences* (SP), or *revealed preferences* (RP) of consumers (and producers). In the case of SP a *hypothetical market* is created, and respondents are asked to make direct or indirect valuations of non-market commodities. The Contingent Valuation Method (CVM) is among the most common SP-techniques mainly used for the valuation of environmental commodities. RP studies try to elicit the implicit value of non-market commodities from people's behaviour in actual (surrogate) markets where the consumption of the non-market commodity has an influence on this behaviour. A typical example would be the estimation of values of time from studying the choice between two alternative road routes. One route is fast, but this is a toll-road, the other route is slower but has no toll. Here, by studying how drivers choose between saved time and paying or not paying the toll, one can derive a value of time.

Stated Preference techniques

SP techniques have some advantages over RP techniques because the researcher has absolute control over the variation in the variables. In many real life situations statistical analysis is hampered by strong correlations or limited variation in the central variables. Careful experimental design[17] of the SP-analysis can avoid many of these typical weaknesses of RP data. The fact that one is not limited to exploring currently existing choices is another advantage making SP analysis popular. Finally, in most cases, an SP analysis is far less demanding with respect to resources needed as well. Still the extensive use of SP data is often under debate, due to many pitfalls and possible shortcomings. The intrinsic fact that the preferences are *stated* will always trigger a question about the transferability of

---

[16] The willingness to pay is used for improvements. Alternatively one measures the *compensation* needed for accepting a reduction in availability of some commodity. This measure is called the *willingness to accept* (WTA).

[17] For recommendations on a stepwise approach to a thorough experimental design of SP analysis, see Hensher, D. A. (1994)..

the results to real world situations. People do not always do what they *say* they will do. This may be due to

- *Strategic answers.* SP studies are often conducted when concrete policy alternatives are being considered (e.g. the introduction of road pricing). Respondents could be able to see through the purpose of the survey and adjust their answers in line with their private interest in the policy.
- *Lack of information or realism.* If respondents are not provided with enough relevant information, thus denied a realistic context for the choices they are asked to make, then responses may not be transferable to the real situation.
- *Biased answers.* Studies indicate that results may be very sensitive to the way questions are asked, how the interaction between the interviewer and respondents works, and also to details in questionnaire design etc.

All these elements limit the credibility of SP studies, but the issues mentioned could all be addressed, and the problems limited by a good study design. The problem of strategic answer could be dealt with through concealing the purpose of the study by presenting complex multidimensional choices to the respondents, rather than simple two-factor trade-offs. By linking the choices to a "real world" setting, e.g. by asking questions linked to a real journey rather than a hypothetical one, the realism could be enhanced. New, computer-based techniques could also create an artificial, but realistic setting for the interviews. The potential biasing behaviour of the interviewer could be avoided through introducing direct computer-based presentation and registration of answers.

A special branch of SP studies is the *Contingent Valuation Method* (CVM). It consists of survey techniques in which the analyst, by asking willingness-to-pay or willingness-to-accept types of questions, collects information on individual preferences regarding some sort of welfare change. The information can be

obtained through open-ended or closed questions. In the former, the respondents state the maximum amount that they will be willing to pay for an improvement, e.g. having the option of going to a national park. Alternatively, the equivalent willingness-to-accept question would be: "What compensation would you need to accept loosing this option?". Using closed-ended questions means that one only asks (a sequence of) questions with yes/no answers[18], like: Would you be willing to pay € 50 (€ 60, € 70 etc.) per year for having the option of visiting this national park? CVM techniques are much applied for estimating non-use values, i.e. option and existence values. CVM studies carry all the weaknesses related to SP studies in general. A particular problem with this rather direct approach to estimating willingness to pay is related to realism with respect to the respondent's budget constraint. People may tend to overstate their willingness to pay for individual commodities because the hypothetical setting (i.e. with no real "threat" of actually having to pay) does not make respondents actually set priorities for all the consumption possibilities that should be fitted into their budget constraint. Thus, summing up the stated willingness to pay for different environmental commodities may far exceed the total budget.

For a comprehensive review of the history and development of SP analysis, see Johansson (1999).

Revealed Preference techniques
Typical RP studies involve hedonic pricing, the travel cost method, avoidance cost, lost production/income estimates etc. Many of these are used for pricing environmental externalities. A typical example of *hedonic pricing* would be a study of house prices where one could try to elicit the implicit prices related to e.g. noise strains. Assuming that the market has proper information about the noise conditions related to houses, one would expect this to be reflected in housing market prices. Including noise information in a statistical study (e.g.

---

[18] This type of CVM studies is often called "bidding games".

regression analysis), one could single out the effect of noise, thus deriving a shadow price of noise.

Accordingly, if one wants to assess the use value of a recreational area, one could study the *travel costs* related to visits to this area. That information combined with information about choices of going or not going to visit the area could give an indication about the willingness to pay for this environmental good.

Some environmental effects could be avoided by introducing certain measures. The costs of implementing these measures would then represent the *avoidance cost*. An example would be the cost of double-glazing in order to avoid indoor noise from traffic.

In the evaluation of accident costs, estimates of lost production, or income, forms one of the relevant cost components. Assuming a perfectly functioning labour market, the wage rate represents the alternative cost of such losses.

In some of the examples mentioned above, RP data will not provide information about the total costs or benefits related to these commodities. In all cases they will only represent the use value of the commodity in question. Many environmental commodities also have *option and existence values* that should be added to the use value. People often have a willingness to pay for having the option (possibility) to use the commodity, and some are even willing to pay for the mere existence (without any plans for using) some commodities. These are cost components related to environmental effects that one normally would have to rely on SP surveys to assess.

As RP data, too, will be flawed by errors, it is not a question about choosing RP *or* SP as the most proper instrument for studying preferences. Applying both methods gives an opportunity for "methodological triangulation", which may be

very fruitful.

2.3.2 Estimating the Value of Time

External time costs are probably the most important externality connected to congested road networks. Estimating the value of time (VOT) is therefore extremely central connected to both road pricing and road investments. A number of value of time studies (Nellthorp et al. 2001) has therefore been conducted over the past decade.

Theoretically, assuming a well functioning[19] labour market, wage rates should be a proper indicator of the VOT. The gross wage rate should represent the opportunity cost of labour, thus being a good indicator of resources foregone in a production perspective (i.e. relevant for evaluating work-time). The net wage rate (after tax deductions) should be equally suitable for evaluating the opportunity cost of leisure-time. At the margin, individuals compare the monetary return from an additional unit of work time to the subjective benefits they expect from pursuing an additional unit of leisure time.

However, there is empirical foundation for saying that time is not a homogeneous product. This means that it doesn't make sense to detach the valuation of time from the activity in question. E.g. it is a broadly accepted empirical result that the WTP for avoiding a minute extra waiting for the bus is much higher than the WTP for saving an extra minute in (in-vehicle) travel time.

In empirical passenger VOT-studies a typical segmentation would be by (op.cit.):

- Travel purpose (business, commuting, leisure)
- Mode (car, bus, rail, air)
- Travel distance (urban/local, inter-urban/long distance)
- Travel condition (expected travel time, delay time, in vehicle/walk/wait)

---

[19] By "well functioning" I here mean a labour market without market distortions.

A different segmentation would be more appropriate for freight vehicles:

- Mode (road, rail, inland waterways, sea, air)
- Goods type (high/low value, (non-)perishable, light/heavy , (non-)unitised, etc.)
- Vehicle type (light/heavy truck, with trailer)

Both RP and SP based studies are being applied for VOT-estimation. For example, values of time could be estimated in a RP setting by studying route-choice when road users can choose between a fast but tolled road, and a slow untolled road (Hjelle 1992). The trade-off between the toll payment and the time saved gives an opportunity to study values of time.

Accordingly, values of time representative of airline passengers could be estimated based on a stated choice study conducted in the airports, and linked to the actual travels. Here, respondents could be asked to choose between specified alternatives with respect to access/egress transport (e.g. rail, bus or taxi), each carrying different characteristics with respect to travel time, waiting time and price.

2.3.3   *Estimating the Value of Statistical Life*
The valuation of externalities related to lives lost and reduced health, is probably the most controversial and difficult area among those addressed here. Obviously, when considering certain death for an individual, the willingness to pay for avoiding it would tend to be infinitely large. However, the fundamental externality here is not life or death, but different *risks* of dying or losing full health. Marginal road use may affect the risks of having an accident for other road users, and our aim is therefore to put a monetary value on this incremental risk. To distinguish this from the individual value of life, we are looking for a valuation of a *statistical*

life (Jones-Lee 1994). Generally, the value of a statistical life (VOSL) could be estimated based on RP and SP data. Typical RP approaches would be to study markets where risk has a monetary value, e.g. in the wage rate of jobs with different risk levels. The most used source of RP information in this area would be the insurance business, where pricing risk is a core activity. To elicit such valuations from the transport business itself, one could extract evidence from the WTP for extra risk-reducing equipment (safety belts, airbags, ABS brakes etc.).

Accordingly, SP surveys could also be used, making respondents choose between activities with different risk levels, and prices. Jones-Lee (op.cit.) concludes that evidence from RP and SP surveys should be considered complimentary, and that both approaches are needed. Reviews of many SP based studies (e.g. Nellthorp et al. 2001) raise many questions about the reliability of the resulting estimates. Typical problems are:

- To what extent are the respondents able to grasp the type of effect a certain risk-reducing action would provide. Empirical evidence suggests that there is a tendency of respondents expressing the same WTP for large and small risk reductions.
- Typical risk reductions may be so small to the individual, that the respondents would be indifferent to them.

The total cost of a traffic injury comprises the following items (Elvik 1994):

- Lost quality of life
- Travel time delay
- Medical treatment
- Lost output
- Property damage, and
- Administrative costs

Shadow prices are needed for estimating the first two cost components, whereas medical treatment, lost output, property damage, and administrative costs generally have market prices that might be used in the calculation of external accident costs.

2.3.4   *Estimating the Value of Environmental Impacts*

We have already seen some examples of the estimation of environmental impacts (see Section 2.3.1) can be estimated using SP and RP approaches. Generally we need estimates that cover the use value, the option value, and the existence value of environmental commodities. As stated above, SP techniques would have to be applied for the latter two, whereas RP data could give information about the use value (e.g. through hedonic pricing techniques).

Shadow prices for environmental impact will most often be based on "the impact pathway approach" (Figure 2-1). This means that there are many factors that influence the shadow prices of environmental external effects. The economic evaluation of the impacts only represents the last step in this pathway. Preceding these evaluations, physical emissions from traffic must be calculated (dependent on vehicle types, speed, climatic and topographical conditions, traffic density etc.). Then concentrations of the emitted substances must be estimated, based on the receptive capacity of the environment in question. The impact of the resultant concentrations will depend on population density and how exposed residents, animals and other elements of nature will be to the emitted matter.[20]

The main types of environmental external effects from road use are noise and emissions to air. Some air emissions have local effects (e.g. particulates), some regional (e.g. sulphur) and some global (e.g. CO_2).

---

[20]   In some expositions, intermediate stages between the Emission and Physical Impact phases are specified: Emmissions-Transport & Chemical Conversion-Concentration & Deposition-Response of Receptors-Physical Impact.

38

Traffic flow	→	Emission	→	Concentration	→	Impact	→	Cost
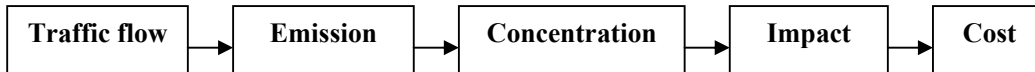
Figure 2-1   The impact pathway approach (Rennings et al. 1999)

A much more detailed exposition on the Impact Pathway Methodology applied to European transport systems could be found in Friedrich and Bickel (2001), which is based on the outcome of the ExternE-project conducted under the 4th European research programme.

I have now given some fairly general comments on possible ways of estimating shadow prices for time, health/life, and the environment. I will return to more concrete evaluations of empirical evidence on the magnitude of these costs when establishing a calculation model for marginal road user costs in Norway (in Section 5).

2.4   NEWBERY'S FUNDAMENTAL THEOREM OF ROAD USER CHARGES

In Newbery (1985) and subsequent papers by the same author, a fundamental theorem of road user charges, focusing on the distinction between the marginal and the average approach to the cost responsibility of the road users, is presented. This problem area is so relevant to the problems addressed in this thesis that I will reproduce Newbery's major derivations and results over the next pages (the comments between the equations do not necessarily match the ones originally made by Newbery, and some intermediate stages of the derivations have been added).

The basic assumptions behind Newbery's model are:

- Traffic is constant at N vehicles per annum

- All road wear stems from the impact of traffic[21]

- Maintenance actions are consistently responsive to pavement condition[22]

There are basically two different elements of road wear costs imposed by road users: First, the costs of maintenance (traditionally borne by the road authority) triggered by the use of the roads, and second, the increased vehicle operating costs that are imposed on other road users stemming from the deteriorated road condition as a consequence of road use. According to Newbery (op.cit.) the second category of costs is between 10 and 100 times as large as the first one, and therefore of great importance. However, focusing on *marginal* figures, the two cost components seem to be of the same magnitude, according to Newbery. The marginal social cost of road wear (relevant for charging Pigouvian taxes) therefore constitutes the costs of increased road maintenance caused by an extra vehicle or axle (ESAL) kilometre[23], plus the increased vehicle operating costs imposed on other drivers due to rougher roads.

Newbery establishes a model based on three basic relations. The first equation establishes the development in accumulated traffic load. Here X denotes the cumulative number of ESAL transits up to date (*z*). N is the number of vehicles per annum, and E is the average number of ESALs produced by one vehicle. At time *t* the accumulated traffic load, Y(*t*), could be expressed as

---

[21] This means that the model does not cover climatic impacts that also may contribute to road wear (e.g. precipitation, sunshine, freeze-thaw cycles etc.). However, Newbery also presents a more general model which also allows for these effects. This modifies the results somewhat (see page 50).

[22] This means that maintenance actions are carried out whenever the road reaches a certain (not necessarily optimal) level of distress, and not periodically or arbitrary.

[23] Newbery uses the common Equivalent Standard Axle Load (ESAL) as the unit for measuring traffic loads. This is based on the 4th power law stemming from the AASHO road test, assuming that any axle load's damaging power is proportionate to the fourth power of its loading. See section 3.2.1 for a closer description of the AASHO Road Test.

Roughness (*R*)
Vehicle operating cost (*v*)
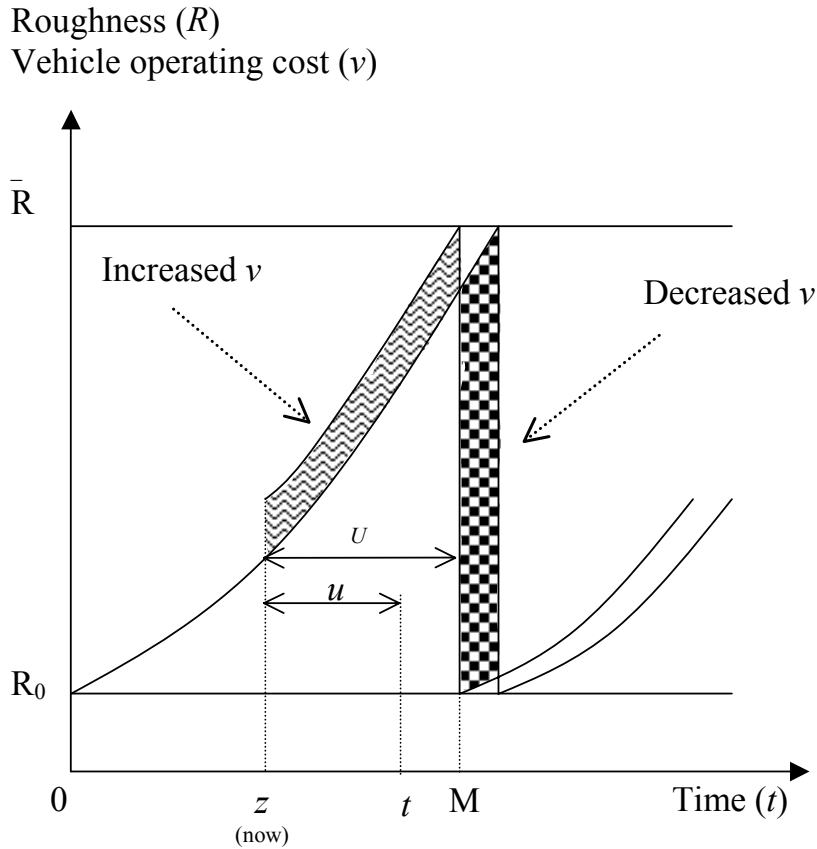
Figure 2-2   Illustration of the external impact of an extra axle[24]

Equation 2-1 $Y(t) = X + N \cdot E \cdot (t - z), \quad t \geq z$

The roughness of the road is a function of this accumulated traffic load:

Equation 2-2 $R = R\left\{ Y(t) \right\}$

Vehicle operating costs, *v*, is expressed as a function of the roughness of the road:

---

24   Based on Newbery, D. M. (1985). and Newbery, D. M. (1988a)..

Equation 2-3 $v = v(R)$

The maintenance strategy of the road authorities is to restore the roughness of the road to the initial level (R_0 in Figure 2-2) whenever the roughness of the road reaches the critical level $\bar{R}$. The cost, C, of this maintenance (e.g. a reseal) is assumed to be constant. Under the assumptions made at the beginning (constant traffic level, and all road wear coming from traffic loads), this maintenance action will recur every T years. Actually T will depend on the strength of the road structure, traffic loads and the maintenance-triggering level of roughness, $\bar{R}$.

If we assume that these maintenance cycles go on infinitely, the present discounted value of all future vehicle operating and road maintenance costs on a road which was restored z years ago, and which has since carried X ESALs, can be written as

Equation 2-4[25]

$$F(X,z) = e^{-r(M-z)} \frac{C}{1-e^{-rT}} + N \int_{t=z}^{M} v\{R(Y(t)\} e^{-r(t-z)} dt + e^{-r(M-z)} \frac{N \int_{t=0}^{T} v e^{-r(t-z)} dt}{1-e^{-rT}}$$

Here, the first section of the equation illustrates the present value of the maintenance costs of the established [R_0, $\bar{R}$] strategy. The expression is based on continuous capitalization and assumes that the overlays represent an infinite geometric series expansion. The middle section illustrates the vehicle operating costs from now, z, until the next overlay, M. The last section has an equivalent interpretation, but represents all the subsequent maintenance cycles after the next overlay. In Equation 2-4 all the future costs are discounted by a discount rate, r.

---

[25]   This equation is slightly altered compared to the ones presented in Newbery, D. M. (1988b) and Newbery, D. M. (1985)., since the variable u is not introduced yet. This is done in the next version of the equation (see Equation 2-7).

Once again continuous capitalization is assumed.

The date of the next maintenance action, *M,* is implicitly defined by the required number of cumulative ESALs, $\overline{Y}$, necessary to reach the triggering roughness level, $\overline{R}$. As shown in the following equations, *M* can be expressed as the sum of the current date, *z*, and a term representing the extra axles necessary to reach the triggering roughness level, divided by the annual number of axles, *NE*:

Equation 2-5 $\overline{R} = R\{\overline{Y}\}, \quad \overline{Y} = X + NE(M - z), \quad M = z + \dfrac{(\overline{Y} - X)}{NE}$

Our interest revolves around the marginal social cost (MSC) of road wear generated by an extra standard axle. For practical purposes an infrastructure charge could not be made dependent on the current age[26] of the pavement on every road section. Estimating an average over roads of all ages, from zero to *T* years, is more useful:

Equation 2-6 $MSC = \dfrac{\overline{\partial F}}{\partial X} = \dfrac{1}{T} \int\limits_{z=0}^{T} \dfrac{\partial F}{\partial X}(X, z)\, dz$

Newbery's Fundamental Theorem claims that the expected (average) marginal social cost (MSC) is exactly equal to the *average* maintenance cost, $\dfrac{C}{NET}$. In (Newbery 1985) this is shown by differentiating Equation 2-4, demonstrating that the road damage externality terms outweigh each other.

In (Newbery 1988b) two other proof strategies are presented:

---

[26] Here, the term "age" really refers to the number of accumulated ESALs the road has been exposed to since the last overlay.

1. Via simplifying variable transformations, and
2. A more direct approach to calculating the optimal user charge for road wear, assuming roads with uniform age distributions.

The first alternative approach starts with introducing the variable U, the time before next overlay (see Figure 2-2), i.e. $U = M - z$. Equivalently, the lower-case u is defined by $u = t - z$. This enables a rewriting of Equation 2-4, expressing the discounted social cost of road wear as a function of X (also substituting for $Y(t)$ according to Equation 2-1):

Equation 2-7 $\quad G(X) = e^{-rU} \dfrac{C}{1-e^{-rT}} + N \displaystyle\int_{u=0}^{U} v\{R(X+NEu)\}e^{-ru}\,du + e^{-rU}\dfrac{N \displaystyle\int_{u=0}^{T} v e^{-ru}\,du}{1-e^{-rT}}$

Since $U = M - z$, and $u = t - z$, it follows from Equation 2-5 and Equation 2-1, respectively, that U and u are functions of X, hence G is also solely a function of X. Since the age of the road can be described as well by U as by z, the expression for the expected MSC (Equation 2-6) could also be based on Equation 2-7, using U as the variable of integration:

Equation 2-8 $\qquad MSC = \dfrac{\overline{\partial F}}{\partial X} = \dfrac{1}{T}\displaystyle\int_{U=T}^{0}\dfrac{\partial G}{\partial X}\,dU$

The effect of increasing X is exactly the same as the effect of increasing the age of the road, U. Since U is a function of X, G could also be differentiated with respect to U:

Equation 2-9 $\qquad \dfrac{\partial G}{\partial X} = \dfrac{\partial G}{\partial U}\cdot\dfrac{\partial U}{\partial X} = -\dfrac{1}{NE}\dfrac{\partial G}{\partial U}$

44

The expression for $\dfrac{\partial U}{\partial X}$ comes from substituting U for M-z in Equation 2-5, solving this equation with respect to U, and then differentiating with respect to X. The result from Equation 2-9 can be substituted into Equation 2-8 to yield:

Equation 2-10 $\qquad MSC = \dfrac{\overline{\partial F}}{\partial X} = \dfrac{-1}{NET} \displaystyle\int_{U=T}^{0} \dfrac{\partial G}{\partial U}\, dU$

The next step is then to evaluate this integral. The minus-sign is dropped by exchanging the upper and lower boundary of the integral. Bearing in mind that $X = \overline{Y} - NEU$, doing the integration yields

Equation 2-11 $\qquad MSC = \dfrac{\overline{\partial F}}{\partial X} = \dfrac{1}{NET} \Big[G(\overline{Y} - NEU) \Big]_{T}^{0}$

Inserting the boundary values for U in Equation 2-7 makes it possible to solve the equation. First we calculate the upper boundary value ($U = 0$):

Equation 2-12

$$G(X \vdots U = 0) = e^{-r \cdot 0} \frac{C}{1 - e^{-rT}} + N \int_{u=0}^{0} v\{R(X + NEu)\} e^{-ru}\, du + e^{-r \cdot 0} \frac{N \displaystyle\int_{u=0}^{T} v e^{-ru}\, du}{1 - e^{-rT}}$$

$\Updownarrow$

$$G(X \vdots U = 0) = \frac{C}{1 - e^{-rT}} + \frac{N \displaystyle\int_{u=0}^{T} v e^{-ru}\, du}{1 - e^{-rT}}$$

Equation 2-13

$$G(X \vdots U = T) = e^{-r \cdot T} \frac{C}{1-e^{-rT}} + N \int_{u=0}^{T} v\{R(X + NEu)\}e^{-ru} du + e^{-r \cdot T} \frac{N \int_{u=0}^{T} ve^{-ru} du}{1-e^{-rT}}$$

Inserting the results from Equation 2-13 and Equation 2-12 (i.e. subtracting them) into Equation 2-11, yields

Equation 2-14

$$MSC = \frac{\overline{\partial F}}{\partial X} = \frac{1}{NET} \left[(1-e^{-r \cdot T}) \frac{C}{1-e^{-rT}} - N \int_{u=0}^{T} v\{R(X + NEu)\}e^{-ru} du + (1-e^{-r \cdot T}) \frac{N \int_{u=0}^{T} ve^{-ru} du}{1-e^{-rT}} \right]$$

Here the two last terms in the brackets cancel, and the first term reduces to C, thus arriving at Newbery's conclusion, i.e. that the marginal social cost of road wear equals the average cost per cumulative standard axle:

Equation 2-15 $\qquad MSC = \frac{\overline{\partial F}}{\partial X} = \frac{C}{NET}$

Newbery's direct approach to road user charges
In the same work (Newbery 1988b) a more direct approach to estimating road user charges is developed. The framework is much the same as the one used to underpin the result of marginal social cost being equal to average maintenance cost. However, the focus will here be on establishing a social welfare function, which includes all relevant benefits and costs related to road wear.

For expositional ease, suppose that the road network consists of a single road of

46

length 1 km, and that this road has a uniform age distribution[27] between 0 and T years (T representing the period between overlays). Suppose also that at date $t=0$, this uniform age distribution is arranged in such a way that the newest road sections (i.e. the ones that have most recently undergone an overlay) are at the beginning of the stretch, and the oldest ones at the end. This means that the age of the road section x km from the start is xT years. Each road user demands q_i trips per annum, and each of these trips imposes a damaging effect of e_i ESALs. The total annual traffic measured in ESALs is then $Q = \Sigma q_i e_i$. The carrying capacity of the road is *NET*, the number of standard axles necessary to reach the overlay criterion. Generally at an arbitrary date t and a distance, x, the age (i.e. number of cumulative ESALs) of the road surface since last overlay could be expressed by the following equation:

Equation 2-16 $$X(x,t) = NETx + \sum_{i=1}^{I} q_i e_i t \bmod NET$$

This relationship is illustrated in Figure 2-3. If t=0 then the second term in the equation becomes zero, and $X(x,0)=NETx$. This case is illustrated by the diagonal in the figure, which has a slope NET. At any other value of t, the cumulative age of the road will reach the overlay criterion over the road section, and the age of the road is then set back to zero (at x^* in Figure 2-3). In the formula this effect is taken care of by the modulo function. The vehicle operating cost is a function of the roughness of the road, and total vehicle operating cost for each individual for traversing the road section can be calculated as the area under this curve, as shown in the following equation

Equation 2-17 $$\overline{v_i} = \int_{x=0}^{1} \left[R\{X(x,t)\} \right] dx$$

---

[27] A uniform age distribution means that the road stretch in question comprises equal shares of each age (e.g. year) class.

Since the road length is 1 km, this could also be interpreted as the *average vehicle operating cost per km*.
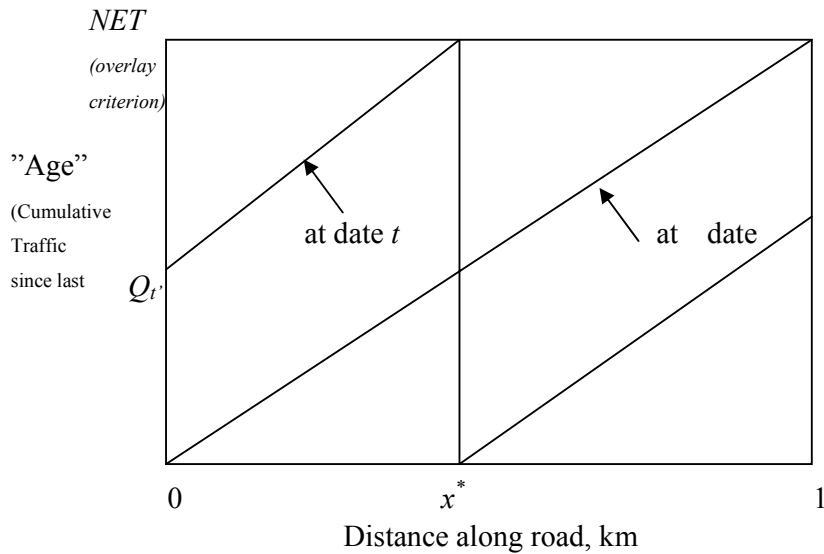
Figure 2-3   "Age" of road as a function of distance and time[28]

If vehicle i is charged a fee f_i per km, the total private cost will be $p_i = \bar{v}_i + f_i$ per km. This is the price parameter that will determine i's demand for trips, q, as follows. If $V_i(p_i, b_i)$ is the indirect utility function, where b_i is lump sum income, then, referring to Roy's identity, the demand functions could be expressed as

Equation 2-18
$$q_i = \frac{\dfrac{-\partial V_i}{\partial p_i}}{\dfrac{\partial V_i}{\partial b_i}}$$

where the denominator is the expression for the private marginal utility of income.

---

[28] From Newbery, D. M. (1988b)., Figure 2.

48

The proportion of the road network reaching the critical state each year (due to the uniform "age" distribution) is $\dfrac{\sum q_i e_i}{NET}$. C is still the overlay cost per km, hence the total annual cost of maintenance per km is $\dfrac{C\sum q_i e_i}{NET}$.

A user charge f is levied on vehicle kilometres. Surplus income, $\sum f_i q_i - \dfrac{C\sum q_i e_i}{NET}$, from the user charge (i.e. income over maintenance cost) is assumed redistributed to users, and vehicle owner i receives a share α_i, $\Sigma\alpha_i=1$. The sum of these surpluses represents the net welfare gain, assuming the current income distribution to be optimal[29]. The necessary components to compose a welfare function, based on the indirect utility functions, to be maximized subject to the user charges f_i, has now been established.

Equation 2-19
$$\max_{f_i} W\left\{V_i\left[\bar{v_i}+f,\alpha_i\left(\sum_j f_j q_j - \frac{C\sum_j q_j e_j}{NET}\right)\right]\right\}$$

Based on this maximation problem, Newbery (op.cit.) concludes that the optimal charges per vehicle km are represented by

Equation 2-20
$$f_i = \frac{Ce_i}{NET}$$

Hence, the optimal charge per standard axle kilometre is equal to the average maintenance cost:

---

[29] Technically this means that $\dfrac{\partial W}{\partial V_i}\cdot\dfrac{\partial V_i}{\partial b_i}=1$.

Equation 2-21 $\dfrac{f_i}{e_i} = \dfrac{C}{NET}$

which is exactly the same result as given in the first approach (see Equation 2-15).

With some further elaboration on this result (extending it to a road network with varying types of roads and traffic levels), Newbery sets forward the following propositions based on the preceding:

- *Proposition 1: If the age distribution of roads of a given type is constant, and the traffic flow is constant, and all road damage is attributable to traffic, then the average road damage cost of a vehicle is identically equal to the average maintenance cost allocated in proportion to its number of equivalent standard axles. The road damage externality is zero.*
- *Proposition 2: With the assumptions of Proposition 1, the optimal flat charge per ESAL km will exactly recover the road maintenance costs of the network if the demand for trips is linear with a slope coefficient uniform across vehicles. If not, the optimal damage charge will be a weighted average of the road maintenance costs.*

Allowing for the effect of weathering and traffic growth

From the vast amount of empirical research conducted in order to identify road damage relationships, it is well known that weathering effects also play a potentially important role in addition to the traffic loads. The assumption of all damage being attributable to traffic therefore clearly needs to be relaxed. The same goes for the obviously unrealistic assumption of constant traffic when we are talking about time intervals of up to 25 years between overlays. Newbery (op.cit.) modifies his model to allow for increasing traffic levels and for allocating part of the deterioration to climatic factors.

50

Allowing for traffic growth, but putting in the extra assumption that there is no road strengthening actions taken in response to this increase in traffic, then the fundamental theorem still holds, as the age distribution still will be uniform. However, this is rather unrealistic. Road authorities will typically reinforce the road in order to keep up its design life under the increased traffic conditions.

The results could be summarised as:

1. *Road damage externalities no longer cancel each other* (as it did with all wear being attributable to traffic), it takes on a positive value. This is also true under the basic assumptions of constant traffic and all damage being attributed to traffic, if maintenance is not responsive to roughness, but arbitrary or periodic. The exception is the case in which demand is linear with a uniform slope coefficient. Then there is still no externality, and the optimal charge will still be equal to the average maintenance cost.

2. *Only a proportion of the average road maintenance cost is now relevant for pricing*.

This is illustrated by the modified expression for the expected marginal social cost (to be compared to the original expression given in Equation 2-15):

Equation 2-22 $$MSC = \frac{\overline{\partial F}}{\partial X} = \frac{\mu \cdot C}{NET} + \frac{1}{T}\int_{0}^{T} \frac{\partial D(X,z)}{\partial X}\, dz$$

The first term is equal to the one found in the original expression, but multiplied by the allocable fraction of maintenance cost (μ). The second term is the road damage externality, where $D(X,z)$ represents the last two terms in the expression

originally given in Equation 2-4[30].

Referring to a study of the Tunisian road network, Newbery concludes that the magnitude of the road damage externality seems to be rather small compared to the magnitude of the overall marginal social cost. This means that the first result above is relatively unimportant compared to the second one.

The allocable fraction of maintenance cost that is attributable to traffic, could, according to Newbery (op.cit.), be estimated based on the initial and final levels of roughness, a climatic parameter[31] (*m*), and the time interval between overlays (*T*)[32]:

Equation 2-23
$$\mu = \left[1 + \frac{mT}{1 - \dfrac{R_0 e^{mT}}{\bar{R}}} \right]^{-1} , \mu \in \{0,1\}$$

The interpretation of the climatic constant *m*, is that the annual roughness increase from weathering is 100*m*%. Newbery reports (without further reference) that the climatic constant typically attains the value 0.025 for humid subtropical climate, 0.01 for arid subtropical areas, and 0.05 for freezing climates. The allocable fraction (*μ*) attains a lower value for the more severe (humid, freezing) areas. Newbery presents a sensitivity analysis with alternative *m*'s and alternative maintenance strategies, resulting in *μ* values ranging from 0.38 to 0.80. The analysis also indicates that the sensitivity to maintenance strategy is more significant in the more severe climates (i.e. with a higher *m*-value).

---

[30]   *X* represents the cumulative traffic up to the current date.

[31]   The parameter stems from a Brazilian study reported by the World Bank in Paterson, W. D. O. (1984)..

[32]   This is appealing because the expression does not include any parameters representing the strength of the road (e.g. structural number), but only parameter that are readily available in many cases.

The calibration of the climatic constant is further discussed in Paterson (1987), concluding that the highest value of *m* occurs in wet freezing climates (represented by Colorado in Paterson's study), and yielding a value of 0.065 (i.e. 6,5% annual increase in roughness). As Nordic climates would tend to fall into the latter category, one might expect this climatic constant to be of considerable magnitude for Norwegian roads. However, this may not necessarily be the case, as suggested by recent Swedish evidence presented in Lindberg (2002b). Weathering may not have any significant impact on deterioration under prevailing Nordic traffic conditions at all[33].

2.5   *ELABORATIONS ON NEWBERY'S MODEL*

Several studies have been based on the model established by Newbery. To some extent Small et al. (1989), reported in 3.2.2, is based on the same framework. Recent European research programmes also refer to Newbery's fundamental theorem of road user charges when dealing with the development of pricing policies for Europe. One of these applications could be found in the nearly finished UNITE-project (see section 3.3.3). Link and Lindberg (2000) and Lindberg (2002b) are both founded on the developments done by Paterson, Newbery and Small.

In the latter reference the model is applied to estimate marginal costs of road maintenance in Sweden. Here, Lindberg reformulates the results of the model, to show that the marginal maintenance costs could be expressed as average costs multiplied by a *deterioration elasticity*. The concrete expressions found in *op.cit.* are:

The marginal maintenance cost of a new or an old[34] road:

---

[33]   Although the Swedish results are based on quite substantial empirical evidence, most of the data collected represent arterial roads with a fairly high durability. Consequently, weathering may still have significant impacts on roads with weaker constructions.

[34]   An old road in this context, is defined as a road with roughness close to the overlay criterion.

Equation 2-24 $MC_{New,Old} = -\alpha\varepsilon AC$

Here $\alpha = (rT)^2 \dfrac{e^{rT}}{(e^{rT}-1)^2}$

and the marginal maintenance cost for an average road, this simplifies to

Equation 2-25 $MC_{Average} = -\varepsilon AC$

Lindberg shows that with a real interest rate of 3-4% and overlay intervals between 5 and 20 years, α approaches unity, hence the expressions for average and old/new roads become the same for plausible values of r and T. In both cases the deterioration elasticity expresses the relative change in the overlay intervals with changes in the annual traffic levels, i.e.

Equation 2-26 $\varepsilon = \dfrac{dT}{dQ}\dfrac{Q}{T}$

If roads deteriorate only because of traffic, i.e. assuming no weathering effect, and the pavement life (T) is constant, then the elasticity will equal negative unity. This brings us back to Newbery's fundamental theorem, saying that average and marginal costs are the same. However, Lindberg argues that the Swedish evidence shows that pavement life changes with traffic levels, in which case the elasticity will be positive, thus breaking the simple relationship in Newbery's model.

54

ESALs per day and direction	SCI[35]						
	50	**75**	**100**	**725**	**150**	**175**	**200**
200	-	-	-	-	-	-	-0.21
300	-	-	-	-	-0.30	-0.40	-0.47
400	-	-	-0.21	-0.37	-0.47	-0.55	-0.60
500	-	-	-0.37	-0.49	-0.58	-0.64	-0.68
600	-	-0.30	-0.47	-0.58	-0.65	-0.70	-0.74
700	-0.10	-0.40	-0.55	-0.64	-0.70	-0.74	-0.77
800	-0.21	-0.47	-0.60	-0.68	-0.74	-0.77	-0.80
900	-0.30	-0.53	-0.65	-0.72	-0.77	-0.80	-0.82

Figure 2-4   Deterioration Elasticity (Lindberg 2002b)

Estimates of the deterioration elasticities from the Swedish case study (see Figure 2-4), show a variation from –0.10 for a high quality road with low traffic levels, to –0.82 for a low quality road with high traffic levels. Combining these figures into the relationships for marginal costs above, shows that one could expect marginal costs to be represented by a rather high proportion of average costs for heavily trafficked weak roads, and vice versa.

Marginal costs for a sample of 249 road sections[36] have been calculated for Sweden. The means, standard deviations, minima, and maxima of central variables describing these road sections, and the resulting average and marginal cost estimates are presented in Figure 2-5.

---

[35]   SCI is the Surface Curvature Index, a measure indicating the solidness of the road. A drop weight is used to measure the deflection (μm) in the centre of the applied loading, and the deflection 300 mm from the centre. SCI represents the difference between these measures. Hence, a low SCI-value indicates a stronger construction.

[36]   These 249 sections are the ones containing "allowed" combinations of SCI and ESALs among the 400 test sections contained in the Swedish long term pavement test.

	Mean	Std.Dev.	Minimum	Maximum
SCI	133.997	44.3632	55.5224	269.104
Vehicles (AADT)	5131.57	2278	1290	10900
Road width (m)	11.72	3.75	7.50	20
ESALs/day/direction	578.94	379.485	137	1320
Overlay cost (k€/km)[37]	4.15	0.98	3.42	7.40
Lifetime (years)	11.81	3.12	3.37	16.97
Deterioration elasticity	-0.43	0.22	-0.80	-0.01
Average cost (€/ESAL-km)	0.00247	0.00179	0.00067	0.01043
Marginal cost (€/ESAL-km)	0.00090	0.00068	0.00002	0.00426

Figure 2-5   Background information and cost estimates for a sample of 249 Swedish road sections (based on Lindberg 2002b)

I have now briefly presented some major contributions to the theory of the marginal costs of road use. After a short presentation of principles related to establishing shadow prices for the various cost components related to road use, I have put the major focus on David Newbery's work related to the estimation of road wear, and some elaborations on this model. Newbery's Fundamental Theorem of Road User Charges, established in the mid 1980s, briefly states that the marginal road damage cost of a vehicle is identically equal to the average maintenance cost allocated in proportion to its number of equivalent standard axles. Gunnar Lindberg shows that the conclusions from Newbery's theorem hold for an average road if pavement life (measured in the number of axles it can withstand) is unaffected by traffic level, and assuming no weathering effect. However, the link between marginal costs and average costs may be broken when studying old or new roads (rather than the average road), and because pavement life empirically seems to vary with traffic levels.

---

[37]   All cost figures were originally presented in SEK by Lindberg. An exchange rate of 1€=8.92 SEK has been applied to convert the figures into Euros.

56

3 STUDIES OF MARGINAL ROAD USE COSTS

3.1 *ALTERNATIVE APPROACHES TO STUDYING MARGINAL ROAD WEAR*

Road infrastructure cost allocation studies has been performed in many states and countries since the 1950s (Jones and Nix 1995). Most of these studies have been carried out to provide information on the total cost responsibility of vehicles, rather than the *marginal* cost responsibility. Still some of the cost allocation studies have been applied for pricing purposes, e.g. in some U.S. states (op.cit.).

By the mid-1990s very few studies included externalities and user costs (i.e. the full social costs) in the cost notion, the same goes for focusing on marginal cost. However, over the last decade there has been an increased focus on the tax-relevant marginal costs of road use. It is also fair to say that the focus on the pricing issue has been stronger in Europe than in most other regions of the world. The United States has a long tradition of performing cost allocation studies, both on the federal level and on the state level. The same goes for Australia, and to some extent Canada. Most of these studies have not traditionally included the marginal social cost approach, but as the interest in developing optimal pricing policies the focus of these studies have changed somewhat over the last 8-10 years. In Europe the focus on pricing has been stronger, at least on the conceptual and general policy-statement level. The European Commission has launched several research programmes with pricing as a primary issue over the last decade. Some of these programmes have also focused on studying marginal road use costs.

The studies that have been performed are based on rather different approaches. This is partly due to the fact that studies have been carried out with different purposes (e.g. planning investments, maintenance or pricing), and within different professional traditions (e.g. engineering or economics). Additionally, the approach

to studying marginal costs of road use has often been decided subject to pure feasibility considerations. Very often the complexity and limited availability of information prohibits the most desirable approaches to be applied.

In the European UNITE-programme (see Section 3.3.3), a suggested typology of marginal cost methodologies is presented (Link and Lindberg 2000). This typology is briefly reproduced in Figure 3-1. The basic classification criteria are:

1. *The choice of functional form* for the costs. The most theoretically adequate approach would be to specify a full cost function including all the relevant cost drivers. The marginal costs are then derived from this cost function. The main, and often prohibiting, disadvantage is the very high requirements this approach puts on data availability for the estimation of the functions. Therefore, many studies end up with a single cost figure (or several figures) as the final product of the study.

2. *The direction of approach* for estimating cost functions. The bottom-up approach focuses on single road sections, and then tries to generalise the results to make them applicable to whole networks. This approach enables the use of engineering based approaches to studying marginal road wear (e.g. like the AASHO Road test presented below). Alternatively one can choose a top-down approach, starting with the real occurred total costs, allocating them by an econometric analysis of cost drivers. This is immediately applicable to the whole road network, but is rather demanding with respect to information needed.

3. *Type of information used*. Generally this could be based on actual costs from accounts and statistics, or it could be based on experiments or simulations. The first category will be directly applicable to whole networks, whereas the latter will have to be aggregated to represent networks. Once again, the AASHO Road Test is the most well known of the experimental studies in the field of experimental road deterioration studies (also called accelerated load tests).

58

In this chapter, the primary aim is to give a brief insight into some of the more influential studies dealing with the marginal costs of road use. Although there are many links between American and European developments in this field, I have chosen to present the American and the European studies separately.

I start with a very brief review of the before mentioned AASHO Road Test. This study is more than 40 years old, but is still an important point of reference for studies in this field. Then Kenneth A. Small, Clifford Winston and Carol A. Evans' proposition on a new highway pricing and investment policy for the United States is presented, partly as a elaboration on the AASHO Road Test, but also partly as a similar framework as the one introduced by David Newbery (see Section 2.4). Finally the (hitherto) last of the three U.S. Federal Highway Cost Allocation Studies is briefly presented, illustrating the typical American tradition for such studies. Finally, I have included a slightly different approach to estimating marginal road wear, represented by a Canadian study.

Among the many European studies in the field I have chosen to focus on two categories: The recent European research programmes linked to transport infrastructure pricing, and Scandinavian studies, which should be the most relevant ones related to the proposed pricing policy for Norwegian roads. Norway and Sweden both have extensive rural road networks with quite low traffic volumes combined with severe climatic environments.

The model established by the Institute of Transport Economics in Oslo forms the starting point for my marginal cost calculations presented in Section 5. I have therefore chosen to give this study a fairly thorough presentation. However, some details about the principles behind the calculations (e.g. Elvik's model for external accident costs) are left to Section 5.

59

Classi-fication criteria	Approach	Characteristics of approach	Scope	Advantages	Dis-advantages
1. Functional form	**Cost function**	Estimation of total cost function and derivating the marginal cost function, either by econometric methods or engineering background.	Whole network, network parts, single sections/lines	Theoretically and empirically adequate approach (first best approach)	High data requirements
	Single cost figure	Linearity assumption for total cost function, pragmatic breakdown approach of variable cost categories to marginal costs.	Whole network, network parts, single sections/lines	Less information necessary, easier (second best approach)	Linearity assumption not confirmed
2. Direction of approach for estimating cost functions	**Bottom-up**	Starting point is costs of basic package, additional costs of successor vehicle categories are stepwise added (discrete approaching of a continuous cost function)	Single lines/sections	Can be done experiment-ally, real world characteristics, use of engineering knowledge	Generali-sation from single sections / lines to whole network complicated, only rough approach to marginal concept
	Top-down	Starting point is real occurred total costs, functional relationship elaborated by econometric analysis of costs and cost drivers (influence factors)	Whole network, network parts, single sections/lines	Easier to elaborate, generalisation better	
3. Type of information used	**Empirical/ac counting cost information**	Observed costs either from official statistics, or from road authorities (ex-post information)	Whole network, network parts, single sections/lines	Information in principle available	
	Experiment-based or simulated engineering information	Observed or theoretical / simulated engineering relationships	Single lines/sections	Proper reflection of engineering knowledge	Generali-sation often difficult, experiments often heavily disputed.

Figure 3-1   Approaches to studying marginal infrastructure costs (Link and Lindberg 2000)

3.2 STUDIES FROM THE UNITED STATES AND CANADA

3.2.1 The AASHO Road Test

The American Association of State Highway Officials (AASHO)[38] carried out a comprehensive accelerated loading test over a two-year period from 1958 to 1960. The AASHO Road Test was performed to provide research data useful in the design of efficient highways at economical cost (Highway Research Board 1962b). The test site was located near Ottawa, Illinois in the USA.

The test facilities consisted of four large road loops, and two smaller loops. Each loop was a segment of a four-lane divided highway whose parallel roadways were connected by a turnaround at each end.. All vehicles assigned to traffic the loops had the same axle arrangement and the same axle load combinations. Each roadway was designed as a succession of structural sections representing different structural pavement designs.

In order to study the effect of climatic factors alone, one of the smaller loops was not subjected to traffic at all. The other loops were subjected to intensive traffic[39] by a number of trucks loaded with concrete blocks running at 35 mph. All the way, the states of the road sections were monitored (measuring cracks, roughness, rutting etc.). Maintenance actions were kept at a minimum. Some of the test sections reached "failure" under the two-year period, and some did not.

The primary aim of the test was to evaluate different structural pavement designs[40], and their performance with respect to durability under various traffic loads. In addition to registering traffic loads and pavement developments, climatic

---

[38]   Now named AASHTO.

[39]   The actual loadings comprised an 18 hour and 40 minutes period each day, 6-7 days a week, totalling 1 114 000 axle load applications over the 25 month period.

[40]   Different bridge designs were also studied, but I will focus this presentation on the results related to pavement performance.

factors, like precipitation and frost measurements, were also logged over the test period.

The findings of the project were presented in the form of estimated equations and graphs, representing the relationship between traffic loads and pavement performance. The most well known formula from the project is the so-called "fourth power law", indicating that doubling the axle weight increases the road damages by a factor of 16 ($=2^4$). This result has been much used and debated, and will also be referred to several times in this thesis (e.g. in the subsequent section, dealing with Kenneth Small's proposed new highway investment and pricing policy).

Several new concepts and notions were introduced during the AASHO Road Test, many of which has become very central in subsequent studies of road wear. Two of these new notions are *the serviceability* of a road, and the *equivalent axle load*.

The serviceability of a road

A road section can be characterised by measuring many different features. However, the most interesting feature of the road is its ability to serve traffic. In order to describe this ability, there is a need for combining may of the technical characteristics of the road. The *Present Serviceability Index* (PSI) is a composite index based on measurement of longitudinal profile variations, the amount of cracking and patching, and transverse profile variations (rutting). The weights on each factor is determined by a statistical correlation (multiple regression) between observed values of these factors and subjective scores recorded from a Rating Panel established under the AASHO Road Test. Longitudinal profile variation (represented by the logarithm of the slope variance) received the heaviest weight among these factors influencing serviceability.

The following equation was used to determine the PSI in the AASHO Road Test (Highway Research Board 1962a):

62

Equation 3-1 $p = 5.03 - 1.91 \cdot \log(1 + \overline{SV}) - 0.01\sqrt{C + P} - 1.38\overline{RD}^2$

in which

p   = the present serviceability index (PSI)

$\overline{SV}$ = the mean of the slope variance in the two wheelpaths (roughness)

C+P   = a measure of cracking[41] and patching[42] in the pavement surface

$\overline{RD}$ = a measure of rutting in the wheel-paths[43]

Equivalent axle loads (ESALs)

The typical traffic pattern of a road will comprise a combination of axles with different loads. In order to combine these different axles into a composite measure representing all the different loads (and axle combinations), the term *equivalent axle load* (ESAL) was established under the AASHO Road Test.

ESAL is a standardisation unit used to bring all kinds of axle configurations down to one denotation. One ESAL is defined as one 18 000 pound (80kN) load on a single axle with dual tires. Axles that contain more or less weight (or with different tire configurations) are related to the ESAL using *load equivalency factors* (LEFs). The LEFs are calculated based on studies of the trade-off between axle load/configuration and road wear (e.g. the fourth power law of the AASHO Road Test).

---

[41] Cracking (C) is measured as the area, in square ft per 1000 square ft of pavement surface, exhibiting class 2 or class 3 of cracking. Class 2 cracking is defined as that which has progressed to the stage where craks have connected together to form a grid-type pattern. Class 3 cracking is that in which the bituminous surfacing segments have become loose.

[42] Patching (P) is the repair of the pavement surface by skin patching or deep patching expressed in square ft per 1000 square feet of pavement surfacing.

[43] Rut depth (RD) is defined as the mean depth of rut in both wheepaths of the pavement where the rut is the depression under the center of a 4-ft straightedge. The mean rut depth was estimated by sampling in each wheelpath at 25-ft intervals.

There is a huge literature discussing many aspects of The AASHO Road Test. Apart from the fact that this is a very old study, representing old technology, both with respect to pavement design and vehicle design, much of the debate revolves around the transferability of the results to other situations. In many cases the application of the test results clearly represents extrapolations beyond the actual observed variations in the study. Even though the accelerated loadings represented quite high traffic volumes, many of the pavements did *not* reach the failure criteria under the test period. Accelerated tests like this will always suffer from the fact that the interplay between climatic factors and traffic may not be representative of "real world" applications.

Transferability with respect to climatic factors is always an issue. The test area had an average annual precipitation of 864 mm, of which about 64 mm occurs as 635 inches of snow (op.cit.). The average mean summer temperature was 24° C and the equivalent figure for winter was –3° C.[44] The precipitation figures should be fairly comparable to the situation in East Norway, but they are significantly lower than typical figures for West and North Norway. Temperatures seem to be more extreme than what is typical for Norwegian conditions.

One of the studies that not only criticizes the AASHO Road Test, but also suggests improvements, could be found in (Small et al. 1989). This study is presented and commented upon in the following section.

---

[44]   Figures converted from inches and Fahrenheit by the author.

3.2.2   Small's Proposed New Highway Pricing and Investment Policy

A comprehensive and consistent proposal with little factual impact
Kenneth A. Small, Clifford Winston and Carol A. Evans proposed, in 1989, a comprehensive new pricing and investment policy for the US highway sector (Small et al. 1989). The proposal addressed a number of important issues related to the socially efficient management of the highway system:

- Better **pricing regimes** for road use, both related to free-flow conditions, and to congested road networks
- Improved framework for decisions related to optimal investments related to highway design and optimal maintenance
- Financial and political viability of the proposed pricing and management schemes

The main achievement of the study is its integrated treatment of all these closely related issues. The proposal should therefore form a good foundation for changing the pricing and investment regimes related to the US highway sector. However, more than a decade after the publication of the proposal, little has changed and only a very limited part of the proposed actions has been implemented[45].

The relationship between road pricing and investment

> *"Road User Charges and optimal investment, though often treated separately by policy analysts, are facets of the same problem: both are aimed at minimizing the total costs of building, maintaining, and using a road system." (Small et al. 1989)*

There is a close link between marginal cost pricing and decisions related to road investments and optimal maintenance. This is perhaps best illustrated by the obvious trade-off between road design and user costs. The better the road, the

---

[45]   The exception may be the state of Oregon, which has adopted a fourth-power axle-load based tax rate for extra heavy vehicles.

smaller the marginal costs of using the road. An optimal investment policy is directly related to the marginal user costs, because one should invest in road quality as long as the marginal savings (i.e. reduced costs), are bigger than the incremental investment.

In uncongested road networks the marginal external costs related to road use are

- Road wear costs
- Accident costs, and
- Environmental costs

The latter two are omitted by Small et al., so the focus is set on road wear costs. In congested areas a fourth component of road use becomes important as well: External time costs, sometimes denoted 'marginal capacity costs'. Small et al. build a model for estimating these costs. This model is presented in the subsequent sections, and illustrated schematically in Figure 3-2.

A model for pricing and investment, based on road-wear and congestion costs
This model is built on empirical knowledge of the trade-offs between road use and road wear. Many empirical studies have been carried out in this field, both based on 'laboratory work' and on 'test roads'. The AASHO road test is by far the most influential among these tests. This is a comprehensive test-scheme carried out in the late 1950s in Illinois in which different combinations of pavement designs and truck loads were tested in order to establish trade-offs. Kenneth Small and Clifford Winston (Small and Winston 1988) have re-estimated these test-results using modern statistical tools, and this work forms the basis for the applied trade-offs between axle loads and road wear in (Small et al. 1989).

66

Step 1: Assess typical time-intervals between resurfacing for each road class
Because of the wide variation in climates and soil types, the typical time-intervals are used instead of basing the calculations on actual road designs (layer thicknesses).

Step 2: Calculate the total no. of ESALs on the road network
Based on:
- A re-estimation of the AASHO road test data (i.e. a 3rd power rule)
- FHWA estimates of truck vehicle miles on each road class
- An approximation of the distribution of ESALs on each surface type

Step 3: "Back-calculate" the typical road thicknesses (i.e. Structural Numbers) for each road class
Based on the re-estimated AASHO-model, the assumed time-intervals between resurfacing and the traffic load (ESALs), the road thicknesses are calculated for each road class.

Step 4: Calculate marginal maintenance costs per ESAL-mile for each functional road class
Depending on the calculated road thickness, functional class and surface type, the maintenance costs are calculated for 40 road classes divided into 5 traffic volume levels.

Step 5: Calculate marginal road wear costs for each vehicle class for the existing road design
Based on the estimated costs per ESAL-mile, and the reestimated "AASHO"-factors marginal costs per vehicle mile can be calculated.

Step 6: Re-estimate optimal road durability under the demand effects of a new highway pricing, investment, and maintenance regime

Step 7: Calculate marginal road wear costs for each vehicle class under an optimal road design and maintenance regime
Based on the re-estimated costs per ESAL-mile, and the reestimated "AASHO"-factors marginal costs per vehicle mile under an optimal road design regime can be calculated.

Figure 3-2   The Small-model for estimating marginal road wear costs

The model is further based on theories in the economics of congestion pricing and investment rules developed by Dupuit, Pigou, Knight, Mohring and others (see Winston 1985). Small et al. develop this theoretical framework further, relating road wear charges to scarce durability (i.e. pavement thickness), and congestion charges to scarce road capacity.

The full explanation of the model is given in Small et al. (1989) pp. 22-36, I will only cover the major equations and conclusions here. The point of departure is the maximation of the annual consumer surplus from travel less annualised costs. It is maximised subject to hourly traffic flow volumes (*q*), road capacity (W) and road durability (D):

Equation 3-2

$$\underset{q,W,D}{Max}\ y\sum_h\sum_i\int_0^{q_{ih}}P_{ih}(q')dq'_{ih} - y\sum_h\sum_i q_{ih}u_{ih}(V_h,W,Q,D) - rM(Q,W,D) - rK(W,D)$$

Comments on the maximation problem:

- $P_{ih}(q)$ represents the inverse demand functions for vehicle class *i* in hour *h*. The price (*P*) includes user-incurred time and money costs plus user charges. This means that the first part of the expression above illustrates the total annual user benefits for all vehicle classes and all time periods.

Subtracted from these benefits are *three cost elements*:

- The first cost item is the *user costs* (*u*). These costs comprise two cost elements: The first depends on road deterioration, which in turn is dependent on the number of annual traffic loadings (Q) and the road durability (D). The second element of the user costs is related to congested areas only and

depends on the hourly traffic volume (V_h) and the road's capacity $(C)^{46}$. Total user costs are obtained by multiplying by traffic volume (q), summing over vehicle types (i) and hours a day (h), and finally multiplying the double sum by y days a year.

- The second cost item describes the present value of the *cost of maintenance* for one mile of highway. These maintenance costs are expressed as an infinite series of traffic-responsive maintenance actions, whose costs are discounted by a rate r to represent the present value. These costs are dependent on traffic loadings Q, road capacity W, and road durability D.

- Accordingly, the third cost item describes the present value of *capital costs* for one mile of highway. These costs are generally dependent on the discount rate r, road capacity W (i.e. width / number of lanes) and road durability D (i.e. layer thicknesses and qualities).

First order conditions for this maximation problem are:

Equation 3-3 $$P_{ih} - u_{ih} = \Phi_i \left(\frac{\partial c_h}{\partial V_h} \right) + \mu_i \left[r\partial \frac{(M+U)}{\partial Q} \right], \quad i=1,...,I \quad h=1,...,H$$

This is the *pricing rule* indicating that a user charge is needed to make up the difference between the costs u_{ih} already borne by the user, and the "social" price P_{ih} required to adjust demand to its optimal level. The components of this required charge is made up of the expressions on the right hand side of the equation above. The first element represents the *congestion charge*, proportional to ϕ_i which is the number of capacity car-equivalents, vehicle i represents. V_h represents the hourly traffic volume measured in capacity car-equivalents, and c_h are the hourly

---

[46]   Congestion could also depend on traffic volumes at earlier times, because of queing; Small claim that this would complicate the equations, but not the concept behind them.

aggregate congestion costs. The second element of the optimal charge in the pricing rule represents the *charge for road wear*. μ_i represents the number of ESALs of vehicle *i*. The rest of the expression illustrates the change in maintenance (*M*) and user costs (*U*) as traffic loads (Q) change.

The second first order condition yields:

Equation 3-4 $$r\frac{\partial(M+K)}{\partial W} = -\frac{\partial c}{\partial W}$$

This condition illustrates the *optimal capacity rule*. Capacity is here represented by the width of the road (*W*). The marginal agency costs of expanding capacity (maintenance and capital, i.e. the left hand side of the equation) should equal the marginal savings in user costs at the optimum (the right hand side).

Finally, the third first order condition for the maximation problem yields:

Equation 3-5 $$\frac{\partial K}{\partial D} = -\frac{\partial(M+U)}{\partial D}$$

This could be labelled the *optimal durability rule*, indicating the optimum condition for providing more solid roads. Once again, durability should be enhanced until the marginal savings in maintenance and capital costs equal the marginal capital costs of the increased durability.

My primary interest in this thesis is related to the optimal pricing rule, given in Equation 3-3 above. I will therefore elaborate on Small *et. al.*'s development of this relationship, also putting the main emphasis on the latter part of the right hand side, i.e. the marginal road wear (thus leaving out the congestion charge part). In *op.cit.*, the road wear relationship is developed from a renewed analysis (Small

70

and Winston 1988) of the AASHO Road Test data, applying improved estimation techniques on these 40 year old test data. Since many of the road sections in the test did not reach failure under the accelerated loading test, the database represents a censored sample. Small & Winston applied a Tobit estimation method based on the following equation:

Equation 3-6 $N = A_0 (D+1)^{A_1} (L_1 + L_2)^{-A_2} (L_2)^{A_3}$

Here N is the number of axle passages that will cause the pavement to reach the overlay criterion. D represents the durability of the pavement, here represented by the Structural Number of the flexible pavements. L_1 represents the axle weight, and L_2 is a dichotomous variable acquiring a value of 1 for single axles and 2 for tandem axles. The parameters A_0 to A_3 are estimated in the Tobit regression procedure.

The estimation results for flexible pavements are reproduced in Figure 3-3. The so-called "4th power rule" is represented by the A_2-coefficient. The results for *rigid* pavements (not reported here), suggests a "3rd power rule" as a result from the re-estimation of the AASHO Road Test data. However, for flexible pavements there still may be a case for keeping the "4th power rule" based on Small & Winston's re-estimations, although there is a quite significant reduction in the estimates (from 4.79 in AASHO to 3.65 in Small & Winston).

Coefficient A_1 represents the relationship between pavement life and pavement "thickness" (i.e. Structural Number). Here the re-estimations performed by Small & Winston indicate a less steep relationship than the original AASHO estimates. Calculating the expected lifetimes (number of cumulative ESALS to failure) of thicker pavements based on the new estimates, shows much shorter lives for thick pavements than the original predictions from AASHO. This finding is, according to Small et al. (1989), corroborated by several other studies. However, the

71

application of the results for thick pavements represents extrapolation well beyond the range of direct observation, as none of these pavements actually reached failure under the AASHO test runs.

Coefficient	Small & Winston estimate	Corresponding[47] AASHO estimate
ln A_0	12.062	13.65
A_1	7.761	9.36
A_2	3.652	4.79
A_3	3.238	4.33
Number of observations	284	
Number of censored observations	45	
Std. error of regression	0.651	
Std. error of prediction	0.629	0.673

Figure 3-3   Results of the re-estimation of AASHO Road Test in (Small and Winston 1988)

The overall fit of the model, here represented by the standard error of prediction, does not indicate a much better fit from the re-estimation procedure (as was the case for the rigid pavements).

The re-estimations on the AASHO Road Test data are subsequently put into the further analysis of the marginal maintenance costs in Small et al. (1989). In line with assumptions made in Paterson (1987) and in Newbery (1988b), Small *et.al.* assumes that pavement roughness grows linearly with cumulative ESALs and exponentially with time[48]. The relationship between intervals between resurfacings (*T*), and durability, traffic loads and aging effects could then be expressed by Equation 3-7.

---

[47]   Note that the dependent variables in Small & Winston and in AASHO are slightly different. The pairs of estimates are therefore not strictly comparable.

[48]   For a closer presentation of the models developed by Paterson and Newbery, see section 2.4.

Equation 3-7 $$T = \frac{N}{\lambda Q} e^{-mT}$$

Here N still describes the pavement durability measured in the number of ESALs it could withstand before repaving is required. λ is the proportion of the annual traffic loads (Q) that occur in the most heavily travelled lane. This relationship is entered into an expression for a marginal maintenance costs, based on a differentiation of the expression for the present value of an infinite sequence of regular overlay expenditures. This yields the following expression for marginal maintenance costs given in Small et al. (1989):

Equation 3-8 $$MC_m = \alpha \beta \left(MC_m^0 \right)$$

Here

$$\alpha = (rT)^2 \frac{e^{rT}}{\left(e^{rT} - 1 \right)^2}, \quad \text{and} \quad \beta = \frac{e^{mT}}{1 + mT}$$

$$MC_m^0 = \frac{C}{\left(\dfrac{N}{\lambda} \right)}.$$

m is the aging effect on pavement roughness

r is the rate used for discounting

C is the cost of an overlay

In Small and Winston (1988), α is shown to lie between 0 and 1, and β will be greater than 1. This means that marginal maintenance costs will be *higher* as aging effects become more prominent (i.e. with increasing m), and not *lower* as assumed in many studies!

This concludes Small *et.al.'s* elaboration on the maintenance cost part of the

pricing relevant marginal costs in Equation 3-3. The other element in this equation (apart from the congestion costs) is the user costs. With repaving with fixed time intervals, the average user costs will vary in a cyclical way (see Figure 3-4) because user costs increase with increasing roughness. Roughness increases up until the road is repaved, then it falls to the initial level. Evidence suggests that the shape of the average user cost curve should be convex (as the solid lines in the figure indicates), but Newbery and Paterson concludes that a linear approximation is plausible for the relevant intervals of roughness (indicated by the dotted lines in the figure).
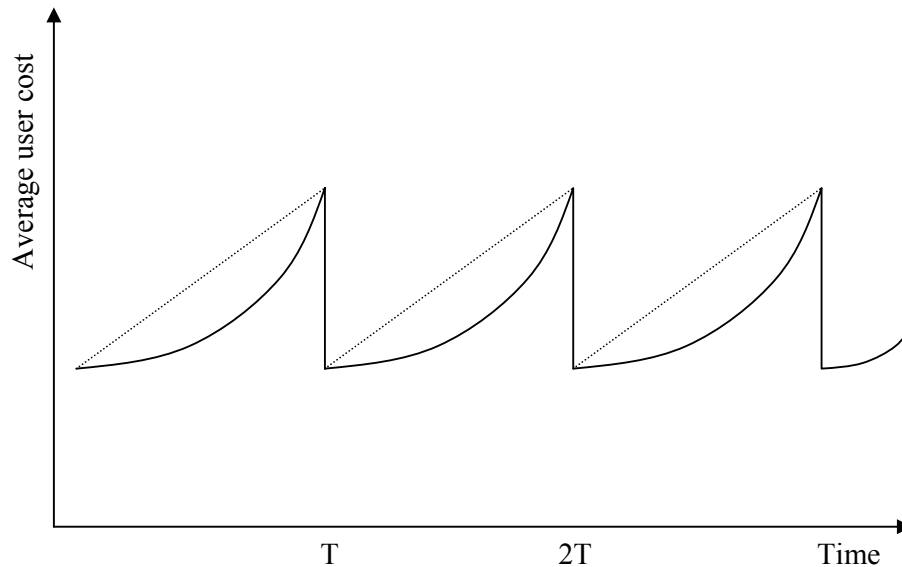
Figure 3-4   Cyclical variation of average user costs with time

Assuming traffic growth is constant at a rate *g*, Small *et. al.* concludes that the annualised marginal user-costs MC_U could be expressed by the following equation:

Equation 3-9 $$MC_U = r\frac{dU}{dQ} = r\frac{dU}{dT}\frac{dT}{dQ}$$

from Equation 3-7 we can derive that

Equation 3-10
$$\frac{dT}{dQ} = -\frac{T}{Q} \cdot \frac{1}{1+mT}$$

and according to Small *et. al.* the change in user costs resulting from a change in overlay intervals T could be expressed by

Equation 3-11

$$\frac{dU}{dT} = \frac{rV_0v_1}{y}\left[-\frac{600}{y^5} + ye^y\left(e^y-1\right)^{-2}\left(1+\frac{5}{y}+\frac{20}{y^2}+\frac{60}{y^3}+\frac{120}{y^4}\right)+\left(e^y-1\right)^{-1}\left(\frac{5}{y}+\frac{40}{y^2}+\frac{180}{y^3}+\frac{480}{y^4}\right)\right]$$

where

y	=	aT
a	=	r-g
$V(t)$	=	*Traffic volume in year t*
$v(t)$	=	*Average user cost in year t*

Combining the three last equations (i.e. inserting Equation 3-10 and Equation 3-11 into Equation 3-9) yields Small *et. al.'s* expression for the marginal user costs with traffic growth at rate g and responsive repaving. This combined expression and Equation 3-8 could then be entered into the pricing rule in Equation 3-3 (along with an expression for the congestion costs) to yield the total pricing relevant marginal costs.

3.2.3   The US Federal Highway Cost Allocation Study

In US DOT (1997a), the latest Federal Highway Cost Allocation Study (FHCAS), conducted by the Federal Highway Administration, is reported. This study is based on a series of previous cost allocation studies at the federal level, the previous one published in 1982. The aim of this study is to relate highway-related

costs to different user groups, and to compare these costs to the user fees actually paid by the various groups.

Road wear	TASK 1: Which costs are attributable to road *use* ? -------------- TASK 2: How should these costs be allocated to different vehicle classes ?	ALTERNATIVE 1: Calculate the *average* costs per equivalent standard axle (ESAL)	RESULT I: Equity-based costs per vehicle-kilometre
Con-gestion			
Accidents		ALTERNATIVE 2: Calculate the *marginal* costs per equivalent standard axle (ESAL)	RESULT II: Efficiency-based costs per vehicle-kilometre
Environ-ment			

Figure 3-5   Schematic illustration of the calculation of average and marginal external costs per vehicle kilometre

Such a comparison could be made in an *equity* setting, where the main issue would be to assess the "fairness" of the current charging system, based on an underlying "*cost occasioned*" principle ("Result I" in Figure 3-5 above). Alternatively one could focus on the *efficiency* of the current pricing regime as an allocative device, ensuring the internalisation of external costs. This would necessitate the calculation of social marginal costs, as indicated by "Result II" in the figure.

The calculation of costs is quite different in the two alternative settings. In the first case, one would make the calculations in a "total cost" setting, in the second approach "marginal costs" should be addressed. The same distinction goes for the charges. Only marginal charges that vary with actual road use should be considered in the latter case. FHCAS focuses to some extent on both these issues, although the main emphasis is put on the question of equitability.

Traditionally the FHCASs have dealt with the so-called *agency costs* of highway

76

provision only (i.e. the costs of building and maintaining the road network). In the 1982 study this was supplemented by an appendix also considering the other external costs of road use, mainly related to accidents and environmental impacts. In the 1997 study these costs are, to a very high extent, incorporated in the reported figures[49].

The basic steps of the 1997 FHCAS

Step 1: Supplementation of data from the HPMS sections with climatic data and enhanced traffic data

The HPMS-database contains information about 99 000 road-sections from 45 US states on:

- Number of lanes
- Type of pavement
- Pavement thickness
- Current pavement condition
- Average daily traffic
- Percentage of trucks
- Predicted 20-year traffic levels
- Climatic zone
- Rudimentary information on pavement base

This was supplemented with state-characteristic data on:

- Freeze-thaw cycles
- Freezing index
- Thorntwaite moisture index
- Modulus of subgrade reaction
- Average annual rainfall
- Thickness of base

---

[49]   Some of the cost elements related to air emissions were presented in a later appendix to the study, US DOT (2000)

- Estimates of annual vehicle miles travelled (VMT) by vehicle class, highway functional class and state
- Operating weight distributions on groups of highway types in groups of states
- Axle weights for the midpoint of each weight group for each vehicle class

Step 2: Deterioration models used for determining the age of the road section

A set of mechanistic-empirical deterioration models was used to determine the age of the road sections. The models cover:

- Traffic related PSR-loss
- Expansive clay-related PSR-loss
- Fatigue cracking
- Thermal cracking
- Rutting, and
- Loss of skid resistance

The models are partly based on the results of the US Long Term Pavement Performance (LTPP) Study. (See Step 4 for an example of the models.)

Step 3: Estimated traffic-loads assigned to the road-sections until rehabilitation-triggering levels of distress is reached

Using the same deterioration models in Step 2, an iterative procedure is conducted to estimate the remaining pavement life. A composite measure of pavement condition was established, called the *overall pavement condition score (OPCS)*. See Figure 3-6 for an explanation of this measure.

A typical one-year traffic load is assigned to the road section, and physical distress parameters are calculated. This procedure is repeated until one of the

78

following rehabilitation-triggering criteria is met:

- PSR ≤ 2,5 or

- OPCS ≤ 10

The calculation of the Overall Pavement Condition Score (OPCS)
The OPCS is calculated based on the following procedure:

First the following "deduction point maxima" for flexible pavements are established:

- PSR Loss: 50 points
- Cracking: 25 points
- Rutting: 30 points
- Skid Resistance Loss: 20 points

Then the calculated physical distress measurements were converted into an index, ranging from zero for a newly-installed, distress-free pavement, to a value of one at the defined critical distress level. Critical distress levels are:
- PSR=2,5
- 20% fatigue cracking
- 1,5 inch rut depth

For example, the critical PSR-value (i.e. the value that triggers rehabilitation of the pavement) is 2,5. The maximum value for a distress-free pavement is assumed to be 4,5. The index for PSR would then be 0 when PSR equals 4,5, and 1 when PSR reaches the critical value 2,5. By linear interpolation we can find that e.g. a PSR-value of 3,5 corresponds to an index value of 0,5.

This is done in the same way for all the four deterioration measures noted above. Then the index values are multiplied by the deduction point maxima above, and the sum of these products is finally deducted from 100 to obtain the estimated OPCS-value.

Figure 3-6   The calculation of the Overall Pavement Condition Score

Step 4: Assigning relative cost responsibility for rehabilitation costs to vehicle class

The mechanistic-empirical models for the different distress mechanisms noted in Step 2 were also used for allocating cost responsibility to vehicle classes. The PSR loss is partly regarded as traffic related, and partly non-load related. Fatigue cracking is regarded load-related, as is rutting and loss of skid resistance. Thermal cracking is regarded non-load related.

An example of the models used is the NAPCOM Flexible Pavement Traffic-related PSR loss model:

Damage=(LEFS/RHZ)$^{(BEZ/(LEFS/RHZ))}$

where:

- LEFS= the summation of accumulated load equivalents, RHZ or RH(ax)
- RHZ= the number of applications to failure (standard axles)
- RH(ax)=applications to failure of axle load "ax"
- BEZ=coefficient of exponent (beta)

RH(ax) is calculated from equations that include the following environmental and design factors:

- Axle load (kips)
- Axle type (single or tandem)
- Thickness of asphalt layer (inches)
- Subgrade modulus (psi)
- Structural Number of pavement base (as in AASHTO Pavement Design Guide)

Different equation-parameters were used to reflect the four different climatic zones (i.e. wet freeze, dry freeze, wet no-freeze and dry no-freeze).

Step 5: Calculation of absolute levels of marginal cost responsibility for vehicle classes by road functional class

The FHCAS was primarily conducted to allocate agency costs to vehicle classes with *equity* as the governing aim. However, a supplementary study was conducted

to provide *marginal user costs* as well, enabling the current charging regime to be evaluated in an allocative efficiency setting too. This study comprised marginal costs related not only to road wear and congestion, but also to accidents and environmental hazards (the latter included in a later addendum to the report). The reports and appendices of the FHCAS do not describe the actual procedure for singling out the marginal costs related to road wear, but presumably the costs related to road *use* were isolated and calculated in the same manner as the agency costs. As noted above (Step 4), some distress mechanisms are load related, and some are not. This should provide a basis for isolating the truly *marginal* costs related to wear.

Vehicle class / Registered Weight (GVW[50])	Cents per Mile
Autos	0.80
Pickups/Vans	0.76
Buses	3.20
SUT[51] <25 001 pounds	2.20
SUT 25 001-50 000 pounds	5.46
SUT >50 000 pounds	18.12
CT[52] < 50 001 pounds	3.43
CT 50 001-70 000 pounds	5.21
CT 70 001-75 000 pounds	7.62
CT 75 001-80 000 pounds	8.65
CT 80 001-100 000 pounds	15.32
CT >100 001 pounds	20.28

Figure 3-7   Federal highway program costs allocated by vehicle class (Source: US DOT 2000)

Generally this represents a top-down approach because the procedure is based on accounted expenditures for whole road networks, rather than actual road wear on single roads. However, the *distribution* of the marginal cost responsibility of the vehicle classes, are based on empirical observations of road wear (e.g. the LTPP-

---

[50]   GVW=Gross Vehicle Weight

[51]   SUT= Single Unit Trucks

[52]   CT= Combination Trucks

project noted above).

Results from the 1997 FHCAS

The main output from the 1997 FHCAS could be illustrated by tables presenting estimates of costs and charges paid by various highway user groups. In order to make the figures comparable to other numbers presented in this thesis, I have transformed some of the figures to metric units of measurement.

In Figure 3-7 the Federal Highway Program Costs are allocated to autos, pickups/vans, buses, and different configurations and weight classes of trucks.

Tax type	2000 tax rate
Fuel	
- Gasoline	18.3 cents per gallon
- Diesel	24.3 cents per gallon
- Alternative fuels	0-18.3 cents per gallon
Vehicle Excise Tax	
- Heavy Trucks >33 000 pounds GVW - Trailers >26 000 pounds GVW	12 percent of retail sales for new vehicles (trucks, tractors or trailers)
Tire Tax	
- 41 to 70 pounds	15 cents per pound over 40 pounds
- 71 to 90 pounds	$4.50 plus 30 cents per pound over 70 pounds
- over 90 pounds	$10.50 plus 50 cents per pound over 90 pounds
HVUT[53]	
- Annual tax on vehicles, 55 000 pounds GVW or more	$100 plus $22 per 1 000 pounds over 55 000 with an annual cap of $550

Figure 3-8   2000 Federal Highway User Tax Rates (Source: US DOT, 2000)

The current taxation regime related to highway use is presented briefly in Figure 3-8. These payments, transformed into average tax rates per vehicle mile, should then be compared to the allocated agency costs presented in Figure 3-7. This comparison brings about the so-called *equity ratios* presented in Figure 3-9. Here,

---

[53]   HVUT=Heavy Vehicle Use Tax

not only the taxes and costs at the federal level are considered, but also equivalent figures for the state and local governmental levels. The question of equitability will be answered quite differently depending on what level of government one considers.

Vehicle class (Gross weight)	Equity Ratios				
Level of Government	Federal	State	Federal and State	Local	All levels of Govern-ment
Autos	0.9	1.0	1.0	0.1	0.7
Pickups/Vans	1.2	1.2	1.2	0.1	0.9
Buses	0.1	0.8	0.5	0.0	0.4
SUT[54] <25 001 pounds	1.4	2.2	1.9	0.1	1.5
SUT 25 001-50 000 pounds	0.6	1.0	0.8	0.0	0.6
SUT >50 000 pounds	0.5	0.5	0.5	0.0	0.4
CT[55] < 50 001 pounds	1.4	1.7	1.6	0.1	1.3
CT 50 001-70 000 pounds	1.0	1.3	1.1	0.1	0.9
CT 70 001-75 000 pounds	0.9	1.1	1.0	0.1	0.8
CT 75 001-80 000 pounds	0.9	0.9	0.9	0.1	0.8
CT >80 000 pounds	0.6	1.0	0.9	0.0	0.7

Figure 3-9   Ratios of 2000 User Charges to Allocated Costs by Vehicle Class for Different Levels of Government (Source: US DOT, 1997a)

The general picture from Figure 3-9 is that autos, and medium sized combination trucks, cover their calculated cost responsibility with respect to agency costs at federal and state level. Pickups/vans and the lightest classes of single unit and combination trucks seem to pay more than an equitable share of these costs at the same level of government. Buses and the heaviest truck types do not cover a fair share of federal and state agency costs through taxes levied at these government levels. Adding the agency costs and taxes paid at the local governmental level, makes all vehicle groups but the lightest trucks pay less than their equitable share

---

[54]   SUT= Single Unit Trucks

[55]   CT= Combination Trucks

of expenses. This is due to the fact that local governments levy a very small share of total user fees. Actually, these fees only cover about 10 percent of local highway agency costs.

Vehicle class / Highway class	US Cents per Vehicle Mile					
	Pavement	Con-gestion	Crash	Air Pollu-tion[56]	Noise	Total
Autos / Rural Interstate	0	0.78	0.98	1.14	0.01	2.91
Autos / Urban Interstate	0.1	7.70	1.19	1.33	0.09	10.4
40 kip[57] 4-axle S.U.[58] truck / Rural Interstate	1.0	2.45	0.47	3.85	0.09	7.86
40 kip 4-axle S.U. truck / Urban Interstate	3.1	24.4	0.86	4.49	1.50	34.4
60 kip 4-axle S.U. truck / Rural Interstate	5.6	3.27	0.47	3.85	0.11	13.3
60 kip 4-axle S.U. truck / Urban Interstate	18.1	32.6	0.86	4.49	1.68	57.7
60 kip 5-axle Comb.[59] truck / Rural Interstate	3.3	1.88	0.88	3.85	0.17	10.0
60 kip 5-axle Comb. truck / Urban Interstate	10.5	18.3	1.15	4.49	2.75	37.2
80 kip 5-axle Comb. truck / Rural Interstate	12.7	2.23	0.88	3.85	0.19	19.8
80 kip 5-axle Comb. truck / Urban Interstate	40.9	20.0	1.15	4.49	3.04	69.6

Figure 3-10   2000 Marginal pavement, congestion, crash, air pollution and noise costs for illustrative vehicles under specific conditions (Source: US DOT 2000)

In Figure 3-10, estimated marginal costs from the US 1997 (2000) FHCAS are reproduced. Here *pavement costs* represent the contribution of a mile of travel by different vehicles to pavement deterioration and the costs of repairing the damage. *Congestion costs* reflect the value of added travel time due to additional small increments of traffic. *Crash costs* include medical costs, property damage, lost productivity, pain and suffering, and other costs associated with highway crashes. *Air pollution costs* are measured in terms of the cost of premature death, illness,

---

[56]   Note: Air pollution costs are averages of costs of travel on all rural and urban highway classes, not just Interstate. Available data do not allow differences in air pollution costs for heavy truck classes to be distinguished.

[57]   KIP=Kilopounds ~ 0.54 metric tons, e.g. 40 kips ~18.16 metric tons.

[58]   S.U. = Single Unit.

[59]   Comb. = Combination vehicle (Tractor+trailer).

and other effects of various highway-related emissions. *Noise costs* reflect changes in the value of adjacent properties caused by motor vehicle-related noise.

Vehicle class / Highway class	US Cents per Vehicle Mile		
	"Federal share" of Marginal costs[60]	Federal Program Costs	Federal Fee
Autos / Rural Interstate	0.81	0.3	0.6
Autos / Urban Interstate	2.91	1.4	0.6
40 kip 4-axle S.U. truck / Rural Interstate	2.20	1.7	8.5
40 kip 4-axle S.U. truck / Urban Interstate	9.64	3.7	8.5
60 kip 4-axle S.U. truck / Rural Interstate	3.72	7.1	9.2
60 kip 4-axle S.U. truck / Urban Interstate	16.18	12.7	9.2
60 kip 5-axle Comb. truck / Rural Interstate	2.82	2.7	6.4
60 kip 5-axle Comb. truck / Urban Interstate	10.44	6.7	6.4
80 kip 5-axle Comb. truck / Rural Interstate	5.56	7.8	6.9
80 kip 5-axle Comb. truck / Urban Interstate	19.50	17.5	6.9

Figure 3-11   Comparison of assumed federal share of marginal highway costs to federal agency costs and federal user fees (partly based on US DOT, 1997a and on US DOT, 2000)[61]

In Figure 3-12 the picture from Figure 3-11 becomes clearer. The estimates of marginal costs exceed the allocated agency costs for all vehicle classes but the 80kip 5-axle combination truck on a rural interstate. The line in the diagram indicates the current level of Federal taxation. If one considers all these taxes to be marginal, it seems that with the current level of taxation no vehicle class covers its marginal cost responsibility in the urban interstate case. The picture is the other way around for the rural interstate traffic. Here all vehicle classes seem to cover their cost responsibility under the current taxation regime. Many vehicle classes

---

[60]   Including Accident Costs presented in US DOT (2000).

[61]   The assumed federal share of marginal costs is 28 percent of total marginal costs, due to the fact that 28% of total Highway User Revenues come from Federal user fees (cf. US DOT (1997a) p VI-22). Assumed marginal costs are collected from US DOT (2000), Federal Program Costs and user fees are collected from US DOT (1997a), Table VI-22.

pay taxes that exceed the estimated marginal costs by as much as 100-300 percent.
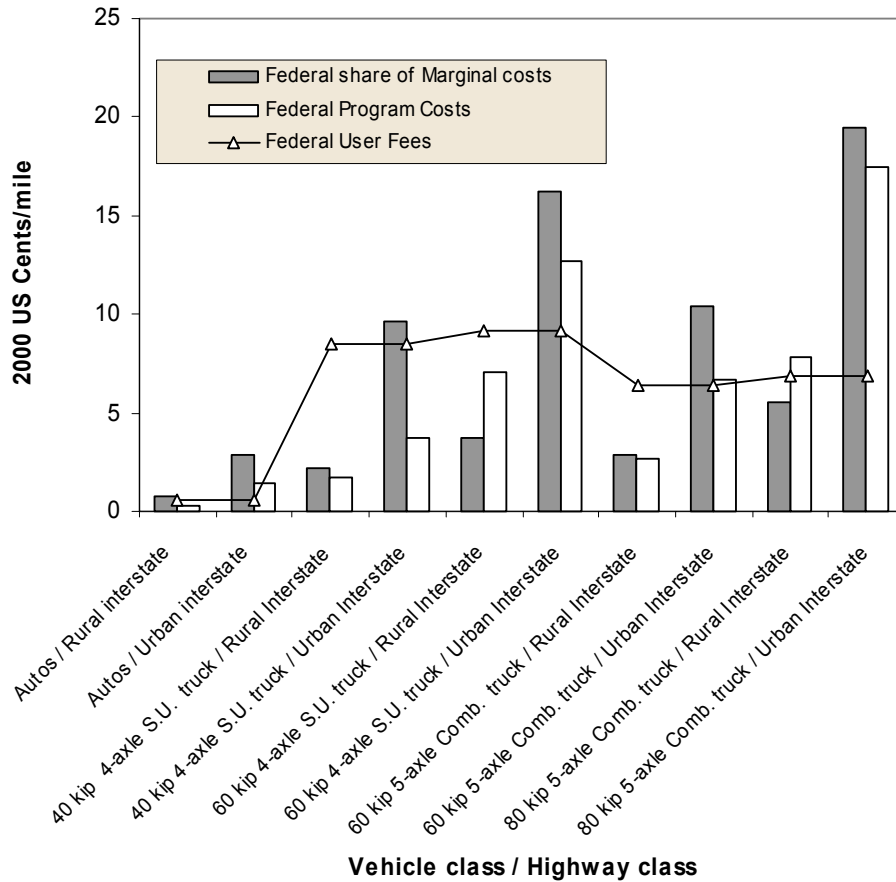
Figure 3-12   Marginal costs, Agency costs and user fees at Federal level

3.2.4   A Canadian Study with a different approach

In (Hajek et al. 1998) a study, based on similar principles as in FHCAS method, is presented. This study was basically conducted with the aim of estimating cost differentials between alternative regulation scenarios with respect to truck weights and dimensions in Ontario, Canada. A significant part of the study was to estimate marginal pavement costs related to heavy vehicle traffic. The major methodological steps are illustrated in Figure 3-14, and comprise three phases

towards the estimation of total lifecycle pavement costs over a 20 year[62] period of analysis. Such lifecycle costs are calculated for different scenarios regarding truck weight and dimension regulatory regimes, and for 20 representative categories of the highway network in Ontario.

Initially, new traffic volumes are estimated for each scenario and road network. The traffic projections were based on individual expected growth rates (expressed in ESAL-kilometres) for 25 different vehicle classes subjected to regulatory changes. These projected traffic volumes were then assigned to the 20 categories of roads based on the size of each network and typical volumes of six-and-more-axle trucks serviced by these roads.

When the estimated traffic levels on each category of roads is calculated, the lifecycle costs of providing a pavement structure to carry the necessary traffic loads are computed, relying on the regular road design and maintenance standards in Ontario. These lifecycle costs are also attributed to different vehicle classes. The methodology is similar to the FHWA method because both methods involve considering a minimum pavement, and then allocate the costs associated with necessary additional pavement thickness to the trucks. The main difference between these methods is the fact that the FHWA method considers *average* ESAL costs, whereas the Ontario approach calculates *marginal* ESAL costs. These marginal lifecycle pavement design costs are a falling function of traffic levels because an incremental increase in pavement thickness permits a significant increase in traffic loads. This means that average ESAL costs will tend to be higher than marginal costs.

Four different ESAL cost functions were developed to distinguish between:

---

[62]   Actually a 60 year period was used for calculating the annuities based on life-cycle costing, but a 20 year perspective was used for the entire analysis.

- New pavements versus in-service pavements
- Southern versus Northern Ontario (to accommodate for differences in climatic impacts)

All costs and maintenance strategies were based on design and maintenance models for asphaltic concrete pavements. The estimated function for "Equivalent Uniform Annual Cost" (EUAC[63]) for new pavements is presented in Equation 3-12. A similar function was estimated for in-service pavements with the same functional form, but with other parameters. The equations gave a fairly good general statistical fit ($R^2 > 0.80$), and all but the N-parameter for in-service pavements appeared to be significant at a 95% test level.

Equation 3-12 $\qquad EUAC_{NP} = 1601 + 311 \cdot \left[\log_{10} ESALs\right]^2 + 1394N + \varepsilon$

Here

ESALs   is the number of equivalent axle loads per lane per year, and

N   is a dummy variable indicating whether the network in question

      is located in Northern (N=1) or Southern (N=0) Ontario

ε is an error term

To obtain an expression for marginal costs per ESAL, the total cost functions were differentiated. In the case of the "new pavement" function the result turns out as illustrated in Equation 3-13.

Equation 3-13 $\qquad\qquad MCOST_{NP} = 622 \cdot \dfrac{\log_{10} ESALs}{\ln 10 \cdot ESALs}$

Here

---

[63]   Costs in Canadian dollars.

88

$MCOST_{NP}$ is the marginal equivalent uniform annual cost per ESAL for one lane of a new pavement structure

This gives a development of marginal costs as a function of ESALs as illustrated in Figure 3-13. Due to the "economies of scale" with respect to increasing pavement thicknesses, this analysis assigns a very high marginal cost per ESAL to roads with low traffic levels (more than 1 dollar per ESAL for new pavements when traffic levels are very low).
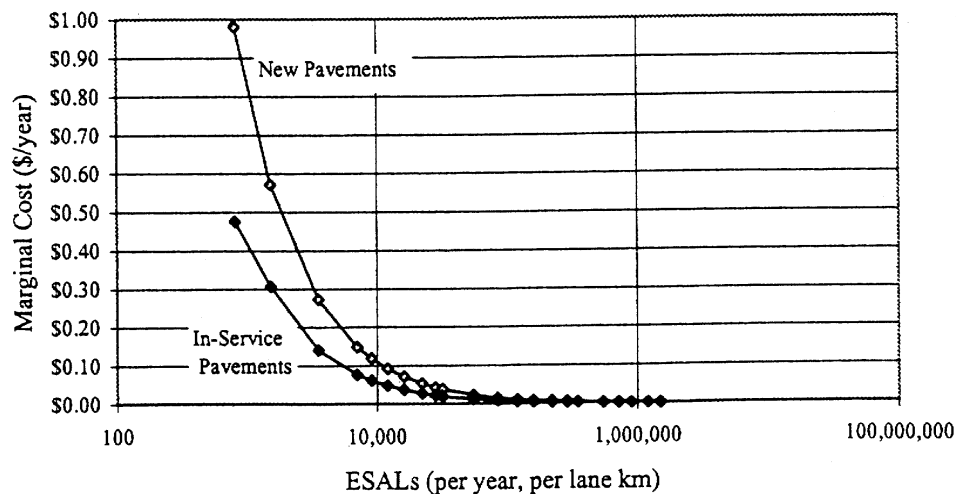
Figure 3-13   Marginal pavement costs as a function of ESALs (Hajek et al. 1998)

Hajek et al. (1998) have also applied these functions to the highway networks defined in the study, and calculated typical marginal costs for an average 5 axle truck. The figures obtained range from the very low estimated marginal cost per kilometre of 0.002 dollars for in-service pavements on the urban freeway network, to an estimated cost of 0.895 dollars on a new pavement on the local road network. Comparing these figures to the ones presented in (Small et al. 1989), the figures from the Ontario study turn out to be considerably lower than the ones calculated by Small *et al*. This may be due to the fact that Small *et al*. made their

calculations based on the actual road designs, which were deemed weaker than optimal, and pavement deterioration due to climatic impacts is higher in Ontario than in the United States.
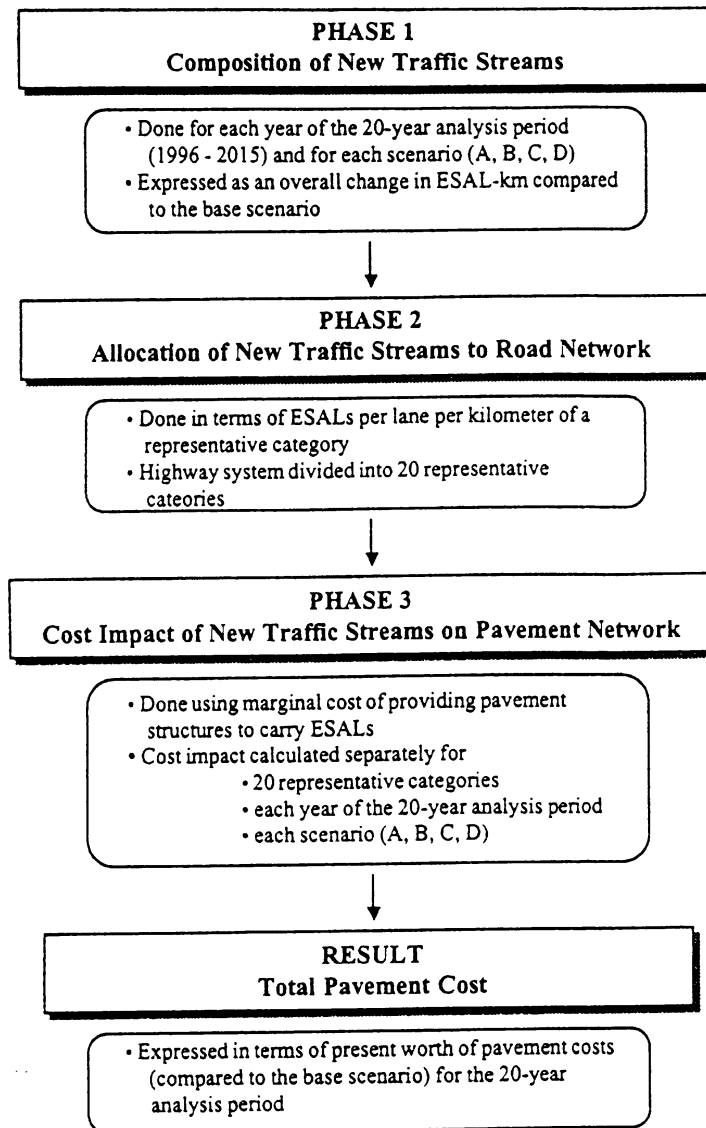
Figure 3-14   Outline of methodology used for allocating pavement damage costs in the Ontario study (Hajek et al. 1998)

Assessment of the Ontario study

This study represents an interesting approach to estimating marginal road wear, and from a pricing point of view, it is certainly an improvement over the *average cost* per ESAL focus of the FHWA. However, as far as I can see, this represents an attempt to unmask the *implicit marginal cost per ESAL* contained in the current maintenance, rehabilitation and design procedures in Ontario, rather than an estimation of actual marginal costs. If these estimates should be regarded as estimates of actual costs, one would have to assume that the current maintenance, rehabilitation and design procedures represent an optimal strategy, also containing consideration of user costs in a full economic perspective. Unfortunately, (Hajek et al. 1998) do not provide any information about the real content of the existing guidelines, i.e. the factors influencing the definitions of the applied maintenance and investment scenarios. It is therefore difficult to assess the validity of the estimates in a "actual road wear"- setting.

3.3   RELEVANT EUROPEAN STUDIES

3.3.1   A review of current cost allocation practices and recent research action on marginal cost estimation

A study by Link et al. (1999), commissioned by EU/DG VII, provides an important background for the EU White paper on transport infrastructure charging (EC-DG7 1998). The project reviews current European models for highway cost allocation and available statistical data related to these issues.

The accounting procedures related to road infrastructure costs differ a lot, as does the frequency of updates and levels of detail. This is to some extent summarised in Figure 3-15.

Criteria	Description	Practice in the following countries
Regular road account	Regular calculation of costs (or expenditures) for different categories, Comparison with revenues, Calculation of cost coverage	Annually: D[64], UK, CH Periodical updates: D[65], A, DK, NL, IRL, E, F, SF, S, N[66] No estimates available: B, L, P, GR
HGV road account	Differentiation of costs (or expenditures) for vehicle categories	D, UK, A, DK, IRL, E, F, NL, S, SF, CH, N
Expenditure accounts	No capitalisation of the road investments	D, UK, NL, IRL, E, F, S, SF, I, CH, N
Cost accounts: Estimation of capital costs	Capitalisation of road investments by using assumptions on life expectancies and interests	D, A, DK, S, SF, F, CH
Cost/expenditure allocation to vehicle categories	Use of a specific method to allocate costs/expenditures to different vehicle categories, especially for HGVs	Own method: D, UK, A, DK, F, NL, S, SF, CH, UK, N Method adopted from another country: IRL, E
Distinction between fixed and variable costs / expenditures		D, DK, S, SF, F, I, N

Figure 3-15   Practice of road infrastructure cost accounts in Europe[67]

From Figure 3-15 we can see that a lot of European countries have developed their own methodology for allocating costs (or expenditures) to vehicle categories. Figure 3-16 provides an overview over these different allocation methods.

One of the aims of this EU-project was to establish a framework for a harmonised method for cost accounting and allocation. It seems that the European states fall into three groups characterised by different degrees of data availability and type of methodology applied (see Figure 3-17).

Even though many countries seem to have the prerequisites for a sophisticated methodology, this does not mean that the *results* from the national models are comparable. Link et al. (1999) also illustrate that even if one applied the same methods to the available data from different countries, one could not expect to obtain comparable figures because of different accounting practices.

---

[64]   Frequently elaborated cost data.

[65]   Annual expenditure data.

[66]   Norway was not part of this study. The N is included based on the author's assessment of the Norwegian methodology described in chapter 3.3.4.

[67]   Source: Link, H., Dodgson, J. S., Maibach, M., and Herry, M. (1999)., supplemented with Ns (for Norway) where appropriate by the author. See footnote 66.

92

Country	Method used
Austria	• Regression analysis • Adaptation of German method
Denmark	Differentiation of capital and running costs into: • Fixed costs • Vehicle-km dependent costs • Space dependent costs • Weight dependent costs Use of specific weight and space factors by type of vehicle
France	Differentiation between fixed and variable expenditures Use of different allocation factors such as: • Vkm • weight-vkm • standard axle-vkm
Germany	Differentiation between marginal costs and capacity costs. Allocation of: 1. Marginal costs by AASHO-Road factor*vkm 2. Capacity costs by (speed-dependent) equivalent factor*vkm
Italy	Differentiation between marginal and capacity expenditures. Use of different allocation factors such as: • Vkm • Axle-weight-km • Standard-axle-km
Netherlands	Differentiation of investment expenditures and running expenditures into different allocation factors such as: • Vkm • PCU-km • Axle load-km
Norway	Differentiation between non-traffic related costs, traffic-related costs and traffic volume related costs. Cost allocation based on: • Vkm • PCU-based vkm • Standard axle-vkm (AASHO, 2,5 power)
Finland	Differentiation between fixed and variable expenditures. Use of different allocation factors such as: • Vkm • Weight-factors
Sweden	Differentiation of fixed and variable expenditures into: • Vkm-dependent expenditures • Space- and speed-dependent expenditures (allocated by PCU-km) • Weight-dependent expenditures (allocated by AASHO-factor-km)
Switzerland	Allocation of: 1. Weight dependent costs of new investment (estimated by percentages per road type) by weight-factors 2. Weight dependent costs for pavement and investive maintenance by axle-load-vkm 3. Capacity costs: 80% by vehicle-length*vkm. 20% by vkm. 4. Current costs by vkm
United Kingdom	Allocation of: 1. Capital expenditure: 15% by max. GVW-km, 85% by PCU-km. 2. Maintenance expenditure: Further differentiated by types of expenditures, different allocation factors applied. 3. Policing and traffic wardens: By vehicle-km

Figure 3-16   Cost allocation methods in Europe[68]

---

[68]   Source: Ibid., supplemented with comments by the author on the Norwegian model. See footnote 66.

Group	Category	Members
1	Countries with detailed data and sophisticated methodology	D, A, DK, S, F, SF, CH and N
2	Countries with a good database (accounts), but no sophisticated methodology	UK, NL, I, IRE, B, E
3	Countries with only sporadic estimates and no own methodology	L, P, GR

Figure 3-17   Categorisation of current level of sophistication with respect to models and data availability for infrastructure costs in European countries

As in the US Federal Highway Cost Allocation Study (US Department of Transportation 1997a), the main focus of this EU-study seems to be on the allocation of *average* costs, rather than marginal costs. However, the EU-study clearly recognises the need for marginal cost estimation in order to achieve a proper foundation of pricing that leads to allocative efficiency. The study concludes that the empirical foundation of estimating proper marginal infrastructure costs is not present in any European country, and emphasises an urgent need for further research in this area. This recognition is carried on in the EU white paper on infrastructure charging (EC-DG7 1998), and therefore a major part of the "phased approach" towards harmonised infrastructure charging in Europe is dedicated to further research in the area of cost estimation. A set of research programmes has been launched within the 4th and the 5th research frameworks to augment the knowledge in this area. Apart from these research programs the so-called High Level Group on Transport Infrastructure Charging has been put together to provide expert advice on these issues to the Commission.

3.3.2   *The EC High Level Group on Transport Infrastructure Charging*

The EC High Level Group on Transport Infrastructure Charging (EC-HLG) comprises a number of experts on infrastructure charging from many European countries. The group has delivered three reports on estimating and charging for transport costs. The first report (EC 1998) formed the background of the general

recommendations given in the EU white paper on infrastructure charging (EC-DG7 1998). Later the EC-HLG was asked to reconvene and define in more detail the way in which transport costs can be defined and estimated (EC 1999a). The third report (EC 1999b) deals with policy options for charging. This is commented upon in section 1.2. Here, I will focus on the two reports dealing with the definition and estimation of costs relevant to transport infrastructure charging.

Basically, the first report recommended that the EC should take the necessary steps to ensure that infrastructure charging principles are the same in all member states. The level of the charges should ideally be based on *marginal* costs, but if cost recovery was politically necessary one should allow for *two-part tariffs* or *Ramsey-pricing* in order to raise enough revenue for infrastructure funding.

In its second report the EC-HLG comes closer to actually recommending a more detailed procedure for estimating infrastructure costs (including external costs related to accidents, congestion and environmental damages). The general framework of this procedure comprises 5 steps to reach a monetary value for each cost component, and an additional 6th step to consider charging mechanisms (see Figure 3-18).

```
┌─────────────────────────────────────────┐
│           Step 1: Define costs           │
└─────────────────────────────────────────┘
                     │
                     ▼
┌─────────────────────────────────────────┐
│        Step 2: Define cost categories    │
└─────────────────────────────────────────┘
                     │
                     ▼
┌─────────────────────────────────────────────────────────┐
│  Step 3: Determine impact and causal links (cost "drivers") │
└─────────────────────────────────────────────────────────┘
                     │
                     ▼
┌─────────────────────────────────────────┐
│           Step 4: Cost attribution       │
└─────────────────────────────────────────┘
                     │
                     ▼
┌─────────────────────────────────────────┐
│      Step 5: Calculate monetary value    │
└─────────────────────────────────────────┘
                     ┊
┌─────────────────────────────────────────┐
│    Step 6: Consider charging mechanisms  │
└─────────────────────────────────────────┘
```
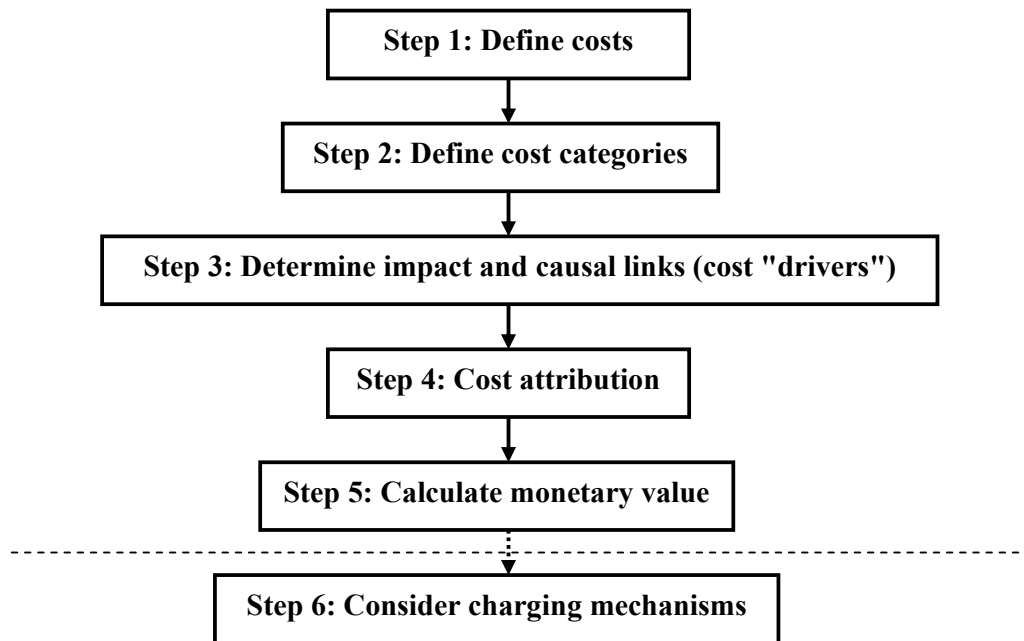
Figure 3-18   The cost estimation and allocation process[69]

Defining costs is necessary because there is a great deal of confusion in this area. Policy papers are not always clear regarding the use of notions like variable and marginal costs, average or *marginal* costs, costs and expenditures etc. All efforts on estimating infrastructure costs should therefore start out with clearly defining the notions used, as was my aim in Chapter 2.1.

Cost categorisation is needed in order to separate marginal costs from total/average costs, and to prepare the ground for cost attribution. An example of such categorisation could be the division of infrastructure costs dependent on road type (e.g. national, regional, local, urban, or four-lane vs. two-lane, concrete vs. flexible pavement etc.)

The third step is to determine the *cost drivers*, i.e. to establish causal links between marginal use and marginal cost. In some instances this is a quite simple

---

[69]   Source: EC, H. (1999a)..

96

task. In most situations, however, the conclusions regarding causal links rely on a mixture of theoretical cause-effect considerations and more or less sophisticated empirical "evidence" for statistical correlation patterns. Clearly identifying the marginal costs stemming from a particular use can be very difficult, especially when dealing with complicated interaction patterns. An example of such an issue related to the topic of this thesis could be the *combined* effects of traffic loadings and climatic factors on road wear.

Having established the proper cost categorisation and the causal links, the fourth step is to *allocate marginal costs to user groups*. Theoretically, marginal costs will vary almost continuously with time of day and location and also with the individual characteristics of the user. For practical purposes (e.g. charging for infrastructure use) a certain level of averaging is necessary of course. The availability of detailed information will also limit the degree of detailing. Sometimes data are primarily available on an aggregated level (e.g. road expenditure accounts) and needs to be disaggregated and allocated to user groups. In other situations we have better knowledge of the impacts of individual use, and need to aggregate this to proper average cost estimates. The different cost allocation studies reviewed in this thesis, represent both approaches to cost attribution.

Trying to reflect external user costs in a Pigouvian tax necessitates a monetary valuation of the identified effects. For some cost items this is fairly straightforward. In the absence of market imperfections, commodities traded in markets could be valued using market prices (adjusting for fiscal taxes). In this manner most costs related to infrastructure investments, maintenance and operation could be determined in monetary terms. The issue is much more complicated when the impacts of congestion, accidents and environmental hazards are to be monetized. In this case willingness-to-pay / willingness-to-accept studies have to be applied in order to provide the alternative values and shadow prices of

the commodities in question.

When the monetary value of all impacts is established, the final step would be to evaluate which charging schemes one should adopt. Here both efficiency and practicability should be considered. This is the issue of the third report from the EC-HLG (EC 1999b), and also chapter 1.2 in this thesis.

The EC-HLG has had three working groups producing reports on the following sub-items:

- WG1: Estimation of infrastructure costs that vary with use
- WG2: Estimation of costs associated with congestion and environmental pollution (2 reports)
- WG3: Estimation of costs associated with transport accidents

These reports provide even more detailed considerations on how to proceed when trying to identify and value the different effects related to transport infrastructure use.

Recently, several research projects have been conducted to augment the knowledge in the field of charging for infrastructure use. Two of the projects from the EU 4th framework research programme are central to the issues focused in this thesis: CAPRI (Concerted Action on Transport Pricing Research Integration) and PARIS (Performance Analysis of Road Infrastructure). Interesting projects have also been launched within the 5th framework programme, like UNITE (Unification of Accounts and marginal costs for Transport Efficiency).

3.3.3   Recent European research programmes

The CAPRI project

The objective of the Concerted Action on Transport Pricing Research Integration (CAPRI) was to bring together the results of research studies that relate to the role of pricing in transport policy development.

In addition to examining research programmes at the national level, the aim of CAPRI was to facilitate the exchange of results from European level research, including:

- Strategic Research, Urban Transport and Road Transport tasks in the Fourth Framework, DGVII programme
- "The Pricing and Financing of Urban Transport" from the APAS study
- The Telematics Applications Programme (DGXIII)
- The Green Paper "Towards Fair and Efficient Pricing in Transport"
- The White Paper "Fair Payment for Infrastructure Use"
- CARD-Me (Concerted Action for Research on Demand Management in Europe)
- The Joint Scientific Committee established for pricing projects in the strategic and road sectors.

The project establishes the general economic principles of pricing transport services (Calthrop and Proost 1998) in much the same manner as I have done in chapter 1.1. Rennings et al. (1999) give more detail on the valuation of transport externalities. This is dealt with in chapter 2.3 in this thesis. The project also focuses on items relevant to urban road pricing, which is not a focal point in my thesis (Vougioukas et al. 1999). The CAPRI project covers both road, rail and air transport pricing. Finally, the last focal point is an assessment of the likely impact of implementing efficient pricing. The end-product of the programme is set of recommendations for (Nash et al. 2001):

- Pricing principles*
- Valuation of Externalities*
- Road Pricing – Urban and Inter-Urban*
- Rail and Other Public Transport
- Air Transport
- Likely Impacts of Implementing Efficient Pricing*

The recommendations given in the areas marked with an asterisk, are reproduced in the Appendix for further reference.

The UNITE project

UNITE is a part of the European Union's Fifth RTD Framework Programme in the thematic programme "Competitive and Sustainable Growth". The program is not finished (February 2003), but a lot of the deliverables are already published. The program sets off where CAPRI ended, by trying to approach more concrete estimates of transport costs. Since the political agenda is concerned with both equity and efficiency, UNITE focuses on transport accounts for establishing the costs-to-tax ratios for user groups, *and* on calculating marginal costs suitable for efficient pricing.

UNITE has thus three core objectives (Sansom et al. 2000):

1. To develop *pilot transport accounts* for all modes, for the EU15 and additional countries;

2. To provide a comprehensive set of *marginal cost estimates* relevant to transport contexts around Europe; and

3. To *deliver a framework for integration of accounts and marginal costs*, consistent with public finance economics and the role of transport charging in the European economy.

100

The project involves the production of *pilot accounts* for 18 countries (EU15, Estonia, Hungary and Switzerland), covering the years 1996, 1998 and 2005 for all significant passenger and freight modes. These pilot accounts are supposed to reflect the best practicable accounts under currently available amount and detail of statistics and models[70]. However, one of the tasks that UNITE has addressed is also identification of key features of *ideal* accounts, to head for better estimates in the future. According to Sansom et al. (2000), the key features of ideal accounts should comprise:

- **A high level of disaggregation** – reflecting factors such as location and time period at the transport link or terminal level;

- **Full information about the financial and social cost structure** – including marginal, variable and fixed costs;

- Similarly, **full information on the charging / taxation structure** – including variable and fixed components

- Use of a basis of **social cost accounting** – as opposed to a purely financial or business accounting basis;

- Dynamic – **examining changes in response to new charging structures** / levels through the use of transport modelling and enabling the non-linearity of cost functions such as congestion to be taken into account by means of demand and supply interactions;

- **Capable of aggregation to the appropriate level of decision-making** – to enable examination of who incurs costs and how much they pay, for different geographic areas, modes, income groups etc.

- There should also be **no arbitrary allocation of costs** to user groups.

---

[70] The ideal is to be able to forecast cost levels under future tax regimes and future traffic levels. Comprehensive transport modelling tools are a necessity for achieving such estimates, and not all the counties possess such model instruments.

Costs that are truly joint should not be fully allocated at the most disaggregate levels.

▪ **Cost information should relate to the present and the future**, and not to the past, thus including costs of new infrastructure and renewals, but excluding sunk costs.

In addition to the pilot accounts the UNITE project also involves studies of marginal costs in more concrete settings through a number of case studies. These include studies for all transport modes, and in the areas of infrastructure costs, supplier operating costs, transport user costs, accident costs and environmental costs. A full list of case studies could be found in Bossche et al. (2001), Table 7.2.

The final step of UNITE, was to give advice on how to integrate the cost recovery/equity approach with the efficiency approach to head for possible acceptable concrete policy options.

Presently (February 2003) some of the case studies are published, and also the pilot accounts for Germany and Switzerland, along with the more general deliverables concerning valuation conventions, details on principles for the pilot accounts and on alternative frameworks for the integration of marginal costs and transport accounts.

The PARIS project
Under the EU's Fourth Framework Research Programme, a comprehensive project for Perfomance Analysis of Road Infrastructure (PARIS), was conducted. This brief presentation is mainly based on EC-DG7 (1999) and Bastiaans (1998b).

The PARIS project collated data from 15 participating European countries, normalised them, and put them into a database of in-service road test sections. These data were gathered in so-called Real Time Loading Testing (RLT). The data were augmented with information obtained from Accelerated Loading Testing (ALT) of pavement sections using dedicated research facilities. A total of

102

700 test sections were entered into the database for estimation purposes, and another 200 for validation of the estimated models. These contained data for both flexible and semi-rigid pavements[71].

The primary goal of the PARIS project was to give road managers the tools needed for cost-effective management of the European road infrastructure. As the models were meant to be applicable for use in the national pavement management systems, the explanatory variables explored were limited to factors that most road authorities already register regularly.

Four types of distress were modelled (see section 4.3 for a further description of distress types):

- Cracking
- Ravelling
- Rutting
- Roughness (longitudinal profile)

Typically two phases in the development of pavement distress can be identified; a *initiation* phase, and a *propagation* phase. This is illustrated in Figure 3-19. Different mechanisms apply to these different phases of distress, hence the PARIS-project (to some extent) developed separate models for distress initiation (cracking only) and propagation (all four distress types).

---

[71]   Both categories included only asphaltic concrete (AC) pavements, but the semi-rigid category contained those sections that contained cement bound layers.
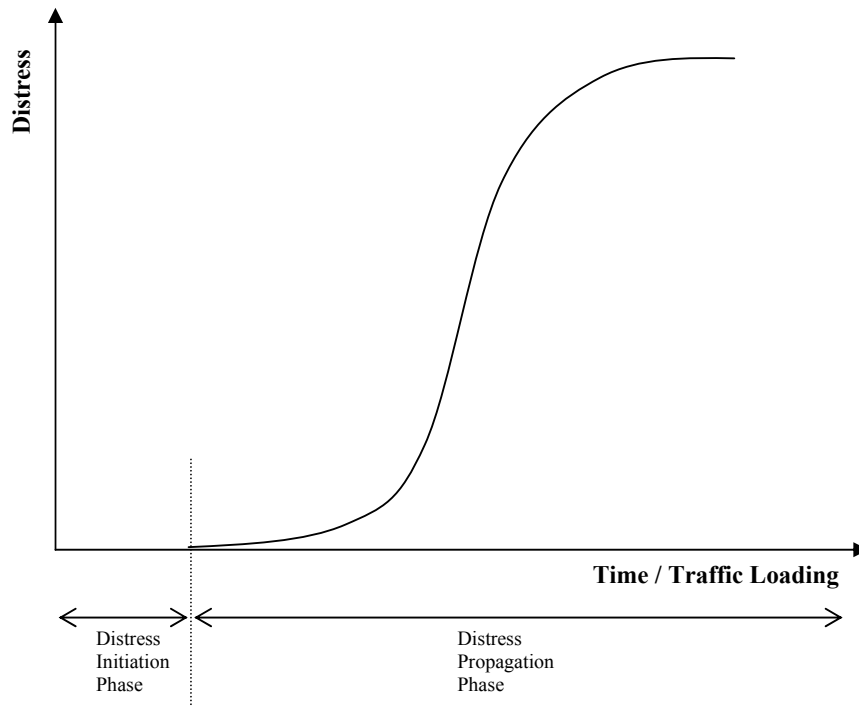
Figure 3-19   General form of development of distress (EC-DG7 1999)

The modelling work comprised the following steps:

1. Identification of response and explanatory variables

2. Examination of variables independently

3. Analysis of paired relationships among variables

4. Model building (initiation, propagation)

5. Illustration of models

Data sources, dependent and independent variables for the different distress types are given in Figure 3-20. As can be seen from the right column, the range of explanatory variables is quite limited (apart from the cracking initiation models). This is partly due to the amount of available information, but also a result of the statistical modelling where relationships including other factors have been tested,

Distress type	Data source		Response (Dependent) variables	Explanatory (Independent) variables
	RLT data	ALT data		
Cracking: *Initiation*				
▪ Wheel path	X		⎫ Cumulative traffic loadings or age	Traffic / Age Pavement response
▪ Total cracking	X		⎭	Traffic / Age Pavement response & Climate
▪ Semi-Rigid Transverse Reflection Cracking	X		Age	
				Asphalt thickness, climate, strength and thickness of rigid layers
Cracking: *Propagation*				
▪ Wheel path	X	X	⎫ Extent & severity of cracking (%)	Traffic / Age
▪ Semi-Rigid Transverse Reflection Cracking	X		⎭	Climate, Traffic / Age
Rutting	X	X	Rut depth (mm)	Traffic / Age
Longitudinal unevenness	X		Roughness (IRI, mm/m)	Traffic / Age
Ravelling	X		Level of ravelling (%)	Traffic / Age

Figure 3-20   Response and explanatory variables used in PARIS modelling (Bastiaans 1998a)

but rejected. Distress initiation models have not been developed for rutting, roughness and ravelling because this was expected to require laboratory testing of

pavement materials to establish the resistance to distress. For cracking, deflection data (which is very often available) could be used for expressing resistance to distress, hence separate models were developed for cracking initiation.

Since the focus in the FAMAROW-model, presented in section 4, is put on rutting and roughness modelling, I will also give a closer description of the PARIS distress propagation models for these distress types here.

PARIS rut propagation models
Initially four functional forms were considered (linear, power, logarithmic and polynomial), but after the initial analysis the linear and the power functional forms found to give the best fit. The independent variables were either the number of ESALs (100kN) or the pavement age since last overlay (or construction). The analysis is based on maximum rut depth for Hungary, and on the rut depths from the outer wheel path for all other countries. Separate models have been estimated for flexible and semi-rigid pavements. In my context it is important to notice that test sections from Sweden and Finland were excluded from the analysis, due to the extensive use of studded tyres in these countries. *This means that the results from the PARIS rut propagation models generally will not be representative of Nordic roads.*

A closer study of the data from the ALTs, showed that a linear relationship gave a good statistical fit over the region of interest. Linear regression analysis was carried out on the time series data for each section independently, to determine the rate of development of rutting with traffic to produce a set of slope values. The resulting slope values were very different. This is not very surprising, when this is only a univariate analysis, and knowing that some of the test-sections had down to 5 observations each. However, this fact is somewhat concealed in (EC-DG7 1999), as one has (for some reason) chosen to present the *Log10 transformation*s of the slope values, and focused on the country-specific *medians* when commenting on the variation. The individual estimates are not presented, but

106

judging from the figures[72] illustrating the Log10-transformed slope estimates (and their medians), the range of slope values were from 0.3 ($=10^{(-0.5)}$) to 125 ($=10^{2.1}$), meaning that one million ESALs could result in 0.3 mm extra rut depth and up to 125 mm! However, the median value for all the flexible pavement sections is reported to be 0.65, corresponding to a marginal wear per MESAL of 4.5 mm. The equivalent figure for the semi-rigid pavements is 3.1.

The exactly same model is also used replacing MESAL with age (years) since last overlay, alternatively since construction. Once again the estimated slope values cover a quite wide range. This time the slope values themselves are presented (graphically); for flexible pavements they span from almost 0 to 4, indicating that the rut depth increases between 0 and 4 mm over a year. The range is almost equal for the semi-rigid pavements. However, the median figures for each country, lie within the range of 0.2 (Hungary) and 0.9 (France), indicating an annual increase in rut-depth of less than 1 mm.

According to (EC-DG7 1999) a range of variables, comprising construction, deflection, climate, geographical location etc. were tested as additional explanatory variables. The test method is not very well documented, but it seems that this may have been done merely by studying the bivariate correlations between the individual variables and the estimated slope values. The report also concludes that the development of other distresses, like cracks, did not appear to have any influence on the evolution of rutting before the pavement is totally cracked (result from the ALT data only).

PARIS Longitudinal unevenness (IRI) propagation models
Most of the PARIS test sections contained longitudinal unevenness data. Mainly these data were in the form of th International Roughness Index, IRI, others were converted into IRI. A total of 645 test sections were used in this analysis, most of

---

[72]   See Figure 9.6 in *op.cit.*

them from flexible pavements. Almost 78% of the test sections were located in Sweden and The Netherlands. Once again a simple linear functional form was chosen. The IRI-data proved to be very different from the rut-depth measurements, as the annual change typically was very low. In 90% of the flexible pavement test sections this annual change was less than 0.1 mm/m per year, for the semi-rigid pavements this percentage was 100% (Bastiaans 1998a). With typical IRI-values ranging from 1 to 6, this means that the majority of test sections had almost negligible changes in the IRI-values. This was also confirmed from the regressions, where *a large number of the slope estimates were not significantly different from zero*.

As for the rut propagation models, a large number of other independent variables are reported to have been tested using cluster analysis, but none came out significant. There is, however, a weak indication that roads with low annual traffic, thin asphalt pavements and narrow carriageways are more susceptible to change in longitudinal unevenness with time, than the more solid high volume roads. The sample of roads contained in the PARIS-project is based on national primary road networks in Europe, which are generally quite high standard networks. This may explain the discrepancy between the findings in the PARIS-project and other studies, where roughness has been found to be a rather important parameter to pavement maintenance procedures.

A conclusion from the longitudinal unevenness models in PARIS, would be that roughness measured by changes in IRI may *not* be a significant parameter for road maintenance models. An alternative conclusion would be that IRI is not sensitive enough, or measured accurately enough to be applicable to high standard roads. The crucial question is, however, whether road user costs are sensitive to such small variations in roughness at all. If they are not, we could probably disregard longitudinal unevenness as a factor affecting optimal road maintenance and optimal pricing policies.

108

Although indicated at the start of the PARIS-reports (see Figure 3-20), no estimation efforts of relating IRI to *traffic loads* are reported, neither in Bastiaans (1998a), nor in EC-DG7 (1999).

3.3.4   Norwegian marginal cost studies

The Institute of Transport Economics (TØI) in Oslo has on several occasions conducted cost allocation studies commissioned by Norwegian authorities. The last effort is documented in Eriksen (2000) and Eriksen et al. (1999), which comprise a review of marginal cost responsibilities of all major modes of transport in Norway (road, rail, sea, and air transport). External costs considered in this report cover:

- Air emissions (SO_2, NO_x, NMVOC, Particles - PM10, Greenhouse gases)
- Noise
- Accidents
- Infrastructure wear
- Congestion

In Figure 3-22 the average figures for all public Norwegian roads are presented. In a marginal cost setting, external costs will vary very much with the traffic conditions and the local environment, so average figures are really not that useful for evaluating the suitability of the current taxation level. Eriksen et. al. *(op. cit.)* has taken the calculations one step further by presenting figures for three levels of density of population. Here I have chosen to present the figures for "Major cities[73]" and "Rural areas". These figures are shown in Figure 3-23 and Figure 3-24, and also graphically illustrated in Figure 3-25 and Figure 3-26.

---

[73]   Note that the term "Major cities" in a Norwegian setting would mean the biggest cities of Norway, i.e. cities in the area of 100 000 to 500 000 inhabitants.

Step 1: Calculate *share* of total maintenance budget for national roads that relates to traffic volume
Based on estimates of shares related to wear from the use of studded tyres and deterioration of roads, bridges and ferry-quays

Step 3: Calculate the average main-tenance cost per ESAL-mile
By dividing the total share of maintenance cost by the total number of ESALs on the national roads

Step 2: Calculate total no. of ESALs on national roads
Based on:
- "2,5 power rule"
- Vehicle class stocks
- Typical axle configurations for each class
- Average annual distance driven in each class
- Average load factors

Step 4: Calculate marginal costs of road wear for each vehicle class
Based on the number of ESALs per vehicle group the marginal road wear costs are calculated from the estimated average maintenance cost per ESAL

Figure 3-21   Conceptual calculation process of the TØI-model

From Figure 3-25 it is quite clear that the dominant elements of marginal external costs in major cities are local emissions and congestion costs (i.e. external time costs). Noise is also a quite significant problem related to road use in such surroundings because many residents live quite close to the heavily congested roads, and their loss of welfare from noise becomes an important part of the disutilities of road use. Other cost items like road wear, global emissions and accidents become relatively small compared to these effects.

The picture is quite different for road transport in rural areas. Note that in Figure 3-26 the scale is very different from the figure representing the urban environment. This enables us to get a better picture of the distribution and magnitude of the "non-urban" external cost items. Road wear becomes dominant only for the heaviest vehicles (buses and HGVs) in these figures. The air emissions are more severe for the petrol-driven vehicles, and, of course, equally

110

dependent on the size of the vehicle.

Vehicle class	NOK per Vehicle Kilometre, All public roads.						
	Greenhouse gases[74]	Local emissions	Noise	Congestion	Accidents	Road wear	Total[75]
Car, petrol	0.02-0.07	0.12	0.04	0.07	0.22	0.00	0.46-0.51
Car, diesel	0.02-0.07	0.12	0.05	0.10	0.22	0.00	0.51-0.56
Bus	0.11-0.37	0.86	0.28	0.11	0.50	0.29	2.15-2.41
Truck 3,5 t+, petrol	0.05-0.18	0.72	0.14	0.06	0.28	0.01	1.25-1.38
Truck 3,5-7,5t, diesel	0.06-0.19	0.37	0.14	0.11	0.30	0.01	0.99-1.12
Truck 7,5t-16t, diesel	0.08-0.26	0.54	0.28	0.11	0.33	0.14	1.47-1.65
Truck 16t-23t, diesel	0.11-0.35	0.72	0.24	0.15	0.35	0.56	2.12-2.37
Truck 23 t+, diesel	0.13-0.44	0.90	0.31	0.18	0.36	1.01	2.89-3.20

Figure 3-22   Marginal external costs of road transport. *All public roads*, Norwegian figures (Source: Eriksen et al. 1999)

Vehicle class	NOK per Vehicle Kilometre, Major cities.						
	Greenhouse gases	Local emissions	Noise	Congestion	Accidents	Road wear	Total
Car, petrol	0.03-0.09	0.17	0.14	0.84	0.22	0.00	1.40-1.46
Car, diesel	0.042-0.12	0.59	0.14	0.84	0.22	0.00	1.82-1.91
Bus	0.16-0.54	4.58	1.39	1.68	0.50	0.29	8.60-8.98
Truck 3,5 t+, petrol	0.07-0.24	1.96	0.70	0.84	0.28	0.01	3.85-4.02
Truck 3,5-7,5t, diesel	0.07-0.25	1.89	0.70	1.68	0.30	0.01	4.66-4.84
Truck 7,5t-16t, diesel	0.10-0.34	2.92	1.39	1.68	0.32	0.14	6.56-6.80
Truck 16t-23t, diesel	0.17-0.58	4.51	1.39	2.52	0.35	0.56	9.50-9.90
Truck 23 t+, diesel	0.17-0.58	4.51	1.39	2.52	0.36	1.01	9.96-10.36

Figure 3-23   Marginal external costs of road transport, *Major cities*, Norwegian figures (Source: Eriksen et al. 1999)

---

[74]   The estimates for external effects from the emission of greenhouse gases are given as an interval, based on the Alternative A and Alternative B in Eriksen e. al (1999), indicating a high and a low alternative for CO_2 emissions. This table is therefore a synthesis of "A Tabell 7" and "B Tabell 7" in the appendices *(op. cit.)*.

[75]   The total is given as an interval representing the low and high alternatives for the effects of greenhouse gas emissions.

Vehicle class	NOK per Vehicle Kilometre, Rural areas						
	Green-house gases[76]	Local emissions	Noise	Conges-tion	Accidents	Road wear	Total[77]
Car, petrol	0.02-0.07	0.11	0.00	0.00	0.22	0.00	0.34-0.39
Car, diesel	0.01-0.05	0.01	0.00	0.00	0.22	0.00	0.24-0.27
Bus	0.10-0.33	0.36	0.00	0.00	0.50	0.29	1.24-1.48
Truck 3,5 t+, petrol	0.05-0.16	0.42	0.00	0.00	0.28	0.01	0.75-0.87
Truck 3,5-7,5t, diesel	0.05-0.18	0.15	0.00	0.00	0.30	0.01	0.52-0.65
Truck 7,5t-16t, diesel	0.07-0.24	0.22	0.00	0.00	0.32	0.14	0.76-0.93
Truck 16t-23t, diesel	0.09-0.30	0.28	0.00	0.00	0.35	0.56	1.27-1.49
Truck 23 t+, diesel	0.12-0.40	0.37	0.00	0.00	0.36	1.01	1.86-2.14

Figure 3-24   Marginal external costs of road transport, *Rural areas*, Norwegian figures (Source: Eriksen et al. 1999)

Eriksen et al. *(op. cit.)* also present figures for the current[78] level of marginal taxation related to road use. Only the taxes levied on fuel are considered truly marginal in this report. V.A.T. is deducted from the figures for HGVs and buses because these costs are deductible for private enterprises. In Figure 3-27 I have calculated "marginal equity ratios" for both the urban and the rural case. If motorists paid according to their marginal costs, they would achieve a ratio of one. If the vehicle class gets a ratio below one, that class pays too little in order to internalise the externalities estimated for that vehicle group. In Figure 3-28 I have presented the estimates of marginal external costs together with the current marginal taxation level. The general picture is that all vehicle classes but the buses seem to pay their way compared to marginal externalities in rural areas. Actually, the lighter vehicle groups seem to pay significantly more than their cost responsibility in this situation.

---

[76]   The estimates for external effects from the emission of greenhouse gases is given as an interval, based on the Alternative A and Alternative B in Eriksen et. al (1999), indicating a high and a low alternative for CO_2 emissions. This table is therefore a synthesis of "A Tabell 23" and "B Tabell 23" in the appendices in op.cit.

[77]   The total is given as an interval representing the low and high alternatives for the effects of the emission of greenhouse gases.

[78]   I.e. 1999 taxes.

112
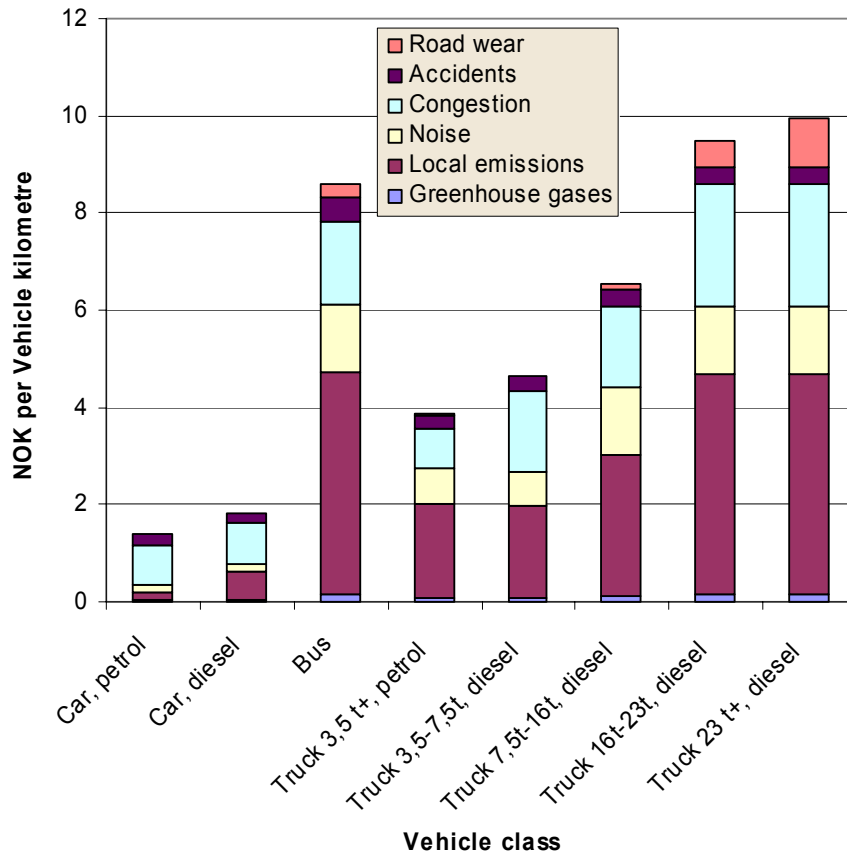
**Figure 3-25   Composition of marginal costs in major cities, Norwegian
figures (NOK 1999)**

The picture is somewhat different when considering the urban case. Due to the
very high costs related to noise, congestion and local emissions, no vehicle group
seems to pay enough to cover their externalities in major cities. It is noteworthy
that in most bigger cities in Norway motorists have to pay a cordon toll when
crossing the city perimeter, and these payments are not included in the figures. On
average, it seems that the lighter vehicles cover their marginal cost responsibility
through the marginal taxes levied on road use. Buses and HGVs pay too little
compared to the average external costs related to the use of public roads in

113

general. This is mainly due to the fact that road wear increases faster with gross vehicle weight than the consumption of diesel. Hence, taxes levied on fuel, is not suitable to cover this part of the externalities.
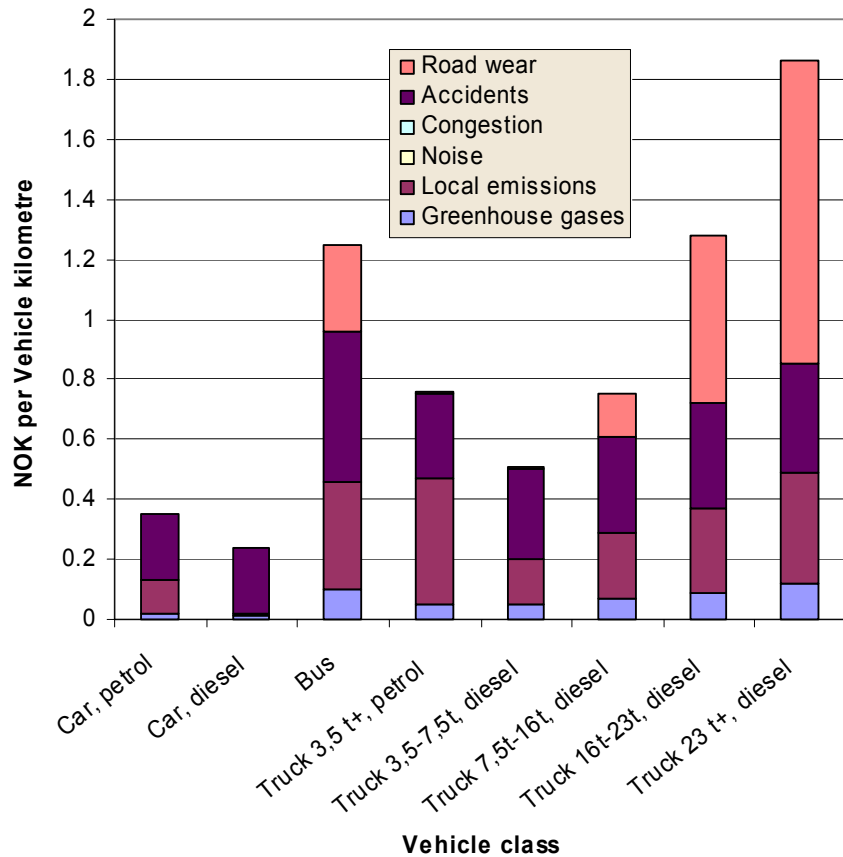
Figure 3-26   Composition of marginal costs in rural areas, Norwegian figures[79] (NOK 1999)

---

[79]   Low alternative for CO_2-valuation

114

Vehicle class	NOK per Vehicle Kilometre				
	Current marginal taxation	Marginal costs[80] Major cities	Marginal costs Rural Areas	Ratio of marginal taxation to marginal costs Major cities	Ratio of marginal taxation to marginal costs Rural areas
Car, petrol	0.53	1.40	0.34	0.38	1.56
Car, diesel	0.37	1.82	0.24	0.20	1.54
Bus	0.18	8.60	1.24	0.02	0.15
Truck 3,5 t+, petrol	1.08	3.85	0.75	0.28	1.44
Truck 3,5-7,5t, diesel	0.79	4.66	0.52	0.17	1.52
Truck 7,5t-16t, diesel	1.07	6.56	0.76	0.16	1.41
Truck 16t-23t, diesel	1.45	9.50	1.27	0.15	1.14
Truck 23 t+, diesel	1.82	9.96	1.86	0.18	0.98

Figure 3-27   Marginal external costs of road transport compared to current[81] marginal taxation. Norwegian figures (Based on Eriksen et al. 1999)

3.3.5   Swedish marginal cost studies

Sweden is one of the very few countries in the world that officially have adopted the principle of marginal cost pricing in the transport sector. The political decision actually dates back to the 1970s. Prior to that the first research programmes aiming at analysing the marginal costs were launched. However, although the *principle* has been adopted, the actual implementation has been limited this far[82]. Since Sweden joined the EU there has been a renewed interest in paving the ground for a more comprehensive implementation of marginal cost pricing[83]. The last research programmes have been conducted by The Swedish Institute for Transport and Communications Analysis (SIKA). The principal considerations on the general conditions for implementing marginal cost pricing in all transport sectors could be found in Hesselborn (2000). Following this report is a more

---

[80]   Based on the lowest alternative of CO_2-valuation.

[81]   I.e. 1999 taxation level.

[82]   However, marginal cost based charges were implemented for the railway sector as early as in 1988.

[83]   Partly as a follow-up of the white paper on fair and efficient pricing.

concrete empirical analysis of the marginal costs in Sweden. This project is not finished (December 2002), but the major results are reported in Hesselborn (2001). Both reports deal with all the transport modes, but I will only focus on the items relevant for pricing roads in the subsequent treatment of the Swedish research work.
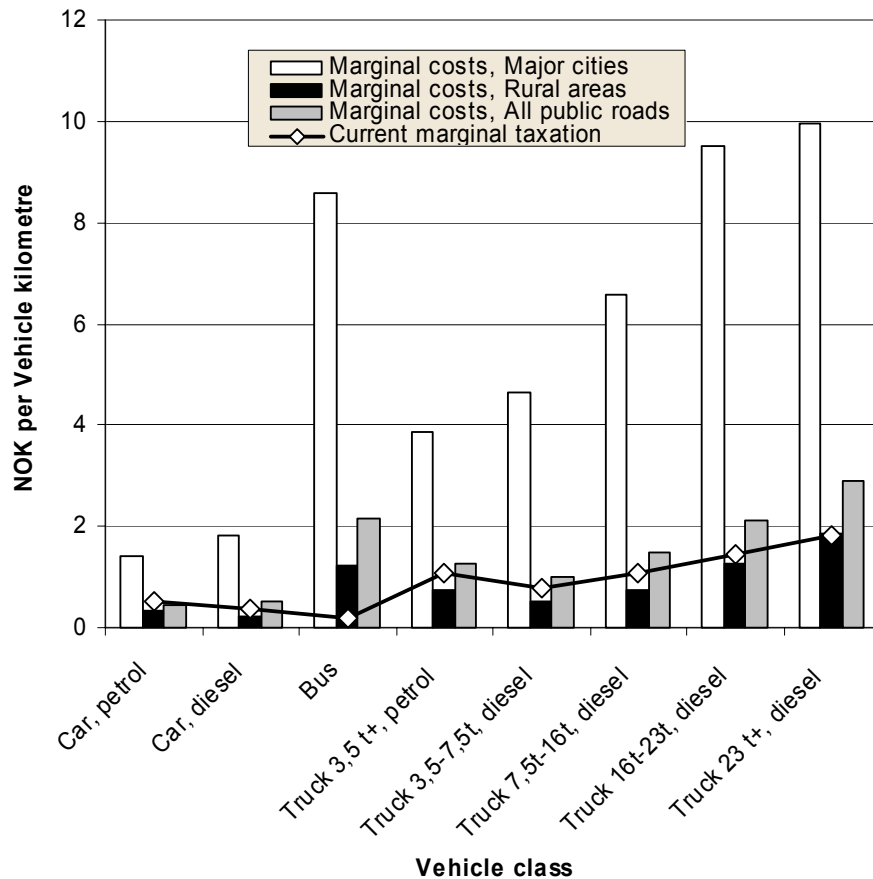
Figure 3-28   Comparison of marginal costs and current taxation level, Norwegian figures[84] (NOK 1999)

---

[84]   Low alternative for CO_2-valuation.

General considerations on the feasibility of marginal cost pricing

SIKA blames the lack of appropriate charging instruments for the limited "success" of marginal cost pricing in the road sector. Earlier studies have often been confined to suggesting appropriate levels for the existing charges levied on fuel. Added to this, the fiscal importance of the charges has limited the political interest in adjusting the charges along with marginal cost pricing principles. Only the effects of CO_2-emissions are really appropriate for pricing fuel. A single focus on fuel charges is therefore not suitable for the purpose, according to SIKA (Hesselborn 2000).

Instead, SIKA promotes a combination of differentiated registration taxes and fuel charges in the short run, and a more sophisticated time, place, and vehicle dependent kilometre charge in the longer run. However, one might also consider congestion charges in the major cities (Stockholm, Gothenburg and Malmö) in the more immediate future.

Swedish estimations of marginal *road wear* have traditionally followed a top-down approach, based on accounted road maintenance costs allocated to different vehicles by ESALs principle. Through a combination of comprehensive engineering research programmes and economic theory, a new (more bottom-up) approach to estimating these costs has been developed. This is based on (Lindberg 2002b) which is treated in Section 2.5 "Elaborations on Newbery's model". The Swedish Public Roads Administration (Vägvärket) has doubts about the outcome of this new approach for estimating marginal costs per ESAL, because the result is increasing marginal costs with increasing design capacity, which contradicts earlier findings. Vägvärket suggests that this, rather contra-intuitive result, may be due to the fact that Lindberg's estimations are based on a sample of rather high volume roads located in southern Sweden. The results may not be representative of the whole road network. SIKA is aware of the problems related to the generalisation of Lindberg's findings, and points out that further research is needed in this field.

Earlier studies have mainly presupposed that *congestion costs* only arise in the major cities. However, recent research efforts suggest that this cost element could not be ignored on certain inter-city road links.

	Road wear	Emissions (ex. CO_2)	Noise[85]	Accidents	Total
Rural areas					
Car, petrol with cc[86]	0.01	0.02	0.01	0.11	0.15
Car, petrol, without cc	0.01	0.25	0.01	0.11	0.38
Car, diesel, with cc	0.01	0.02	0.01	0.11	0.15
Car, diesel, without cc	0.01	0.05	0.01	0.11	0.18
Truck 3,5-16 ton	0.03-0.07[87]	0.32	0.05	0.28	0.68-0.72
Truck >16 ton	0.06-0.17	0.66	0.12-0.26[88]	0.28	1.12-1.37
Densely populated areas					
Car, petrol with cc	0.01	0.10	0.07	0.20	0.38
Car, petrol, without cc	0.01	0.60	0.07	0.20	0.88
Car, diesel, with cc	0.01	0.19	0.07	0.20	0.47
Car, diesel, without cc	0.01	1.05	0.07	0.20	1.33
Truck 3,5-16 ton	0.03-0.07	0.96	0.48	0.49	1.96-1.99
Truck >16 ton	0.06-0.17	1.50	1.10-2.41	0.49	3.15-4.57

Figure 3-29   Calculated marginal cost components for road use. SEK per vehicle km. (Compiled from tables in Hesselborn 2001).

New estimation approaches to determining the risk elasticities for different vehicle types and traffic situations may change the current comprehension of the magnitude of external *accident costs*. The new figures may only constitute one tenth of the previously used ones. However, SIKA (Hesselborn 2001) suggests

---

[85]   The figures for the most densely populated areas are given for the noise cost assessment.

[86]   cc = catalytic converter

[87]   The lower figure represents vehicles without trailers, the upper figure vehicles with trailer.

[88]   The lower figure represents trucks at high speed, the upper figure trucks driving at low speed.

that the established values should still be applied, awaiting further research on the issue.

Much research has been put into the determination of *air emission factors* under different traffic conditions at different locations. It is quite clear that the factors vary considerably depending on population density, traffic density and road category. Generally the damage imposed by air emissions is much higher in cities. The dominating factor is the particle emissions, making the external costs of using diesel cars (without particle traps) much higher than the petrol driven ones. In uncongested rural areas, the CO_2-emissions are the most severe ones (apart from the other emissions from vehicles without catalytic converters). However, the emission of greenhouse gases should be put under a general charging regime, not only covering the transport sector.

Noise may be an important external effect in densely populated areas, especially connected to heavy vehicles. However the effects are very different from location to location, making this effect less suitable for general charging regimes.

Key results from the recent Swedish studies

Based on the considerations given above, the recent estimates on marginal cost components for the road sector are given in Figure 3-29. The emissions figures are given exclusive of CO_2 since the applied shadow price for greenhouse gases would otherwise totally dominate the result. In Figure 3-30[89] two alternative CO_2 shadow prices are applied and added to the other marginal cost components to illustrate the importance of different valuations.

---

[89]   Note that in this figure the marginal costs are converted into SEK per liter fuel. This is merely done by multiplying the vehicle-kilometre figures by the average fuel consumption (cf. Table 4.50 in op.cit.)

	Total marg.cost. Excl. CO_2	Total marg.cost. CO_2 0.50 SEK/kg	Total marg.cost. CO_2 1.50 SEK/kg
Rural areas			
Car, petrol with cc	1.98	3.18	5.59
Car, petrol, without cc	4.69	5.87	8.23
Car, diesel, with cc	2.55	3.84	6.42
Car, diesel, without cc	2.48	3.74	6.25
Truck 3,5-16 ton	3.15-3.34	4.42-4.61	6.97-7.15
Truck >16 ton	2.61-3.20	3.88-4.46	6.41-7.00
Densely populated areas			
Car, petrol with cc	3.27	4.44	6.77
Car, petrol, without cc	7.04	8.20	10.52
Car, diesel, with cc	5.73	7	9.56
Car, diesel, without cc	12.91	14.17	16.69
Truck 3,5-16 ton	9.72-9.92	10.99-11.18	13.51-13.71
Truck >16 ton	6.33-9.18	7.60-10.45	10.13-12.98

Figure 3-30   Calculated marginal costs for road use with alternative valuations of CO_2-emissions. SEK per liter fuel. (Source: Hesselborn 2001).

The new Swedish marginal cost figures do not alter the picture from previous studies much with respect to cars. However, some differences do not show in these figures regarding the use of cars in major cities. Here the emission cost component is significantly higher than before. For trucks the lower limit of the interval has not changed much, but the upper limit has been somewhat reduced compared to earlier studies. This goes both for the rural and the densely populated areas. This effect is mainly due to lower estimates for road wear for truck/trailer combinations.

3.3.6   Comparison of national and international cost estimations

I have now briefly reviewed some central studies containing marginal cost estimations from some American and European countries. It is obvious that the methods used are not quite similar, and certainly the empirical results differ. There is generally no reason to expect marginal costs to be equal in different regions as the impacts of the externalities will depend heavily on factors such as

120

traffic density, vehicle and road characteristics, population density, driving behaviour etc.
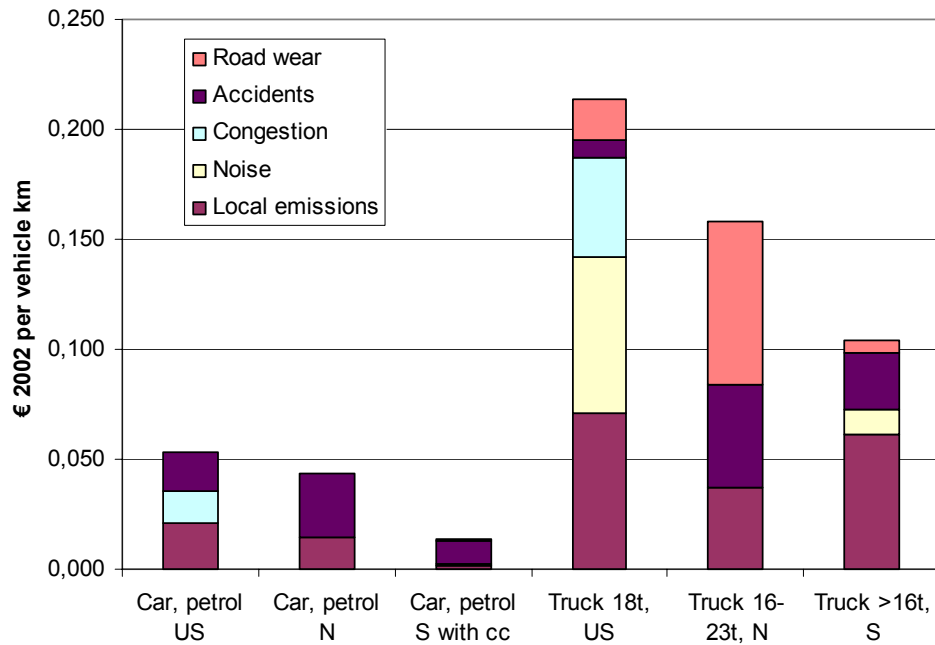
Figure 3-31   Comparison of marginal cost estimates. *Rural areas*. 2002 € per vehicle kilometre

Still, it *is* quite interesting to compare the estimation results to see how estimations of different cost components vary from country to country. I have therefore processed some of the results of the US Federal Cost Allocation Study, the last published Norwegian estimations and the recent Swedish Marginal Cost Study (all presented above). All economic figures have been converted into Euros at a January 2002 price level[90], and evaluated per vehicle kilometre. The figures representing the rural or sparsely populated areas are presented in Figure 3-31. The three first bars represent the external marginal costs for cars, the latter three a

---

[90]   No correction for purchasing power is made, pure currency exchange rates have been applied, and the Norwegian consumer price index has been used for inflating figures to 2002 price level.

mid-to heavy size truck without trailer. Before commenting on the concrete differences, we should note that:

- None of the Swedish figures include any costs related to congestion. This is not due to absence of traffic congestion in Sweden, but because such estimates are missing.

- The Swedish study distinguishes between cars with and without a catalytic converter. Since the majority of Swedish cars have such a device installed I have chosen to include the figures for cars *with* catalytic converters. This has a very significant impact on the local air emissions.

- The truck categories are not strictly comparable. The US figures represent a 40 kip (i.e. 18 ton) 4-axle single unit. The Norwegian figures represent trucks in the range of 16 to 23 tons, but a typical truck in this range would only have 3 axles. The Swedish figures represent all single unit trucks above 16 tons.

- All figures are *exclusive* of greenhouse gas emissions (e.g. CO_2).

For *cars* on *rural roads*, the US figures are higher than the corresponding Norwegian figures, mainly due to the congestion element in the US figures. Norwegian rural roads are normally not congested. Marginal accident costs are evaluated at a lower cost in the US than in Norway, while the opposite is true for the local air emissions. The Swedish figures are significantly lower than both the US and the Norwegian ones. This is mainly due to two factors: Firstly, Swedish figures only represent cars with catalytic converters, which means that local air emissions are almost eliminated. Secondly, the Swedish figures do not cover congestion costs. However, for the rural case, these costs may be assumed to be close to zero anyway.

The overall relative picture is the same for the *trucks*. Leaving out the congestion element, the overall cost estimate of the US study and the Norwegian one coincide

fairly well. However, the composition of the marginal cost element is significantly different. In the US case marginal noise costs constitute more than one third of the costs, whereas this is a zero element in the Norwegian figures. Marginal road wear and accident costs are much higher in the Norwegian study compared to both Sweden and the US. Marginal accident costs are also evaluated to be lower in Sweden than in both Norway and the US. Local emission costs, however, are considered to be more considerable in the US and in Sweden than in Norway.

Moving on to the cost estimates for the *densely populated areas* (cities) given in Figure 3-32, we should first notice that the *scale* is very different to the one used for the rural case. In this setting, congestion costs become far more dominating. This means that the Swedish cost estimates, which are exclusive of this cost element, are somewhat less interesting here. For the cars there is a remarkable similarity between the Norwegian and the US figures, both with respect to composition and overall level. The dominant factor is the congestion costs.

There is a just as striking *difference* in both overall cost level, and the composition of the cost elements for the *trucks* as there is similarity in the car figures. The Norwegian figures (€ 1.2 per vkm) amount to twice the US marginal cost level. The estimated congestion costs are high in both countries, but the main difference is the much higher costs connected to noise and local emissions in the Norwegian study. The external noise costs are almost at the Norwegian level in the Swedish figures, but local emission costs are much lower in the Swedish study as well.

It is beyond the scope of this study to investigate the causes of these observed differences. They may be due to actual differences in the cost levels in the three countries, but they may just as well be due to different approaches to measuring the external effects, and assigning monetary values to them. The US and the Norwegian studies also rely on slightly older statistics than the Swedish one. This will have some impact on the figures, especially with respect to emissions and

123

noise levels, as there has been a significant technical development in this area.
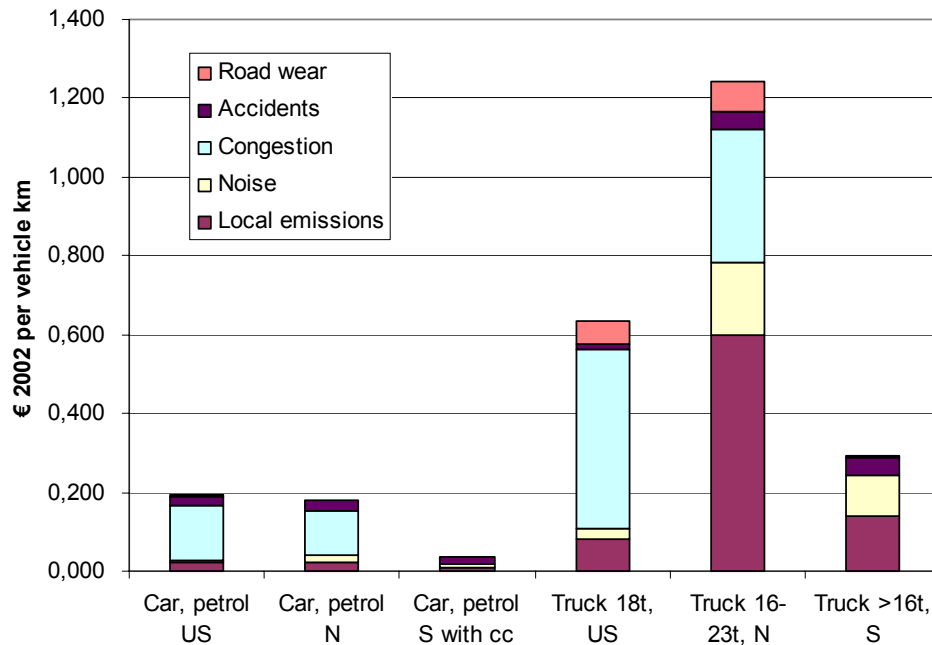
Figure 3-32   Comparison of marginal cost estimates. *Cities/densely populated areas.* **2002 € per vehicle kilometre**

3.3.7   Summary of the findings in the review of marginal cost estimation studies

This concludes my review of recent research in the area of marginal cost estimation. I have covered studies ranging from theoretical developments like Newbery's fundamental theorem of road user charges, via studies more related to political implementation issues, to the more concrete estimation efforts made in this area. I have tried to focus on studies with high relevance for the focus areas of this thesis, but this is by no means an exhaustive review of such research. My hope is that I have managed to illustrate the most important theoretical developments, and the typical nature of research done in this field.

My review starts with an overview of various approaches to studying marginal

infrastructure costs, introducing a typology for classifying the various studies. Three "dimensions" are identified as the main classifiers: The choice of functional form, the direction of approach (top-down, bottom-up), and the type of information used (network accounts/statistics or experimental data).

The first study presented is by far the most cited reference when dealing with marginal infrastructure cost estimation: The AASHO Road Test. This 40 year old accelerated load test conducted in the USA introduced a number of central concepts and notions that have been extensively used by subsequent studies in this field. The most important may be the notion of an "equivalent standard axle load" (ESAL), related to the famous "fourth power law" which indicates that the doubling of an axle weight would increase the "damaging power" with a factor of 16 ($=2^4$).

The second study presented in this review (Small *et. al.*) elaborates on the AASHO Road Test data, utilising modern econometric tools. This modifies the outcome of the AASHO Road Test somewhat, but the "fourth power rule" survives for flexible pavements (a third power rule is suggested for rigid pavements). The re-estimation of the AASHO Road Test data is, however, merely one part of a comprehensive framework established for putting forward a whole new policy proposition for highway maintenance and investments in the USA. The theoretical framework is much in line with the one introduced in Newbery's Fundamental Theorem of Road User Charges presented in Chapter 2. Based on derived optimal durability models, the majority of the US highway network was deemed to be weaker than optimal, thus causing estimated marginal road wear to be higher than it would have been under an optimal construction and maintenance regime.

The US Federal Highway Cost Allocation Study is a very comprehensive study mainly aimed at allocating highway costs to vehicle groups on an equity basis

(cost occasioned approach). However, the last study, completed with an environmental cost appendix in year 2000, also paid some attention to the estimation of *marginal costs*. The methodology for the estimation of these marginal costs are sparsely documented in the publications I have managed to bring about, however the outcome of the study is presented for comparison with European studies.

The final American study is a Canadian one, estimating the marginal road infrastructure costs of the Ontario road network. The main difference between the approach of this study compared to the US studies presented above, is the fact that this one estimates the *implicit marginal cost per ESAL* contained in the current maintenance, rehabilitation and design procedures in Ontario, rather than the *actual* costs on the in-service roads. The fact that Canadian marginal cost estimates turn out to be significantly lower than the ones estimated by the US FHCAS may be due to before mentioned below optimal durability of US highways.

Turning to the European studies, the focus has been partly put on the most recent EU research programmes in this field, and partly on specific Scandinavian studies. For the sake of completeness, I have started this section by providing an overview over current methodologies applied in various European countries. This overview shows that many efforts are made in the field of road cost allocation in Europe, but that the majority of studies have focused on providing equity based figures, and whenever the focus has been put on marginal costs, then the methodology has been rather simplistic in most cases.

The CAPRI and the UNITE project represent the main contributions to the field of marginal external costs of road use in Europe lately. Where the CAPRI project mainly focused on establishing the principles and policies, the UNITE project have had a more concrete approach, trying to establish a unified accounting

126

approach for providing equity based figures, establishing a recommended methodology for estimating marginal costs, and providing a range of case studies containing actual marginal cost estimates.

The PARIS project does not focus on the estimation of the marginal costs of road use as such, but on the development of models for pavement performance. I have included a short presentation of some elements of this project because the main empirical contribution of this thesis is in the estimation of marginal road wear. However, since the Nordic roads unfortunately were left out of the database for the rut propagation models, I find the models developed to be of little relevance to Norwegian conditions (mainly due to the extensive use of studded tyres in Norway).

Previous Norwegian studies of marginal road user costs have been conducted by the Institute of Transport Economics (TØI) in Oslo. The Base scenario of the CATERU model presented in Chapter 5 represents an update of the last TØI-model, and therefore I have given a thorough documentation of the building blocks of this model. Some rather extensive marginal cost estimation projects have been carried out in Sweden lately, and since the climatic, demographic and economic conditions in these two Nordic countries are comparable, these studies should be quite interesting seen from a Norwegian point of view. Perhaps the most interesting part of these studies, seen from my point of view, are the models for road wear developed by Gunnar Lindberg of The Swedish National Road and Transport Research Institute (VTI). These models represent to some extent the same bottom-up approach used in the development of the FAMAROW model presented in Chapter 4 of this thesis. However, it is also interesting to take notice of the fact that the Swedish Public Roads Administration has doubts about the degree of generalization that could be made based on these new models.

It is quite clear that the different approaches used for the estimation of marginal

road user costs in the various studies presented in this chapter, makes it hard to compare the results. Not only do they differ in methodology, but they also differ with respect to scope (cost components covered) and by vehicle categories. Despite this, and despite the fact that there is no reason to expect marginal costs to be the same from place to place, it is interesting to make some comparisons – if not for any other reason – merely to illustrate how different current estimates are. The really interesting exercise from a research point of view, would of course be to further explore the causes behind the differences. Are they mainly due to different methodological approaches, measurement errors, - or do they represent real differences in the cost structures of the countries involved? It is beyond the scope of this thesis to dig deeper into these questions, and I may only notice that current estimates differ a lot, both with respect to overall magnitude, *and* with respect to the relative contribution of the different cost components.

Next, I will turn to a first attempt to estimate a model for marginal road wear on Norwegian in-service roads.

4 THE FAMAROW-MODEL: ESTIMATING ROAD WEAR USING THE NORWEGIAN ATC-UNITS AS WIM-SENSORS[91]

4.1 *THE BASIC IDEA: REGRESSING FACTUAL ROAD WEAR AGAINST FACTUAL ROAD TRAFFIC*

As seen in the previous sections of this thesis, the typical approach to estimation and allocation of marginal road wear costs is founded on theoretical (mechanistic) or experimental analyses (road tests).

The theoretical models are complex, and their ability to predict "real world" effects relies heavily on detailed information on road constructions (layer thicknesses, material characteristics etc.). Even with this information available there are mechanisms working that are difficult to model (e.g. dynamic vehicle-road interactions, the effects of freeze-thaw cycles etc.).

On the other hand, the experimentally based analyses may very well have limited value in applications where the environment is quite different to the test environment. Notably, the bulk of marginal road wear cost allocations carried out is based on the 40 year old AASHO road test – and some variant of "the 4th power law". The transferability of such an analysis may be questionable. This is why the basic idea behind the study reported here is to explore the possibilities of loosening the bonds to these much used approaches, and find a study design more directly related to factually observed road wear, and factual traffic patterns. New methods of weighing vehicles in motion (WIM) enable continuous monitoring of traffic mix with respect to the number and the weight distribution of the axles passing an observation site. Furthermore, there has also been a huge development with respect to more frequent loggings of different statistics describing the state

---

[91] FAMAROW = FActual MArginal ROad Wear, ATC=Automatic Traffic Control, WIM=Weigh-In-Motion

(or serviceability) of the road.

The NPRA[92] has carried out such measurements for some years now, and much of this information is stored in the Norwegian Road Data Bank. This collection of time series data can be used for analysing changes in the state of the network of national roads.

Combined with this database of road condition data, Norway also has deployed a lot of instruments for automatic traffic control (ATC), mainly focused on detecting and fining motorists violating the speed limits. These installations comprise a dual coax-cable and a computer unit that with minor alterations could also be used for detecting axle-loads. A number of these installations have been fitted with extra software and hardware, and axle-loads have been logged for 8-10 years at these observation sites. Recent research has shown that a dual coax-cable installation is far from ideal, especially for capturing dynamic load variations (see Dolcemascolo and Jacob (1998), Argoul et al. (1998), US Department of Transportation (1997b), and Coste (1998). Still trials indicate a fairly good performance tested against static loads, and for statistical purposes the accuracy should be sufficient. There certainly exist more advanced (and costly) WIM-installations, but the advantage of the Norwegian ATC-systems is that there are so many of them out on the real road network already, operating under the real traffic conditions.

This combination of more or less systematic monitoring of the state of the road network, and the number of WIM observation sites on the same network, should give a possibility for a new approach to estimating factual marginal road wear for different axle- (and vehicle) groups.

---

[92]   NPRA=The Norwegian National Roads Administration.

Step 1: Logging traffic mix and traffic levels on selected road sections using WIM-devices
Use the Norwegian ATC-deployments as simple WIM-devices all axle-passages on selected locations on the national road network are logged over discrete time periods.
The average axle-load distribution obtained from the WIM-loggings is combined with general traffic development figures, i.e. AADT estimates from the National Road Data Bank (NRDB), to obtain "continuous" time series of traffic load.

Step 2: Logging factual road wear over the observation period
Registrations of roughness (IRI) and rutting (rut-depth), collected by vans fitted with laser and ultrasound beams, are entered into NRDB. Based on this, a time series of (approximately annual) factual road wear for the relevant road sections is available.

Step 3: Regression analysis: Road wear against traffic data
Control for the impact of climatic factors, relate the recorded traffic data to the observed road wear by regression analysis.

Step 4: Converting marginal road wear into marginal costs
Based on a model for the costs related to maintaining the standard (or performance) of the road network, the costs related to traffic-volume-dependent road wear are estimated.

Step 5: Calculating marginal road wear costs per vehicle-kilometre
Based on the factors resulting from the regression performed under Step 3, and the unit costs established under Step 4, the marginal road wear costs per vehicle-km are derived.

Figure 4-1   The main building blocks of the FAMAROW-model

In Figure 4-1 I have illustrated the conceptual model for estimation of factual marginal road wear and the calculation of marginal costs per vehicle (or axle) kilometre.

Figure 4-2 gives an overview of the various sources of information utilized in the FAMAROW model.

4.2 *WEIGHING VEHICLES IN MOTION*

A significant part of the data for this analysis is based on the Weigh-In-Motion technology. In the Appendix a review of the state-of-the-art of such WIM-

technologies is presented. The major purpose of this review is to provide the necessary background for assessing the functionality of the Norwegian ATC[93] deployments as WIM installations. These are automatic speed surveillance systems comprising a double piezo-electric sensor (nude parallel cables with 3m longitudinal spacing), an automatic camera, and an instrumentation unit. These systems are situated on major roads that historically have a combination of above limit average speed and a high frequency of accidents. Vehicles that exceed the speed limit (by a certain margin) are photographed, and the driver is subsequently fined via an identification of the number plate and the personal information in the vehicle register.

Related to the BUAB[94]-project, conducted by the Norwegian Public Road Administration (NPRA), the possibility of using these ATC-deployments as WIM measurement devices was tested in the early 1990s. Software and autocalibration procedures were developed to turn these deployments into a simple form of a WIM system. The benefit of the autocalibration procedures[95] is that they limit the need to control for environmental factors like temperature and moisture.

The test runs made are unfortunately not very well documented, but the achieved accuracy of the systems is reported in Senstad (1994) and reproduced in Figure 4-3. For the gross weight of the vehicle one found that 84% of the heavy vehicles were registered with an accuracy of +/- 20 percent relative to controlled static weight. The accuracy was even higher for the front axle.

---

[93]   ATC (Automatisk Trafikk Kontroll)=Automatic Traffic Control

[94]   BUAB (Bedre Utnyttelse Av Vegers Bæreevne)=Better utilisation of the bearing capacity of roads.

[95]   Autocalibration means that the system is calibrated through a number of passages of ordinary cars. Knowing the approximate wheel-load of these cars, one can calibrate the recordings according to the signal given from these passages.

132

Data sources and data preparations for
the FAMAROW model

Basic data sources	Data preparations
Information collected for all observation sites	Processing of the basic data collected

Norwegian Road Data Bank
Road Networks 1985-2000

- IRI-loggings
- Rut-depth measurements
- AADT-estimates
- Maintenance actions
- (Road structure information)

IRI-values
Weighted average IRI-values for each 200 m stretch around the observation sites are calculated based on 20m interval loggings. Change in IRI-value since last IRI-observation or re-surfacing calculated.

ATC / WIM-deployments
Discrete time-series 1993-2000

- Number of axles
- Axle spacing
- Dynamic axle loads
- Date & Time
- Vehicle length

Rut-depth values
Weighted average Rut-depth (RD) - values for each 200 m stretch around the observation sites are calculated based on 20m interval loggings. Change in RD-value since last RD-observation or re-surfacing calculated.

Traffic figures

A 2nd power regression equation for Average Annual Daily Traffic (AADT) levels is estimated for each observation site, based on the occasional estimates given in the Road Data Bank. AADT figures are predicted using the estimated AADT-model for calculating differential traffic over the respective observation periods. The general AADT-figures are subsequently combined with the relevant percentages of each axle load category from the WIM-data, to obtain differential traffic loads for each weight class.

Climatic data (DNMI)
Monthly averages 1990-2000

- Precipitation
- Average temperatures
- Snow coverage/depth

Norwegian Public Roads Administration
Road Maintenance Handbooks & Procedures

- Maintenance-triggering IRI values
- Maintenance-triggering Rut-depth values
- Maintenance actions required
- Estimated costs of maintenance actions

Road wear costs
Based on the estimation outcome of traffic-related road wear, marginal road wear costs are calculated for the relevant vehicle classes (relevant for externality charges).

Figure 4-2   Data sources and preparations in the FAMAROW model

Accuracy	Gross weight	Front axle
+/- 10 %	53 %	61 %
+/- 15 %	80 %	69 %
+/- 20 %	84 %	87 %

Figure 4-3   Achieved accuracy at ATC-installation on E6 Øyer in 1993 (NRRL 1994)

These trials indicate that the current ATC deployments in Norway do have a potential as crude WIM measurement devices. However, there is still much research to be done before the expected performance of these systems can be established. Bearing in mind that these tests were conducted under rather ideal (summer) conditions, the expected average accuracy of a typical ATC system would probably be lower than the figures reported above.

In the Appendix, the design of WIM-systems with the necessary accuracy for different purposes is presented. In order to control for dynamic loads (and thereby isolating static loads) much more comprehensive systems based on multiple sensors are required. As could be expected, there is a trade-off between cost and more advanced systems with higher measurement accuracy.

4.3   MEASURING ROAD WEAR

4.3.1   Types of road wear

Road wear could be defined from at least two different angles: One with the major focus on the *physical* state of the road, and one with a more *functional* origin. Focusing on the changes of the physical state of the road, one would typically consider road (i.e. surface) distress resulting from damaging physical processes like[96]:

---

[96]   Other forms of distress may also be important, and potentially maintenance triggering, e.g. ravelling, changes in crossfall etc.

- The development of cracking due to:
    i. Fatigue
    ii. Low temperatures
- Rutting
- Roughness due to subgrade volume change

On the other hand, if we choose a functional approach to road wear, the focus would rather be on the suitability of the road as a piece of infrastructure enabling fast, safe and comfortable road use. From an economic point of view it would be natural to focus on the latter approach, which basically revolves around the ability of the road to serve its major purposes. Since there are multiple dimensions involved when one is to describe the road's ability to serve these purposes, the need of a composite measure of what has been called the road's *serviceability* is importunate. However, the computation of such a composite statistic will inevitably demand a set of weights that must be assigned to each characteristic of the road. The determination of these weights will depend on the intended function of the general statistic.

4.3.2   Factors that affect rutting

Rutting is a result of an interplay between traffic loading and the factual bearing capacity of the road. Apart from the characteristics of the road materials, the latter may also be strongly influenced by climatic circumstances. This is especially so when the pavement subgrade undergoes seasonal variations in bearing capacity, or when bituminious courses are subjected to high temperatures (OECD 1988). Ruts develop within pavement layers when traffic loading causes layer densification and/or when stresses included in the pavement materials are sufficient to cause shear displacements within the materials.

Research has shown that the susceptibility to rutting can be linked to the

following *material attributes* (Archilla and Madanat 2000):

- Excessive asphalt content
- Excessive fine-grained aggregate
- High percentages of natural sand
- Rounded aggregate particles
- Excessive permissible moisture in the mix or in granular materials and soils
- Temperature susceptible asphalt cement
- Cold weather paving, leading to low density

Other factors affecting rutting are (op.cit.):

- Temperature
- Precipitation
- Time, type and extent of loading

In a controlled experimental environment it is possible to study the impacts of these factors, but when studying in-service pavements in a factual environmental setting one is generally not able to measure and control all these variables and mechanisms.

In cold climates road wear from the use of *studded tyres* on bare roads contributes significantly to rutting, and will, on roads with heavy traffic loads, be the most important cause of rehabilitation work being carried out. In South Norway, the use of studded tyres is allowed between November 1st and the first Sunday after Easter.

136

4.3.3 Factors that affect roughness

Roughness is defined as the variation in the longitudinal profile of the wheel-paths. A rough road surface reduces driving comfort and safety, and will also contribute to a higher deterioration rate of the road due to increased dynamic forces from traffic (OECD 1997). There are several factors that could affect the development in roughness, but the link to traffic loads is generally considered to be weaker than for rutting. However, the *interaction* between traffic loads and factors related to climate (temperature extremes, freeze-thaw cycles, precipitation etc.) and characteristics of the road construction are certainly among the candidates of explanatory variables related to roughness.

In the COST 324 project, several European models for pavement performance were reviewed, and among them 5 models that incorporated a relationship for longitudinal profile (EC-DG7 1997). Three of these models used "age" as an explanatory factor, two models included the measured deflection figures, and other factors used in these 5 models were the number of passenger car units, the thickness of the bound layers, the freezing index, and the width of the road.

According to the PARIS-project (reported in Bastiaans 1998a, and in EC-DG7 1999, and reviewed in section 3.3.3), the variation in roughness (measured by IRI) seems to be negligible on fairly solid roads in Europe.

4.3.4 Indicators representing roughness and rut-depth

The International Roughness Index (IRI)

This is by far the most used statistic to describe road roughness. The IRI statistic has its origin from research projects initiated by the World Bank in the early 1980s, and has been implemented in standard procedures for logging of road profiles in most countries around the world (e.g. in 1990 it was made a required statistic to be put into the Highway Performance Monitoring System of the US Federal Highway Administration (Sayers and Karamihas 1998).
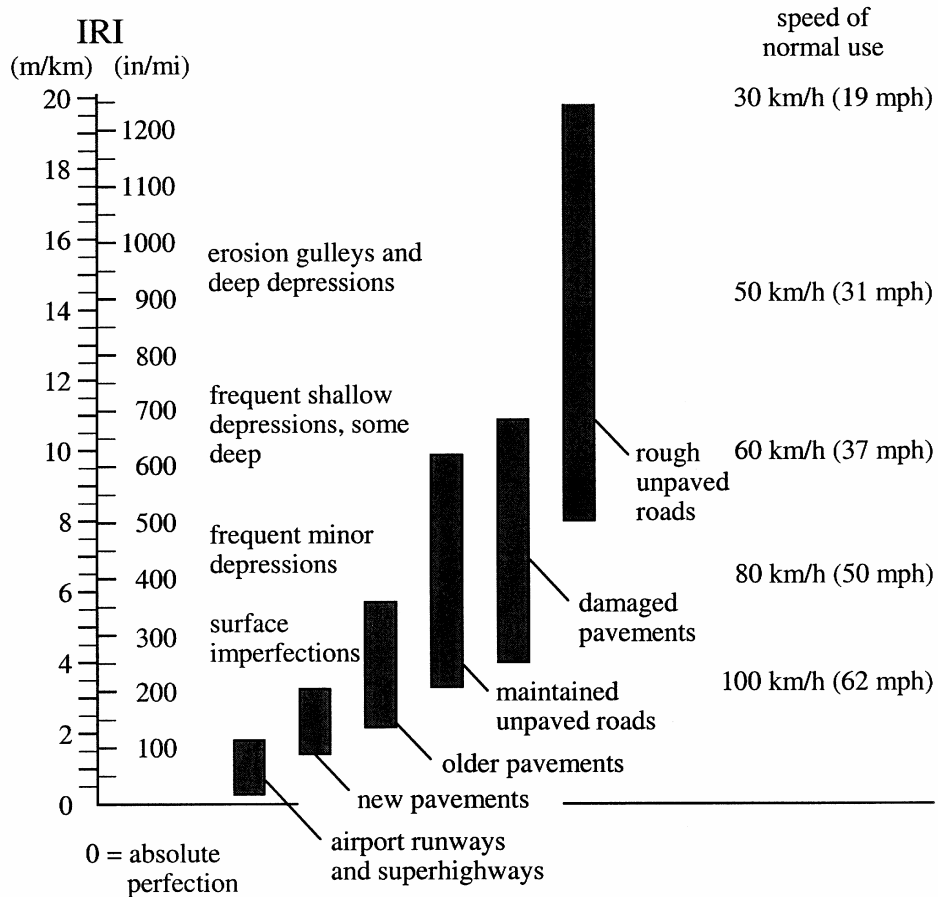
Figure 4-4   Illustration of typical IRI values for different road classes (Sayers and Karamihas 1998)

The index was developed to match the responses of passenger cars, but subsequent research has shown good correlation with light trucks and heavy trucks. Specifically, IRI is very highly correlated to vehicle response variables like:

- Road meter response (relevant for comparability to historical data)
- Vertical passenger acceleration (relevant for ride quality)
- Tire load (relevant for vehicle controllability and safety)

However, it does *not* correlate well with some other response variables like vertical passenger position and axle acceleration (op.cit.).

An IRI value of 0,0 means that the longitudinal road profile is perfectly flat. There is no absolute upper limit to the IRI values, but a road obtaining values above 8 mm/m will be nearly impassable at a speed of 80 km/h (see Figure 4-4).

A more thorough description of the IRI statistic could be found in Sayers (1995). For the purpose of this thesis it is enough to establish the fact that this is a well-established statistic representing the longitudinal roughness well. However, some questions have been raised on the applicability of the statistic for assessing solid roads with rather small changes in IRI-values (see e.g. EC-DG7 1999, section 9.4).

Measuring rut-depth

There is no internationally standardised way of measuring transversal rut depth like the IRI statistic for roughness. The measurements carried out by the NPRA (see section 4.4.2) using profilometers, have applied changing principles over the past decade. The ultrasonic device is composed of 17 ultrasonic sensors mounted across a 2m long beam fixed to the test van. Before 1999 these sensors registered the depth of the outer wheel-path, but from 1999 onwards the peak between the two wheel-paths has been measured instead.

Composite measures of road condition

The ideal statistic in my context would be a composite measure that illustrates the state of the road in more general terms than the specific statistics related to roughness and rutting. However, such a composite statistic will inevitably include a set of weights that must be assigned to each characteristic of the road in order to calculate the composite measure. The determination of these weights will depend on the intended function of the general statistic.

139

There are generally two different approaches to this problem:

- Either the aim is to express how well the road is performing with respect to its major functions as a proper infrastructure for safe, comfortable, and high speed road traffic. This is commonly known as the *serviceability* of the road, an expression introduced in the AASHO Road Test.

- Or the aim is somehow to express the need of maintenance actions. This is (or rather: should be) a highly correlated measure with the serviceability, but it usually includes more absolute minimum or maximum values for the individual statistics entering into the picture. These measures or models are included in various Pavement Maintenance Models.

Under the AASHO Road Test the notion of *serviceability* was developed into more concrete statistics called the Present Serviceability Rating (PSR) and the Present Serviceability Index (PSI). The PSR is a subjective rating originally made by an expert panel which evaluated the pavement condition based on close inspection, the experience of driving over them, and the measures taken from several instruments. A form like the one reproduced in Figure 4-5 was then used to report the overall rating on a scale from 0 to 5. The average rating for a specific road section was then calculated and presented as the PSR for that section.

Based on objective statistics from the same set of road sections, combined with the PSRs, a function for predictions of PSRs based on measured data was estimated. The estimator is then denoted the Present Serviceability Index.

Over the years similar kinds of panel rating procedures have been developed and applied. The basic problem with this kind of statistic is the inherent *subjectivity*, meaning that it is very difficult to transfer these results both in "time and space". A large number of other statistics describing the road condition have been

140

launched and used in various kinds of research work, but it seems that these either correlate very well with the IRI, or they are not well suited for efficient automated measurement procedures (op.cit.) like the NPRA laser-based loggings of rutting and roughness.

Figure 4-5   Example of a typical questionnaire for subjective rating of the road condition (Sayers and Karamihas 1998)

Although it would be preferable to have one composite statistic for road serviceability in my context, it seems that the subjectivity involved in the determination of the weights is an inhibiting factor. At least one would need a solid basis of empirical research and estimations for applying such weights into a common index (similar to the ones used for the valuation of time in a generalized cost statistic). Along this line of argument, I will base my analysis on IRI –values and rut-depth measurements, bearing in mind the rather obvious shortcomings due to the partial nature of these statistics.

4.4 PRESENTATION AND CRITICAL EVALUATION OF THE DATA SOURCES FOR THE FAMAROW MODEL

4.4.1 Traffic data

Length of the observation periods

The composition of traffic at the observation sites is logged by the ATC/WIM-devices described above. The periods of observation add up to more than 6500 weeks with WIM-loggings on these 97 sites, over the period from 1993 to 2000. The length of the observation periods varies considerably from site to site, ranging from 1 week to over 300 weeks of observations. This means that the statistical accuracy of the observations also differs.

24 of the 97 observation sites have been excluded from the analysis due to the following considerations:

- Observation sites with more than 2 lanes are excluded because there is no available information about the distribution of traffic over the lanes.
- Observation sites with less than one week of registrations are excluded because the effective observation period for these sites may only be a few hours on an arbitrary day of the week, thus providing a very biased traffic mix observation.
- Observation sites with known technical problems are excluded (based on information from the NPRA).

These data are only used for calculating average percentages for the different axle- (or vehicle-) groups, and for this purpose even a rather short observation period should give the necessary accuracy.

Observation site	Code	County	1993	1994	1995	1996	1997	1998	1999	2000	2001	Sum
Svinesund	B006	1	3		58	37	62	59	59	49		327
Ørje	B007	1				15	49	58	26			148
Storebaug	B008	1				14	49	53	59	53		228
Feuelstad	BN50	1				4	42	25				71
Høysandvn.	BN5F	1						10	42			52
Solbergkrysset Syd	BN5H	1						6	7			13
Solbergkrysset Nord	BN5I	1						11	7			18
Jarle Syd	BN5J	1						5	11			16
Jarle Nord	BN5K	1						6	10			16
Molteberg	BN5L	1						8	17			25
Patterud	BN5R	1				9	25	5				39
Patterud Syd	BN5S	1				9	61	27				97
Åvangen skole	BN5T	1				11	62	28				101
Nøkkeland Syd	BN5U	1				11	62	28	4			105
Nøkkeland Nord	BN5V	1				2	5	2	9			18
Rom skole	BN66	1								10		10
Ved sengeloftet	BN68	1								10		10
Fossum bru Øst	BN6A	1								12		12
Fossum bru Vest	BN6B	1								9		9
Fosshellinga Øst	BN6C	1								10		10
Fosshellinga Vest	BN6E	1								10		10
Ullens. Hovinmoen	C00M	2						43	21			64
Ørbekk	C05W	2	12	55	61	31	20	52	54	48		333
Strømsenga	C0WA	2	23	51								74
Varingskollen S	CN5J	2	38	59	59	43	11					210
Varingskollen N	CN5K	2	23	59	48	24		8	4			166
Julsrud	CR0B	2					5	18	30			53
Espa	D001	4	10	42	61	21	21	63	50	63		331
Motrøa S.	D004	4								60	1	61
Ebru	D008	4							60	63	1	124
Basterud	D00E	4					15	62	61	63		201
Brevad	D06F	4								60		60
Ebru/Husum	DN5F	4	11	9	21							41
Vikselv Syd	DN5H	4							39			39
Skavabakken	DN5J	4							32	42		74
Kolomoen	DN5L	4							24	25		49
Dombås Sør	E02U	5						9				9
Otta S	E030	5	11	38	59	43	60	63	60	63		397
Skarsmoen	E0BZ	5		35	63	57	55	62	60	54		386
Losna	EN5F	5					13	8	7			28
Fåvang S	EN5G	5					12	16	9			37
Ringebu S	EN5H	5					24	18	7			49
Frya	EN5I	5					56	16	4			76
Hundorp	EN5J	5					11		8			19

Figure 4-6a   Registration periods for the WIM/ATC data

143

Observation site	Code	County	1993	1994	1995	1996	1997	1998	1999	2000	2001	Sum
Kvam S	EN5K	5					14					14
Sjoa S	EN5L	5					25	16				41
Vinnes	FN5D	6		21	61	16						98
Drammensbrua	FN5I	6				13	35	38	28			114
Kjerringland	IN5E	9					5	4				9
Trøe	IN5G	9					6		2			8
Ganddal v/Telefonvn.	LUV7	11						5				5
Ganddal v/Olsokvn	LUV8	11						5				5
Kleppe Nord	LUV9	11						3				3
Ree Nord	LUVA	11						3				3
Ree Sør	LUVB	11						3				3
Gudvangen	S004	14							27	36	1	64
Fodnestunnellen S	SN5D	14							3	61		64
Fodnestunnellen N	SN5E	14							3	61		64
Førde	SNXX	14		45	11							56
Alsgård	W04Y	18						4	31			35
Gulli v/Tønsberg	ZN5H	7			33	60	57	38	40			228
Hem Undrumsdal	ZN5I	7		16								16
Sande	ZN5J	7		12	17	23	37	60				149
Smørstien	ZN60	7				21	45	62	14			142
Sum			131	442	552	464	944	1010	929	862	3	5337

Figure 4-6b   Registration periods for the WIM/ATC data, continued

Geographical spread

As can be seen from Figure 4-6, most (88%) of the observation sites are located in counties 1-7 which are all located in South-East Norway (see map in Figure 4-7). This is partly due to practical considerations related to the fact that trials with WIM-registrations are led from the Norwegian Public Roads Administration in Oslo, but also because the roads in this region are among the most congested in Norway.

From a statistical point of view, the selection of observation sites may bias the results mainly due to two aspects:

- *Firstly*, the sample is located in an area with less precipitation and possibly more extreme temperatures (summer and winter) than the average for all the national roads. Typically there is more precipitation and a milder climate on the West-Coast of Norway. However, winter temperatures in North Norway may be even more extreme than in the South East.

144

- *Secondly*, most of the observation sites are located on roads with a higher traffic volume than the average national road. Higher traffic volumes are also followed by more solid road constructions, and more frequent maintenance.
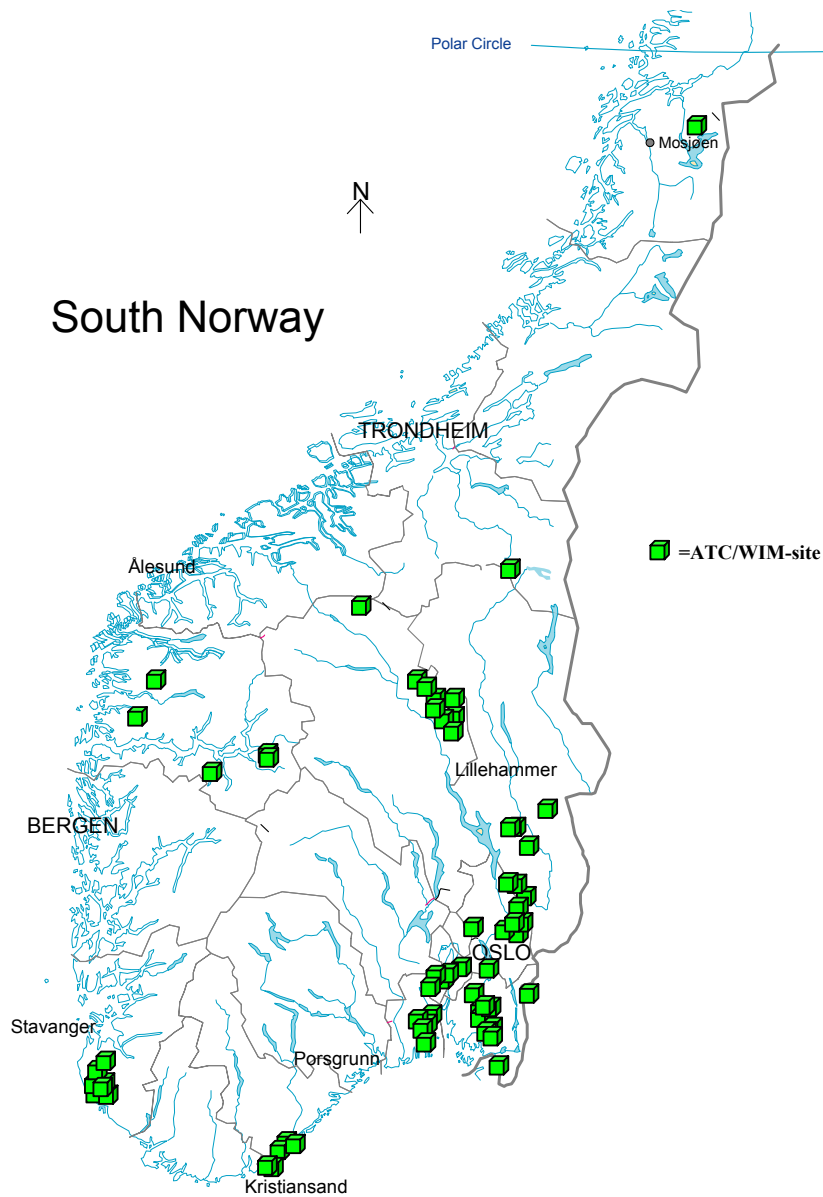
Figure 4-7 The location of the ATC/WIM sites in Norway
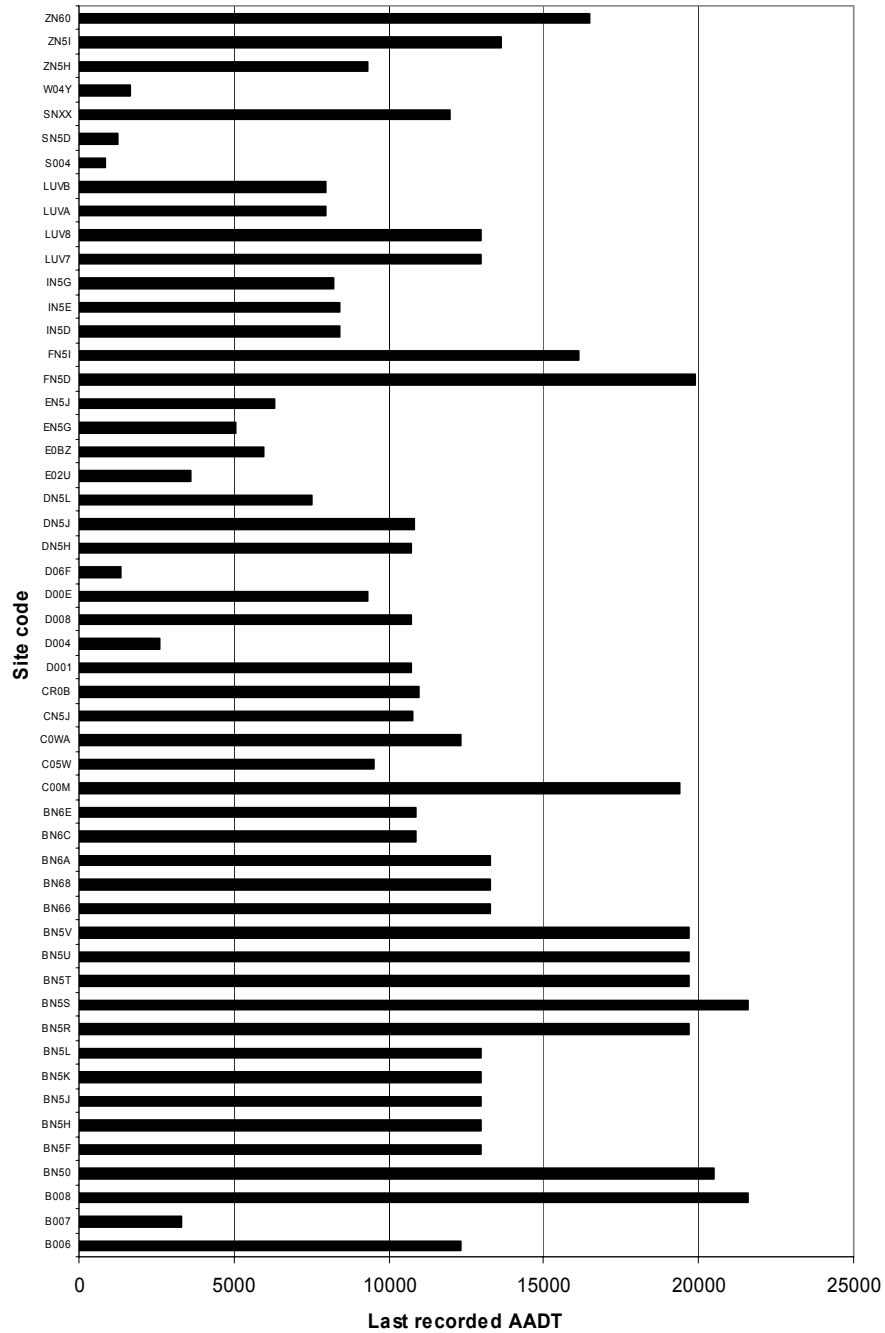
Figure 4-8   Last recorded average annual daily traffic (AADT) volumes at the observation sites[97]

_____

[97] Source: The Norwegian Road Data Bank, observed traffic level from 1998, 1999 or 2000.

146

Traffic levels

The annual average daily traffic (AADT) volumes of the observation sites are presented in Figure 4-8. Most observation sites have a daily traffic level of more than 3000 vehicles, and many sites have a traffic level above 10 000 vehicles per day. As 90 percent of the Norwegian National Road network has an AADT of less than 5000, it is quite clear that my sample mainly represents rather highly trafficked roads according to typical Norwegian standards. The lack of low volume roads in the database also makes the results less applicable to the networks of municipal and county roads.

Regression:
$$AADT = \alpha + \beta_1 * Days + \beta_2 * Days^2$$

○ = Actual AADT logging
◁ = Predicted AADT value
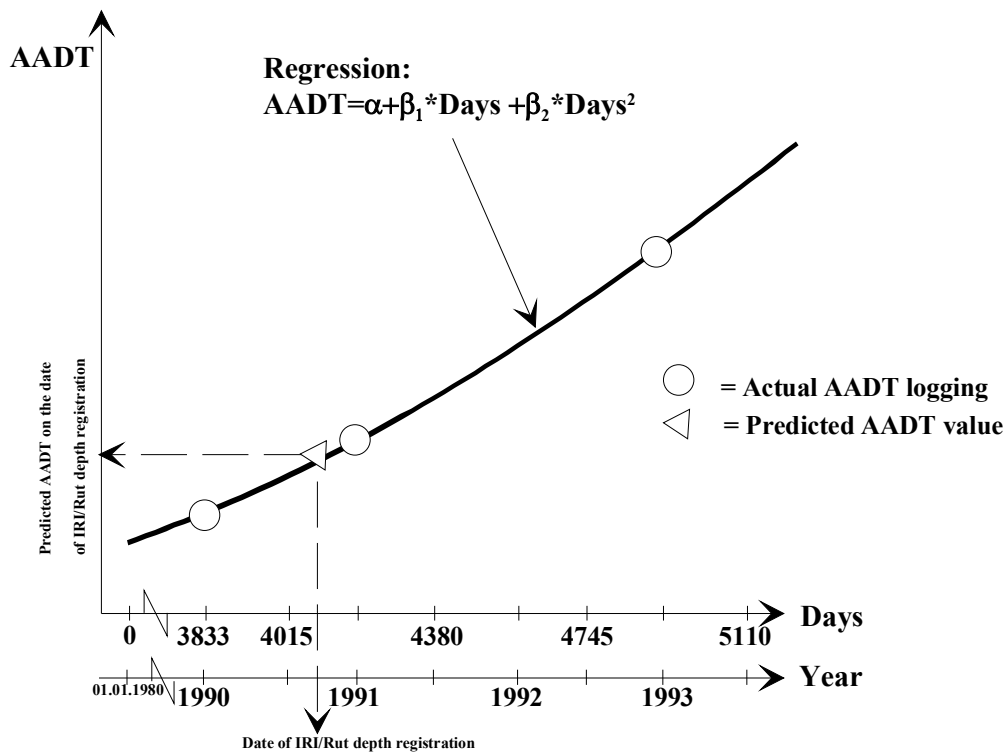
Figure 4-9   Illustration of the principle of predicting the level of traffic at the date of road wear logging

Site-Code	Constant	Beta-1	Beta-2	Mult R-sq
B006	25623.66	-7.98	0.00086	0.99555
B007	11791.39	-3.60	0.00034	0.99361
B008	28749.90	-7.20	0.00087	0.99222
BN50	23133.35	-4.85	0.00063	0.98837
BN5F	28359.96	-8.07	0.00084	0.96950
BN5H	28359.96	-8.07	0.00084	0.96950
BN5J	26751.75	-7.60	0.00080	0.93681
BN5K	26751.75	-7.60	0.00080	0.93681
BN5L	26751.75	-7.60	0.00080	0.93681
BN5R	24855.77	-5.24	0.00063	0.97991
BN5S	46650.48	-12.86	0.00131	0.98951
BN5T	38202.44	-10.34	0.00109	0.97822
BN5U	38202.44	-10.34	0.00109	0.97822
BN5V	38202.44	-10.34	0.00109	0.97822
BN66	61.67	3.17	-0.00018	1.00000
BN68	61.67	3.17	-0.00018	1.00000
BN6A	61.67	3.17	-0.00018	1.00000
BN6C	24232.04	-5.87	0.00056	0.99091
BN6E	24232.04	-5.87	0.00056	0.99091
C00M	26242.81	-6.18	0.00071	0.99259
C0WA	16572.21	-3.15	0.00035	0.96041
CN5J	14402.02	-2.04	0.00021	0.92992
CR0B	3507.09	0.97	0.00000	0.99721
D001	11675.86	-1.77	0.00022	0.99251
D004	3892.58	-0.61	0.00006	0.86712
D008	9223.46	-0.71	0.00012	0.98417
D00E	14275.36	-2.65	0.00027	0.92418
D06F	1548.59	-0.03	0.00000	0.63722
DN5H	11565.83	-1.73	0.00022	0.98879
DN5J	11097.30	-1.54	0.00020	0.99796
DN5L	3519.92	0.45	0.00001	0.96674
E02U	3301.29	0.05	0.00000	0.50617
E0BZ	2828.16	0.83	-0.00006	0.90112
EN5G	6683.06	-0.62	0.00006	0.38436
EN5J	2068.99	1.29	-0.00010	0.92776
FN5D	127697.33	-40.39	0.00374	0.97842
FN5I	-18506.85	11.06	-0.00086	0.93225
IN5D	9146.11	-1.19	0.00015	0.99016
IN5E	9165.34	-1.23	0.00016	0.98954
IN5G	7835.19	-0.92	0.00014	0.99348
LUV7	5983.29	1.00	0.00001	0.90490
LUV8	3699.39	1.78	-0.00006	0.92068
LUVA	13967.37	-3.87	0.00045	0.90781
LUVB	6515.66	-1.07	0.00019	0.83950
S004	-8044.85	2.94	-0.00024	0.88480
SN5D	-3595.36	1.65	-0.00013	1.00000
SN5E	2686.57	-0.40	0.00003	0.82070
W04Y	-482.65	0.54	-0.00003	0.98383
ZN5H	6260.24	0.38	0.00001	0.96462
ZN5I	16615.75	-2.21	0.00025	0.78583
ZN60	10431.31	0.52	0.00005	0.87370

Figure 4-10   Estimators from the AADT-regressions

The observations of traffic levels are not carried out every year, but represent annual averages based on countings carried out at arbitrary dates by the local road authorities. For my purpose I need to predict the traffic volumes on the days where the road wear data were collected, or to be more exact: The amount of traffic that has passed the observation site between the loggings of the road wear (roughness and rutting). For this purpose I have estimated a simple regression equation for each observation site based on the present AADT time series.

I found that Equation 4-1 fitted the traffic data for this observation period well.

Equation 4-1 $\qquad AADT = \alpha + \beta_1 \cdot Days + \beta_2 \cdot Days^2$

Days is a variable denoting the number of days between January 1st 1980 and the observation day (calculated in the same way). The principle for predictions of AADT at the date of the road wear logging, is illustrated in Figure 4-9.

From Figure 4-10 one can see that most regressions obtain a very good fit to the AADT observations with this rather simple model specification. Few regressions turn out with a multiple R-squared below 0.9. For the few observation sites where this model has a poor fit, the models are substituted by the models for the geographically closest observation site along the same road.

The predicted AADT-levels are subsequently multiplied by the corresponding average site-specific traffic distribution data from the WIM-ATC deployments to obtain a number of axle passes (or a number of vehicle passes) in each weight category.

Some descriptives from the WIM-data collected
In Figure 4-11 I have summarized some information about the WIM-data. The collection of observation sites *does* provide a fairly good spread with respect to

traffic mix, as the proportion of car axles varies from 49% to 87%.

Descriptive	
Average proportion of car axles[98]	78%
Minimum proportion of car axles	49%
Maximum proportion of car axles	87%
Average proportion over-weight axles[99]	1%
Average share of tandem axles	4%

Figure 4-11   Sample descriptives of the WIM data[100]

Together with the traffic data now described, the road wear measurements constitute the most important input to my model. I will now turn to a closer description of the rut-depth and roughness measurements.
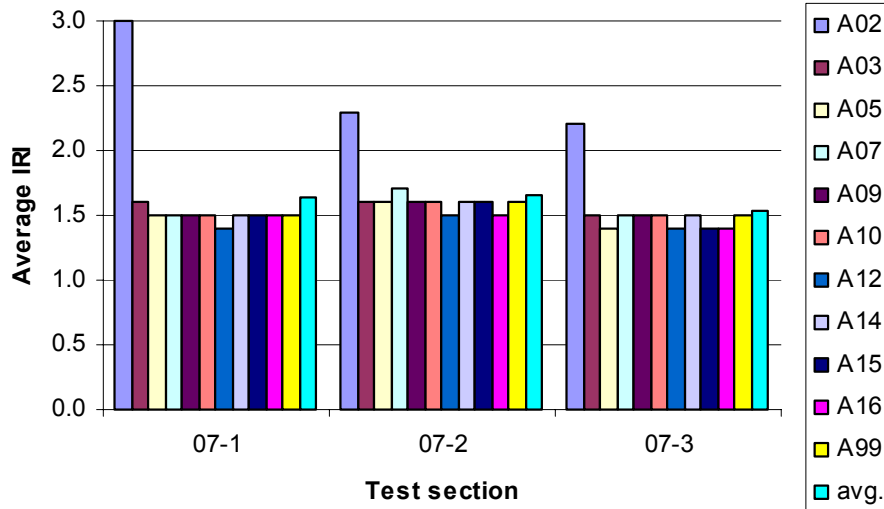
Figure 4-12   Comparison of IRI-loggings by 11 different ALFRED measurement vehicles on three different test sections, April 2001[101]

---

[98]   NOTE: Not weighted by traffic level.

[99]   Calculated as a percentage of all axles above 2 tonnes.

[100]   Based on the sample applied for the model estimation efforts reported at the end of this section.

[101]   Source: Unpublished spread-sheet from Torleif Haugødegård, NPRA

150

4.4.2   Roughness and rut depth measurements

The Norwegian Public Roads Administration (NPRA) has measured roughness (IRI) and rut depth on the Norwegian network of national roads for many years. Most of these measurements are registered in the Road Data Bank, and it is thus possible to extract time series describing the development of road wear. The measurements are now, as a rule, carried out annually. The dates vary from year to year, and some years are also missing from the records.
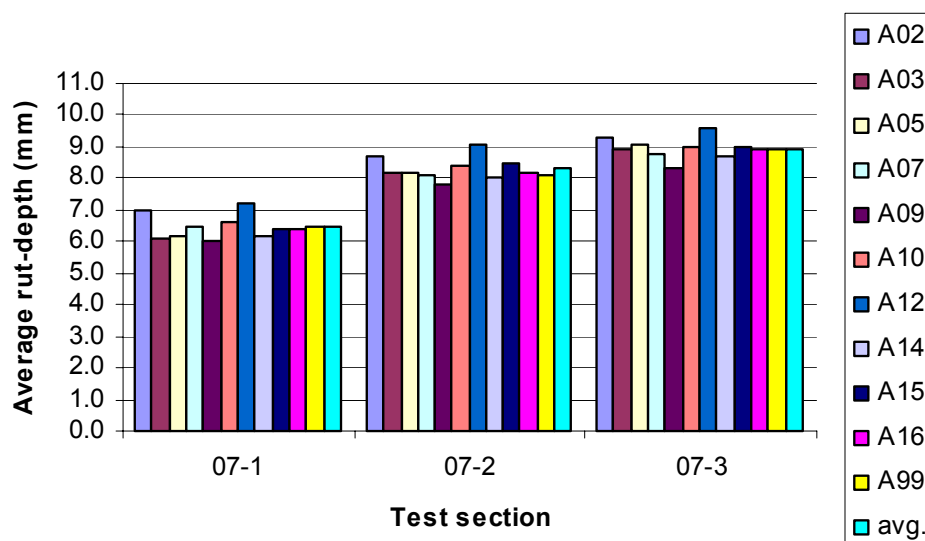
Figure 4-13   Comparison of rut depth loggings by 11 different ALFRED measurement vehicles on three different test stretches, April 2001[102]

From 1997 onwards, a combined laser / ultrasound device fitted to the front of a van is used for these measurements. Before 1997 an ultrasonic device was used. The laser / ultrasound device has one laser and 17 ultrasound-censors, plus censors that control the movements of the van. The technology is developed for the NPRA in co-operation with SINTEF[103].

---

[102]   Source: Unpublished spread-sheet from Torleif Haugødegård, NPRA

[103]   Source: Torleif Haugødegård, NPRA and Even Sund, SINTEF (The Foundation for Scientific and Industrial Research at the Norwegian Institute of Technology)

Figure 4-14   Photo of van with ALFRED laser/ultrasound device for roughness and rut depth loggings[104]

Unfortunately, I have not been able to find any up-to-date documentation on the *measurement accuracy* of these vehicles. However, some early tests are reported in Haugødegård (1993). The tests conducted on new pavements indicate a measurement accuracy of +/- 1 mm on average rut depth measurements, and differences between the tested measurement vehicles of 0.3-0.5 IRI-units. However, the report concludes that there were significant differences in repeated measurements of the longitudinal profile that made the IRI-measurements rather inaccurate. However, both hardware (laser) and software have been upgraded since these early tests, and the measurement accuracy for IRI-loggings is reported to have improved resulting from this[105].

---

[104]   Photo: SINTEF

[105]   Source: Torleif Haugødegård, NPRA.

152

In Haugødegård (1995) the same measurement equipment was tested once again on a new pavement with more "closed structures" than the ones tested in the earlier experiment. The IRI-loggings turned out to be more accurate for these pavement surfaces.

The vehicles are spread over the country most of the year, but once a year they are calibrated at a central gathering. The only recent test of the measurement performance of the equipment stems from such gatherings. The outcome of a comparison of the test-vehicles in 2001 is illustrated in Figure 4-12 and in Figure 4-13. These are the average values for roughness and rutting logged by the 11 vehicles tested on three different test-road stretches (07-1, 07-2 and 07-3). There is one significant "outlier" in the IRI-loggings: Vehicle A02 reported extremely high IRI-values compared to the other vans. According to Torleif Haugødegård of the NPRA, this was due to a miscalibration that was subsequently corrected. Unfortunately, these figures only enable a comparison between the vehicles, and not relative to a test section with known characteristics regarding rutting and roughness.

Based on the reported figures from this April 2001 test, I have calculated the standard deviations reported in Figure 4-15.

Test Section	Standard Deviation IRI	Standard Deviation Rut-depth
07-1	0.4545	0.3668
07-2	0.2207	0.3606
07-3	0.2284	0.3297

Figure 4-15   Standard deviations for all the ALFRED test vehicles (IRI and Rut-depth)

As pointed out above, vehicle A02 is a significant "outlier" and this very much affects the standard deviations for the IRI measurements. Without this outlier the standard deviations are fairly low (see Figure 4-16).

Test Section	Standard Deviation IRI	Standard Deviation Rut-depth
07-1	0.0471	0.3381
07-2	0.0568	0.3534
07-3	0.0516	0.3259

Figure 4-16   Standard deviations for all the ALFRED test vehicles, except A02 (IRI and Rut-depth)

At the most recent calibration gathering all vehicles were tested over the same road section, and each vehicle was also driven three times over the section in order to check the consistency of the measurements.
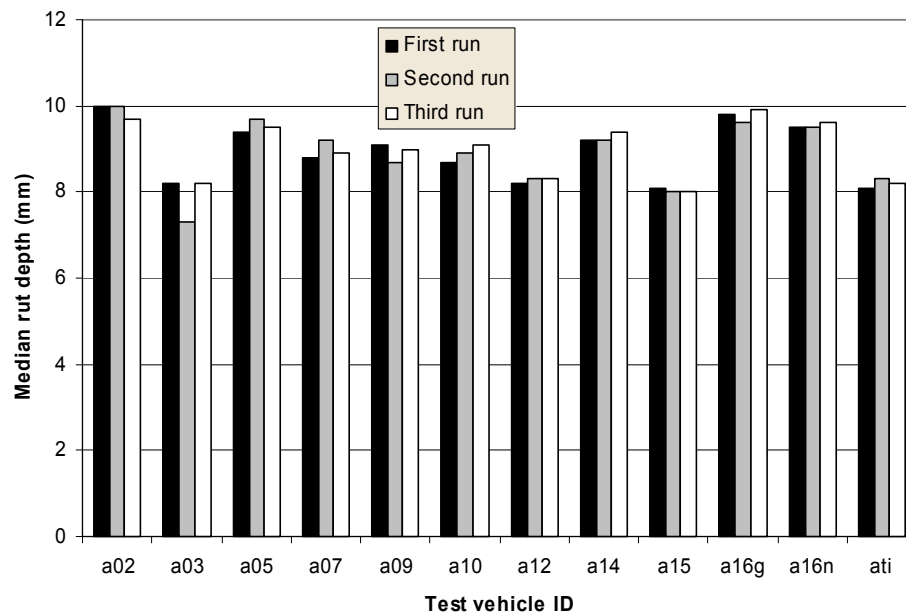
Figure 4-17   Median rut depth loggings for the whole road section, ALFRED calibration meeting, May 2002[106]

In Figure 4-17 the measured median rut depth for the whole test section is

---

[106]   Source: Test results from May 7th 2002, collected at road RV114 in Sarpsborg, Norway. Information supplied by Torleif Haugødegård, NPRA

154

presented for all vehicles, and for the three test runs. Ideally all these measurements should be equal. The consistency of the measurements within each test vehicle seems to be fairly good, with only small deviations between each test run (i.e. typically less than 0.3 mm deviation). The differences between the test-vehicles do, however, seem to be quite high (up to 2 mm, or 25%). Since these figures represent the median for the whole 1 km road section, and loggings were done for each metre, one would expect that most random factors were "filtered out".

The impression of measurements being quite consistent when repeated measurements are done with one test vehicle, is corroborated when plotting the 20m loggings for all three test-runs against one another. Both the rut-depth figures presented in Figure 4-18, and the IRI figures in Figure 4-19 show very small deviations from one test run to another, even when the individual loggings are plotted.

In Figure 4-20 and Figure 4-21 the average loggings for each location are plotted for all the different test vehicles, showing rut depth measurements and IRI values, respectively. For the rut depth measurements, it seems that a few vehicles form the "outliers" quite consistently. Vehicle A03 seems to represent the lowest loggings at most locations, while vehicle A02 is among the top loggings at most locations.

Figure 4-20 may also give an indication of increasing differences with increasing rut depth (see loggings around location 5460). The picture is somewhat less clear when studying the IRI-loggings in Figure 4-21, and it is hard to identify vehicles that constitute the bottom or the top for most loggings.
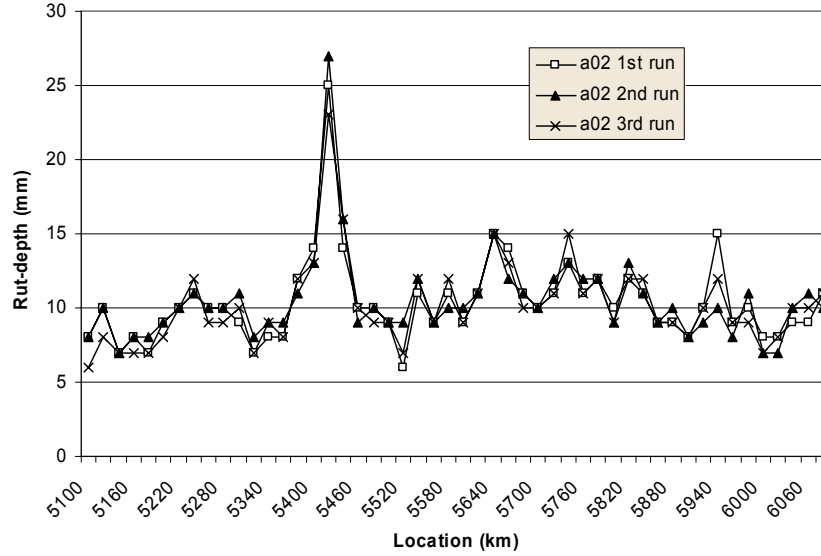
Figure 4-18   Repeated rut depth loggings over the same road section with vehicle A02[106]
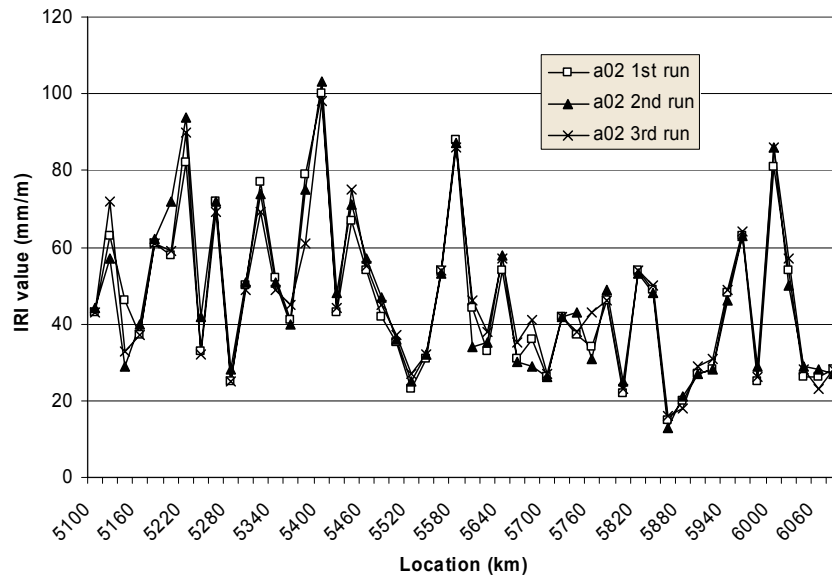
Figure 4-19   Repeated IRI loggings over the same road section with vehicle A02[106]

156
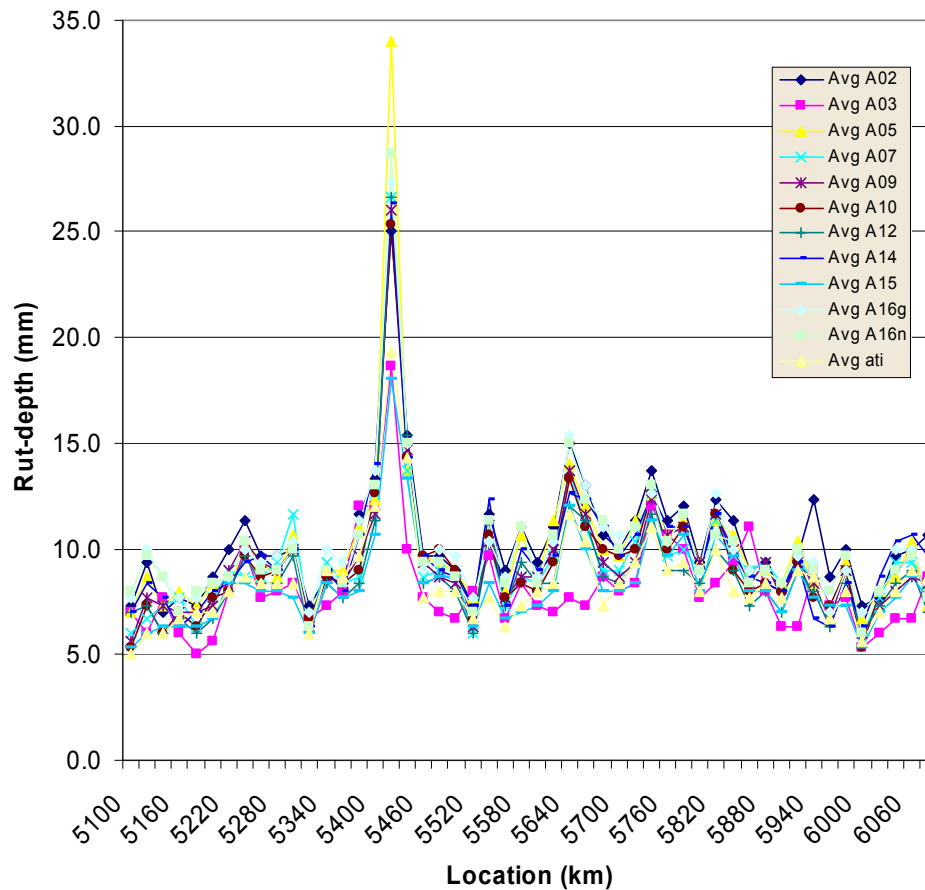
Figure 4-20   Comparison of average rut-depth measurements by all tested vehicles after three test-runs over the same road section[106]

This is how far we can get from these comparative tests. The overall picture is that measurements seem to be fairly consistent with respect to repeated measurements with the same vehicle, but there are significant differences between the vehicles. Once again I would like to point out that these tests should ideally be conducted on test sections with known characteristics with respect to rutting and roughness. In an ideal world, we would also like to have information about how the measurement accuracy develops over time (in between calibrations).
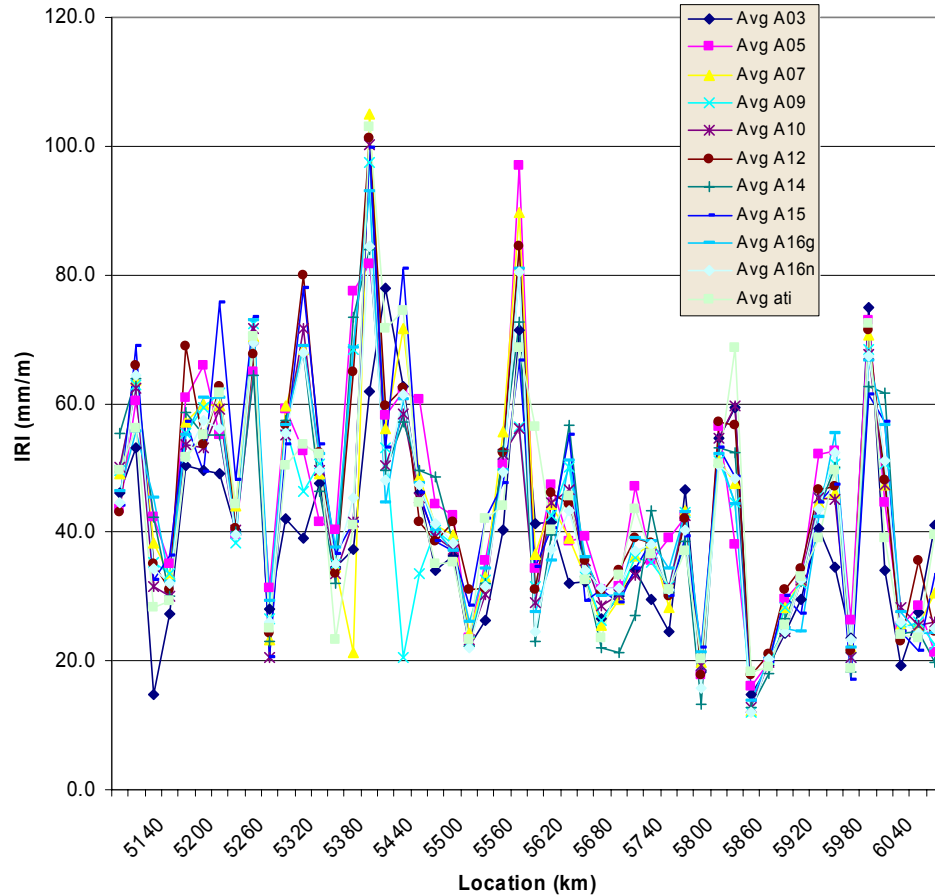
Figure 4-21   Comparison of average IRI measurements by all tested vehicles after three test-runs over the same road section[106]

Rut depth data are actually logged meter by meter by the test vans. These loggings are subsequently converted into a 20-meter average when entered into the Road Data Bank. IRI values are logged and entered for every 20 meters. In my analysis I have extracted these measurements of roughness and rutting from a 200 meters stretch covering the ATC/WIM-deployment sites +/- 100 meters, and calculated an average value for both IRI and rut-depth.

In Figure 4-22 and Figure 4-23 some examples of IRI and rut-depth measurement series are shown for selected observation sites in some counties. The development

158

patterns seem rather confusing when presented like this. For some observation sites the IRI-values fall and rise quite rapidly. Our fundamental expectation would be that both IRI and rut-depth values would rise with time given there were no maintenance actions taken in the time span.
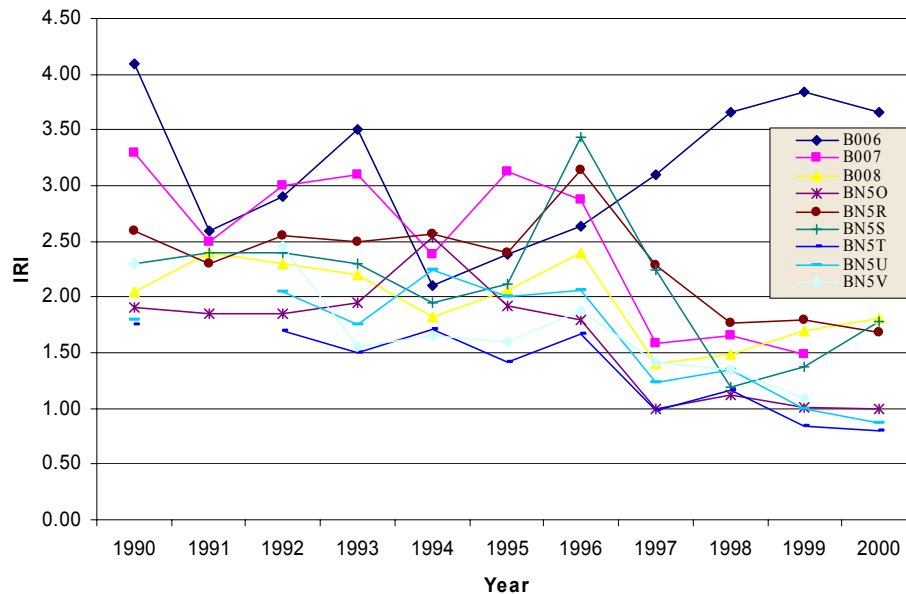
Figure 4-22   Examples of IRI-measurements for observation sites in Østfold county

Another reasonable expectation would be that the frequency of maintenance actions should increase with increasing traffic levels. Graphically this means that we would expect frequent cases of falling IRI and rut-depth values (i.e. improved road surfaces) for the observation sites with high traffic levels. If we confer the traffic levels given in Figure 4-8, most of the sites represented in Figure 4-22 have fairly high traffic levels (around 20 000 AADT). The exceptions are B006 and B007 which have AADTs of 12 000 and 3 000, respectively. B006 seems to have two intervals with falling IRI measurements, whereas B007 seems to have five

such cases[107]. The loggings for the other sites seem to have between four and six cases of falling IRI values.

The picture is not very different for the rut-depth loggings given in Figure 4-23. Again we see a mixture of long positive developments in rut-depth at some observation sites, and some negative developments at others. Compared to the IRI-loggings there are many more sites with rather long unbroken positive developments in the measurements.

Maintenance and rehabilitation actions will of course break the pattern of increasing rut-depths and roughness loggings, so there is a need of exploring the possibilities of identifying the timing of such actions. The Norwegian Road Data Bank also contains a registry of maintenance actions. This should make it possible to exclude from my database observation periods where such actions have been taken. The ideal situation would then be having the rutting and roughness of the newly repaved or milled road logged, but unfortunately there is no such routine. We will therefore have to omit all observation periods with pavement maintenance actions being taken. This will, unfortunately, reduce the number of observation periods in my database.

The pavement maintenance registry of the Road Data Bank contains information about:

- Identification of the relevant road section
- Action date
- Type of material applied (e.g. asphalt mix)
- Layer thickness (i.e. quantity of material consumed)
- Type of action (reseal, patching, milling)

---

[107] The last case is marginal and may very well be caused by low measurement accuracy.

160
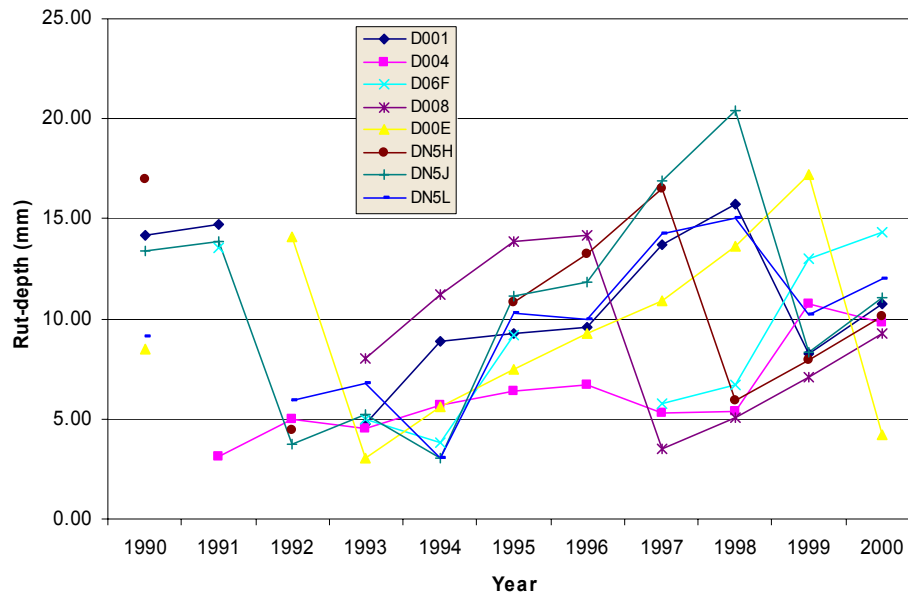
Figure 4-23   Examples of Rut-depth measurements for observation sites in Hedmark county

Having consulted some of the persons responsible for maintaining this registry at several county road administrations, there is reason to suspect that this registry does not cover all maintenance actions taken. Since we would expect such an action to have happened whenever we observe a falling IRI or rut-depth value[108], it is to a large extent possible to control this.

Plotting the registered maintenance actions against the time series roughness and rutting data should give an indication of the magnitude of the problem. This is done in Figure 4-24 and in Figure 4-25. Since the loggings of roughness and rutting are done at different times each year, I have chosen to express the dates as "days after January 1st 1980" (this is also done in the regression equations

---

[108] Theoretically there could be enough traffic between the maintenance action and the IRI-observation to create a positive development in IRI values. Considering the prevailing levels of traffic, this does not seem like a plausible explanation.

following).

The first figure shows that there are not always maintenance actions registered whenever there is a drop in IRI-values from one year to another. From 1990 to 1991 the IRI-values fall for sites B006, B007 and BN5O. There is no registered action taken for B006 and B007. A thin repaving is registered for B008 and BN5O. In this period there was an increase in the IRI-values for site B008.

The rest of the cases in Figure 4-24 show the same lack of cause and effect pattern between maintenance actions and development patterns related to roughness.
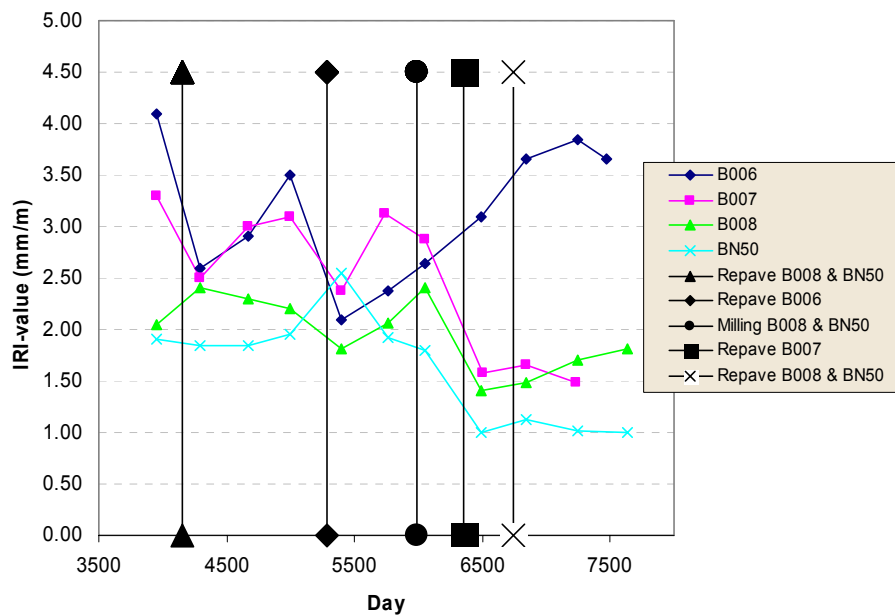
Figure 4-24   Examples of registered maintenance actions plotted against IRI-loggings for four selected observation sites in Østfold county

One would expect to observe a stronger correlation pattern between maintenance actions and the development in rut-depth, since typical maintenance actions like

162

repaving and milling would have a more direct effect on rutting than on roughness. Examining Figure 4-25 gives a slightly improved pattern of correlation, but there are several instances where there is either a diminishing rut-depth combined with "no registered maintenance action", or instances where maintenance has been carried out, and still rut-depth numbers are rising. One example might be a registered repaving in 1997/98 and still significantly increasing registered rut-depth at the B008 site over the same period. Examining this case more closely reveals that milling was performed on this stretch some time before the repaving, resulting in a very significant drop in rut-depth measurements. The following "mismatch" with the somewhat rising (but still low) rut-depth figures combined with the repaving action taken may be a result of measurement error in the logging of rut-depths.

The overall picture from this small sample of rut-depth loggings from Østfold county, is that the major drops in rut-depth measurements are explained by registered maintenance actions. Apart form the example mentioned above, it seems that there are no cases where there is a combination of maintenance actions being performed and falling rut-depth loggings.

Based on this review of my data sources related to road wear measurements and maintenance actions, it seems that there probably are some serious shortcomings in my basic data related to the IRI-loggings. To some extent this may be due to a possible lack of compliance with the directions to properly register every maintenance action. Having consulted some of the persons responsible for such registrations in the affected local road administrations, it seems that some offices meticulously register every bit of maintenance action, whereas others only register the main reseals. Since minor actions, such as slot filling, also could affect our measurements significantly, this may be an important problem related to the quality of my data. Another possible explanation is a poor accuracy in the measurements of roughness, creating questionable development patterns.

163
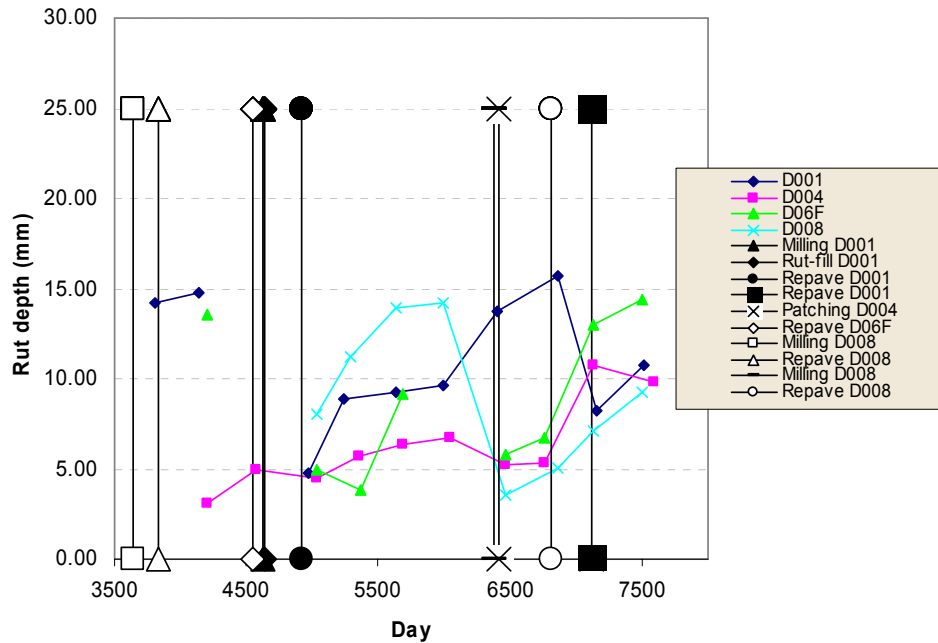
Figure 4-25   Examples of registered maintenance actions plotted against Rut-depth-loggings for four selected observation sites in Østfold county

However, it seems that the consistency of the Maintenance Registry of the Road Data Bank, and the rut-depth loggings are much better than is the case for the roughness loggings. This may not be very surprising, for two different reasons which have been mentioned earlier in this review of the quality of my data. *Firstly*, we have seen that the consistency, and the resulting measurement accuracy of the IRI-measurements performed with the ALFRED-vehicles are not very good. This means that the observed "mismatches" between rising IRI-figures and maintenance may actually be within the limits of typical measurement errors. The rut-depth measurements seem to achieve a higher level of accuracy, and may not suffer from such errors to the same extent. *Secondly*, typical maintenance actions like milling, repaving, rut-filling and patching would have a much higher impact on the observed rutting of the road than on the roughness.

164

Apart from the possible shortcomings of the data resulting from measurement errors and holes in our maintenance register, the basic assumption of a positive correlation between traffic levels and road wear may also be questionable under certain circumstances. One illustrating example has been reported from one of the regional departments of the NPRA[109]: Here one has observed a change in the lateral positioning of the vehicles over the seasons. This may actually even the road surface over a specific time span, thus creating a negative development in rutting figures. Generally, just observing traffic levels and developments in two different measures of road wear, is a very simplistic approach to the problem. A proper analysis should control for a number of other important factors as well.

The interplay between the structure of the road construction and climatic factors (precipitation, temperature levels, temperature swing, sunlight etc.) may be the most important additional factor. Frost heaving cycles may also strongly influence road wear, and possibly dominate the basic relationship between traffic loads and road wear.

Measurements of road characteristics carried out *under* the spring thaw period would   particularly disturb my search for a correlation between traffic loads and road wear. Both rutting and roughness might tend to be higher under these periods than the *permanent* deformations of the road. In Figure 4-26 I have plotted the frequency of observations by week number for all registrations of IRI and rut-depth over the first part of the observation period. Spring thaw problems typically arise in March and April, i.e. between week 9 and 16. From the figure it can be seen that this should not seriously affect the analysis, as a very low proportion of the observations are carried out as early as this. The major part of the measurements is carried out between week 18 and 43 (i.e. between May and October). Normally this is a period without severe frost in the relevant regions.

---

[109] Source: Jostein Myre, NPRA Akershus
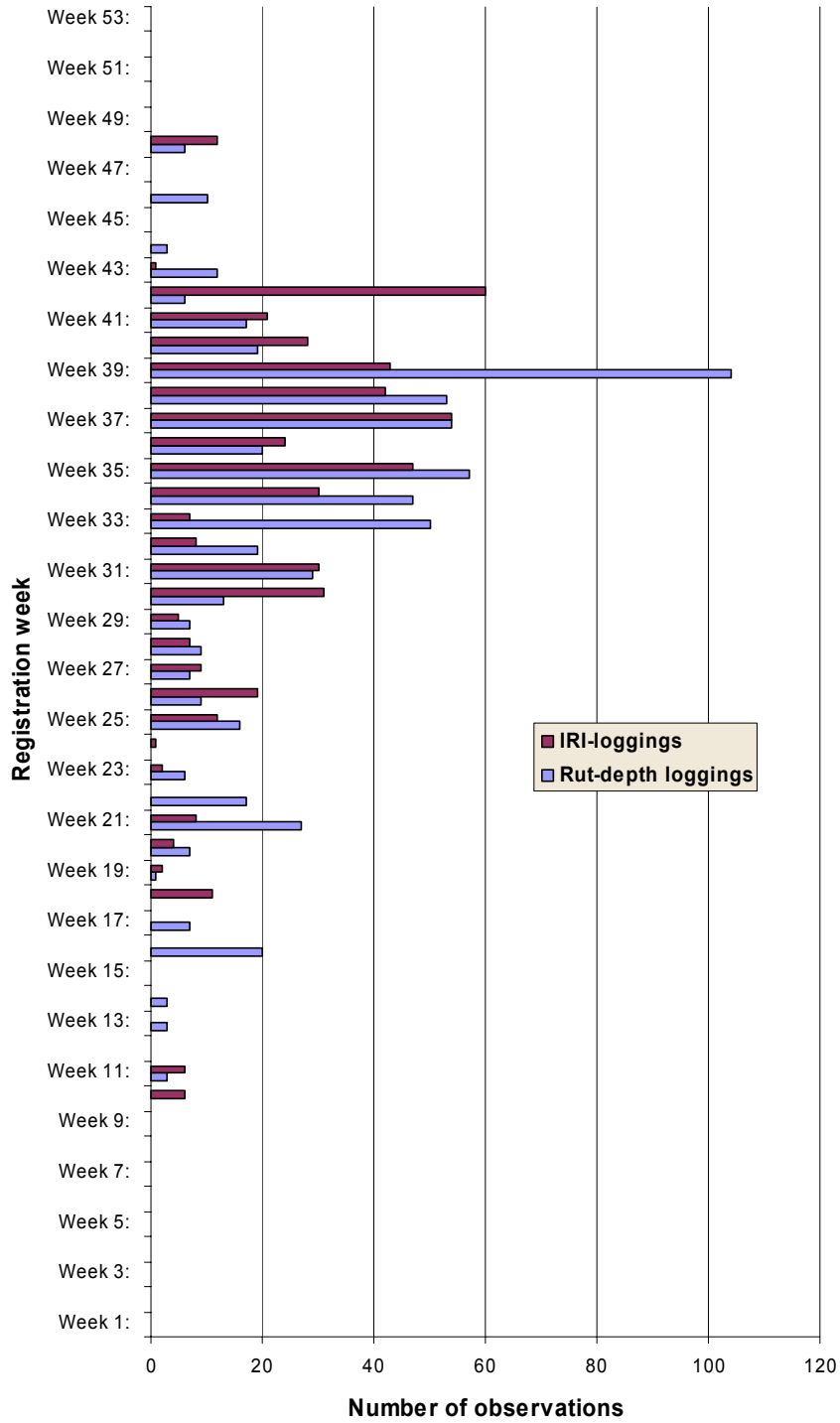
Figure 4-26   Seasonal distribution of the IRI and rut depth loggings at the observation sites (1993-1998)

The observed variation between the performance of the test vehicles used to register IRI and rut-depth values, is only one of many possible sources of measurement error, including the following:

- Errors related to the *primary data collection* phase (technical error related to the laser and ultrasound equipment, the sensor controlling of vehicle movements, the software and hardware that treat the signals received)
- Errors related to the *registration* of the measurements into the Road Data Bank (punching error, linking to relevant road identification)
- Errors related to the *reporting* from the database, and the subsequent *data processing* (e.g. it is not possible to use macros for reporting from the database, and this allows for different procedures being used for producing the output from different registers (counties, years)). Since all output is produced county by county and year by year, a number of Visual Basic macros have been written to put this together in an integrated form (see the Appendix).

The first category of measurement errors is already accounted for to some extent by the calibration test data presented above. The second type, related to registration is beyond my control, but since most of these procedures have been automated, one may hope that there is very limited room for error in this phase. A thorough scheme of spot tests and comparing input and output-data related to the Visual Basic data processing, should limit the probability of the last category of errors occurring.

4.4.3   Climatic data

The third main source of data (in addition to the traffic data and the road wear measurements) are the climatic data supplied by the Norwegian Meteorological Institute (DNMI). Because climatic factors are among the possible explanatory variables related to road wear, I have collected data for precipitation, snow-depth, snow-cover and temperatures.
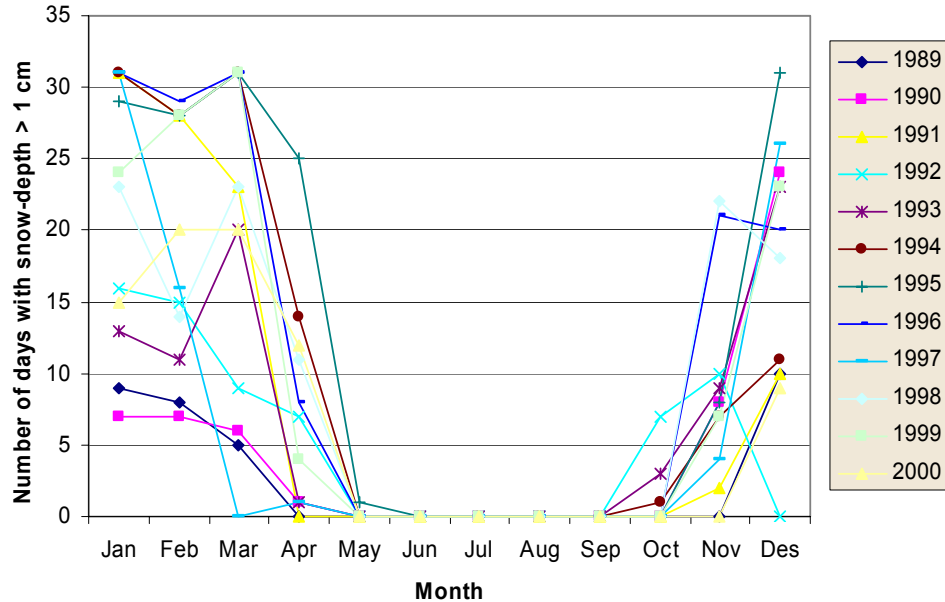
Figure 4-27   Number of days with snow at Igsi / Hobøl, Østfold

The climatic conditions at the observation sites would be most relevant to my analysis, but no registrations are available at these sites. However, the Road Data Bank contains information about the closest DNMI[110] weather observation site for each traffic observation site. Based on an update of this registry, data were commissioned from DNMI for the relevant set of sites. The proximity of the DNMI observation sites to the traffic observation sites varies considerably, and the distance is a significant source of error.

To give a picture of the variations in the climatic factors recorded, I have given some graphic illustrations of snow-depths, temperatures and precipitation at a few selected observation sites.

---

[110] DNMI=The Norwegian Meteorological Institute

168
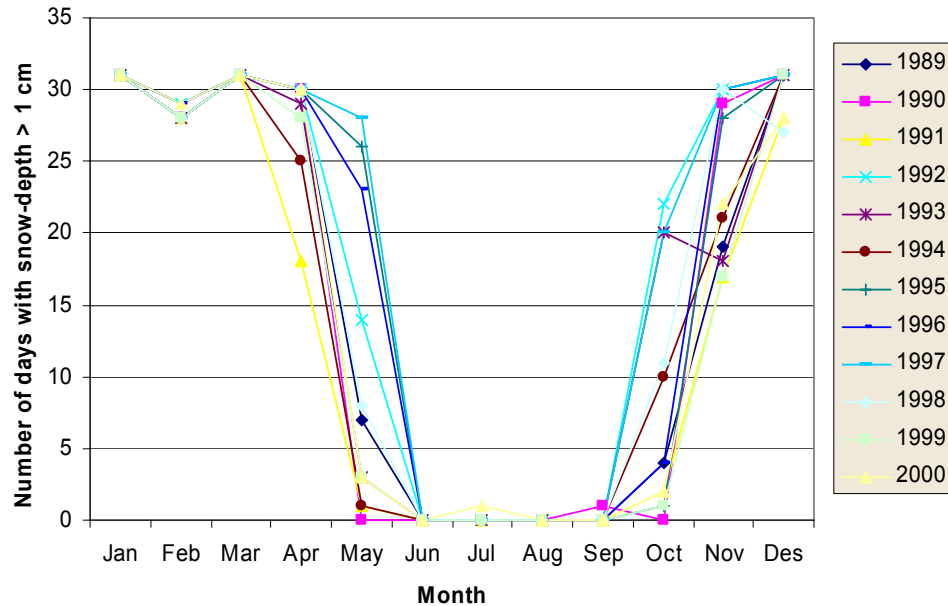
Figure 4-28   Number of days with snow at Kvikne, Hedmark

In Figure 4-27, Figure 4-28, and Figure 4-29 I have tried to illustrate the prevailing variations with respect to snow cover, measured as the number of days per month with snow-depth above one centimetre. The first figure should represent a fairly typical situation for the areas around the Oslofjord (i.e. counties 1,2,3,7,8). Here the typical picture is that there is snow-cover most of the days in December, January and February. However, as can be seen from most of these figures, there are significant changes from year to year.

Moving on to the observations at Kvikne, far north in Hedmark county (northern part of South-East Norway), there is snow-cover most of the time from November to April. Some of my observation sites are located along the coast in South and West Norway. The observation from Kjevik should give a fairly typical picture for most of these sites. In some years there is almost no snow-cover, while in other years there are some days with snow during mid winter.
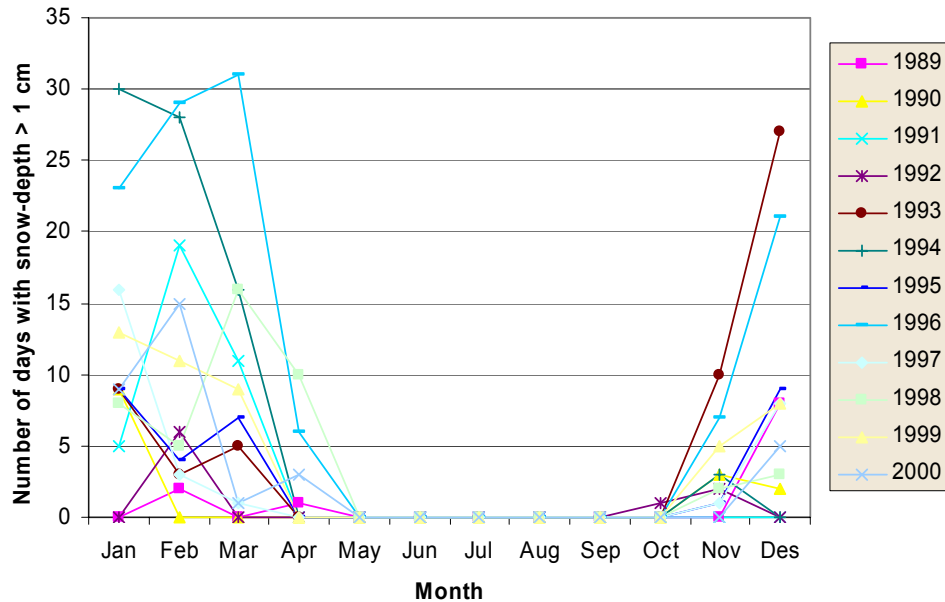
Figure 4-29   Number of days with snow, Site: Kjevik, Vest-Agder

The middle air temperature gives a similar picture to that for snow-cover. From Figure 4-30, Figure 4-31, and Figure 4-32, one can see that the most severe winter temperatures could be found in the inner part of South East Norway (represented by Drevsjø). The climate is somewhat milder in the Oslo area, and the south coast (Kjevik) does not have much frost at all.

Finally, I have chosen four locations for illustrating differences in precipitation levels. Although I have recorded the monthly figures for the whole observation period, the overall picture could more easily be seen from the monthly normals presented in Figure 4-33. The Oslofjord area is represented by Igsi/Hobøl and Oslo which are fairly dry, only beaten by Kjøremsgrende (Lesja, Oppland) representing the rather dry areas in the inner parts of the South East. The south coast (Kjevik) has some more precipitation than the eastern parts, but a really humid climate can be found in the western parts of Norway (Høyanger Verk).

170
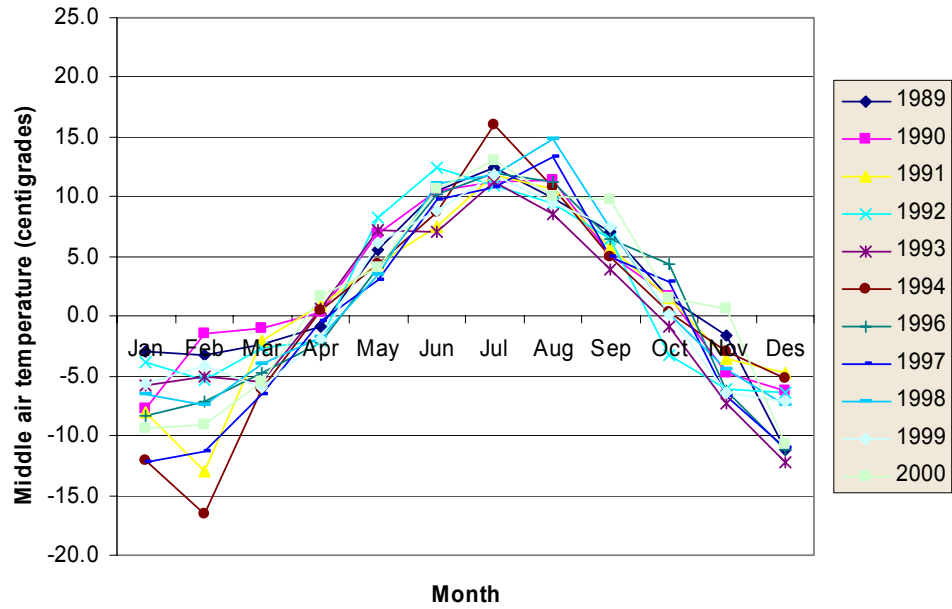
Figure 4-30   Middle air temperature at Drevsjø, Hedmark
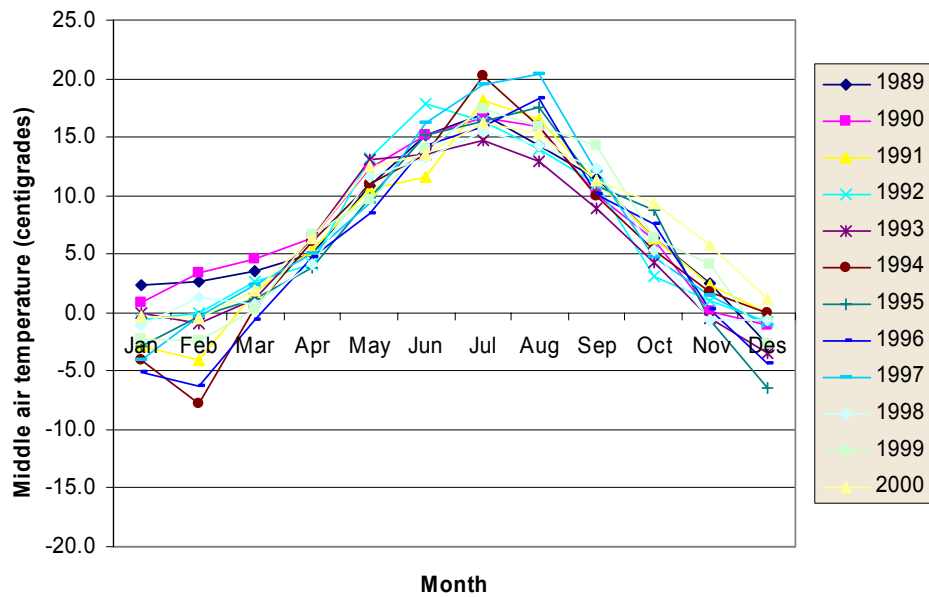
Figure 4-31   Middle air temperature at Blindern, Oslo
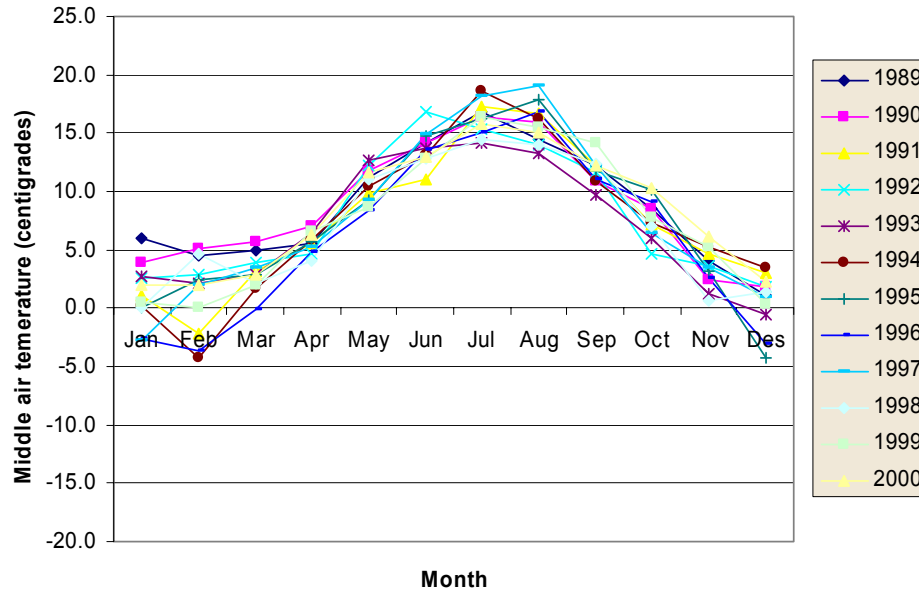
Figure 4-32   Middle air temperature at Kjevik, Vest-Agder

The selection of the climatic data presented above, should give a rough picture of the quite big differences in climatic conditions within the "observation area". This should on principle give a good starting point for identifying some climatic factors affecting the level of road wear. However, when we study the effect of precipitation, snow-cover and temperatures, it is quite clear that very *local* differences may overshadow the regional differences captured in my data set. The ideal situation would therefore be to have direct climatic observations collected along with the traffic data on the observation sites.

172
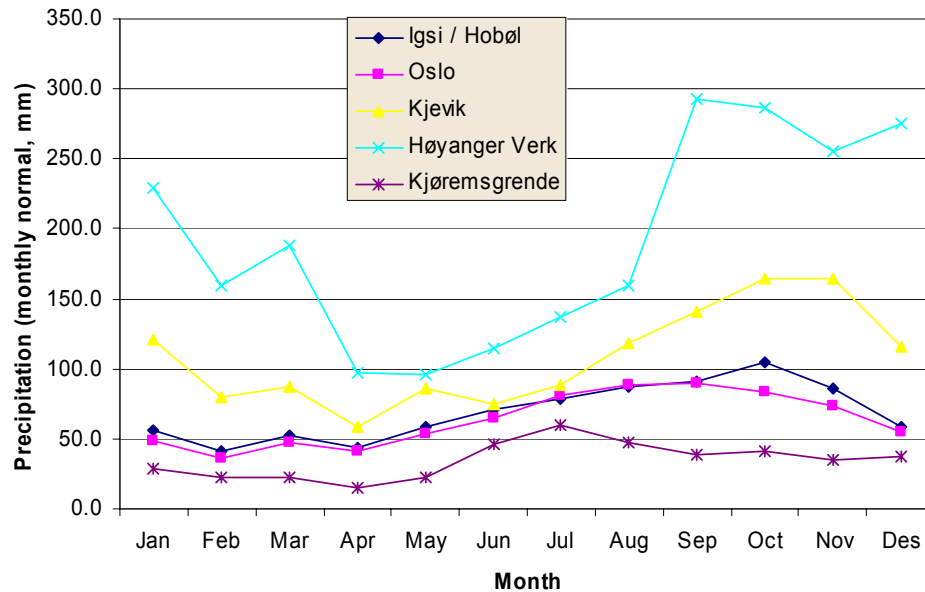
Figure 4-33   Comparison of precipitation levels, monthly normals

4.5   REGRESSION MODELS FOR ROAD WEAR

The primary aim of this analysis is to estimate models for relating observed road wear to traffic levels and climatic indicators. The available data for this analysis are reviewed in the previous sections, and despite the obvious imperfections of the data set, some crude relationships are estimated.

4.5.1   Regression of roughness

I have very briefly discussed some of the factors affecting the development of roughness above. My database gives some control over some important factors apart from measuring the development in roughness (IRI). The information about the traffic loads (axle passages and loads), and about maintenance work is fairly good, and I also have some approximate climate data.

The most prominent shortcoming in my data set is lack of information about the physical structure of the road construction. Functional data like e.g. measurements

173

of deflections[111] etc. are also lacking. This is probably a serious problem when estimating a roughness model. To the extent that the performance of the road construction follows a geographical pattern (by county), or is related to the overall traffic level, I will be able to capture this in my model. The latter phenomenon, i.e. the correlation between traffic levels and the solidness of the road construction, following from the applied road design manuals and procedures, actually represents a genuine identification problem when basing the analysis on in-service roads, rather than test roads. In the model-specification I will try to control for this by differentiating the traffic-parameter by AADT dependent road classes.

I would like to be able to express the marginal contribution of the different factors in terms of a percentage change in IRI-values. This is possible when using a log-linear specification. Such a specification also harmonizes well with a supposition that the development of roughness follows an exponential path. This may be so because dynamic forces (e.g. resulting from HGV "body bounce") typically increase with increasing roughness. The following model is used for the first estimation:

Equation 4-2

$$\ln IRI = \beta_0 + \beta_1 PRECIP + \beta_2 AVGTEMP$$
$$+ \beta_3 DAYS + \left(\beta_4 + \beta_5 D_{5-10} + \beta_6 D_{10-15} + \beta_6 D_{GT15}\right) SUMAXLES$$
$$+ \sum_{i=2}^{18} \beta_{i+5} COUNTY_i + AFTER97 + \varepsilon$$

Where

β_0 - β_{23} are coefficients to be estimated

IRI   the observed average IRI-value for a 200m road section
around the observation site

---

[111]   Some old measurements of deflections exist (see Figure 4-44).

174

PRECIP is total precipitation since the last maintenance action (mm)

AVGTEMP is average monthly temperature at the observation site over the observation period

DAYS is the number of days since the last maintenance action

SUMAXLES is the sum of axles (millions) that have passed the observation site since the last maintenance action

D_{5-10} is a dummy variable for AADT-level between 5 000 and 10 000 vehicles

D_{10-15} is a dummy variable for AADT-level between 10 000 and 15 000 vehicles

D_{GT15} is a dummy variable for AADT-level above 15 000 vehicles

AFTER97 is a dummy variable indication that the observation is made after the new laser-based equipment was installed

$COUNTY_i$ are dummy variables for counties 2,4,5,6,7,9,11,14 and 18

The model is estimated based on 512 observation periods for the estimation of an equation with 17 parameters (not all counties between 2 and 18 are included). Ordinary least squares estimation gives the estimates and t-values presented in the first two columns in Figure 4-34. The estimation of this model gives a multiple regression coefficient, R^2, of 0.692, and a value adjusted for the number of coefficients, R^2-adj, of 0.462.

Since many such models are specified with the number of cumulative equivalent standard axles (ESALs) as the traffic parameter, I also try the same model specification substituting SUMAXLES with ESALS[112]. This yields the estimators and t-values presented in the last two columns in Figure 4-34. This model seems to give a slightly better fit than the previous one, yielding a R^2-adj. of 0.472.[113]

---

[112] Here the traditional AASHO specification of ESALS is used, i.e. based on a 18 kip standard axle and a 4th power equation. NOTE that some (European) studies use a 10 tonne axle as the standard axle.

[113] Models with alternative exponents (2 & 6) for calculating ESALS have also been tested, but these yields a poorer overall fit.

Estimator	Model with SUMAXLES		Model with ESAL4	
	Estimated value	t-value	Estimated value	t-value
Constant	0.908	10.11	0.898	11.77
Precipitation	-3.62E-06	0.37	-1.44E-06	-0.20
Average temperature	3.62E-05	0.88	4.55E-05	1.18
Time (days)	-2.79E-05	1.62	-2.68E-05	1.78
ESAL4	-	-	1.93E-02	4.17
ESAL4*D_{5-10}	-	-	-1.94E-02	-3.43
ESAL4*D_{10-15}	-	-	-1.40E-02	-3.13
ESAL4*D_{GT15}	-	-	-2.53E-02	-5.51
SUMAXLES	1.37E-02	3.56	-	-
SUMAXLES*D_{5-10}	-1.30E-02	3.66	-	-
SUMAXLES *D_{10-15}	-1.17E-02	-3.44	-	-
SUMAXLES *D_{GT15}	-1.67E-02	-4.74	-	-
Dummy county 2	-0.176	-4.86	-0.157	-4.28
Dummy county 4	-0.241	-6.77	-0.232	-6.79
Dummy county 5	-0.287	-6.44	-0.266	-6.00
Dummy county 6	0.211	2.79	0.137	1.91
Dummy county 7	-0.272	-6.09	-0.282	-6.36
Dummy county 9	-0.106	-2.29	-9.49E-02	-2.12
Dummy county 11	-6.08E-02	-1.42	-3.35E-02	-0.80
Dummy county 14	-0.237	-3.08	-0.214	-2.85
Dummy county 18	0.300	2.77	0.330	3.69
After 97	-0.260	-7.90	-0.269	-8.83

Figure 4-34   Estimation results, first IRI-models

The *climatic factors* do not seem to have a significant effect on the IRI-developments in this case. One might suspect that this is due to a strong correlation between the climatic factors and the county dummies, since precipitation and temperature would tend to vary along geographical patterns. However, I have tried the same model specification leaving out the county dummies, and even in this case the estimators for precipitation and temperature do not seem to give a significant contribution to explaining the variation in the IRI-values.

176

The *time component* does not seem significant either, indicating that IRI-values do not grow significantly with the mere aging of the road surface of these roads.

Most of the county dummies come out very significant (apart from counties number 6 and 11). I have left out county number 1 in the model specification, meaning that the rest of the dummies should be interpreted as 'deviations' from county number 1. The possible interpretations of this significant impact of geographical location are several. The most prominent one would be that there are climatic differences that I have not captured in my model specification. There are at least two facts that weaken this hypothesis. Firstly, I have included factors like precipitation and average temperature in my specification. These do not come out significant, and are not the ideal variables for describing the possible impacts of freeze-thaw cycles etc., but I would expect these factors to cater for the coarse, county to county, climatic variation. Secondly, looking at the signs of the estimators of the dummies, one would expect climatic impacts to be more severe in county number 4 and 5 (harsh winter climate) compared to county number 1 (milder climate). The estimators for counties 4 and 5 both have negative signs, indicating that IRI-changes should be smaller in the counties with the more extreme winter temperatures.

Since there are no reasons to expect a systematic variation in the solidness of the roads with respect to county, my remaining suspicion would be that the significance of the county dummies actually represents a systematic measurement error related to the fact that each county has its own measurement vehicle. Any differences in the way each vehicle registers IRI-values would therefore be mirrored in my county dummies. The fact that the vans have been annually calibrated against each other should reduce this potential problem, but may not have eliminated it.

Dropping the factors that did not achieve a t-value above 2, leaves us with a

simpler model specification (Equation 4-3) leaving out time, the climatic variables and some of the county dummies. The estimates of the simplified model are given in Figure 4-35.

Equation 4-3

$$\ln IRI = \beta_0 + \left(\beta_4 + \beta_5 D_{5-10} + \beta_6 D_{10-15} + \beta_6 D_{GT15}\right) ESAL4$$
$$+ \sum_{i=2}^{18} \beta_{i+5} COUNTY_i + AFTER97 + \varepsilon$$

The adjusted R^2 for this model is 0.470, which is only slightly lower than the first model. Comparing the two model specifications shows that none of the estimators change sign or magnitude significantly, and thus prove relatively stable to model specification.

A closer look at the estimators for the standard axles shows that the most significant impact is made on low volume (i.e. low standard) roads with an AADT less than 5000 vehicles. The estimators for the other road classes should be deducted from the general ESAL4 estimator. All of these estimators have a negative sign, indicating that IRI development on high volume roads is less significant than on the low volume roads. In fact the "net" estimator for roads above 5000 AADT seems to hover around zero, indicating a very small, if any, traffic related impact on IRI-development in this sub-sample.

The dummy related to the shift in the measurement technology in 1997 also turns out very significant, indicating that IRI-measurements done with the new laser-based equipment generally produced lower IRI values than the old ultrasound devices did.

178

Estimator	Estimated value	t-value
Constant	0.769	41.15
ESAL4	0.021	4.96
ESAL4*D$_{5\text{-}10}$	-0.020	-3.55
ESAL4*D$_{10\text{-}15}$	-0.015	-3.59
ESAL4*D$_{GT15}$	-0.026	-5.98
Dummy county 2	-0.161	-4.58
Dummy county 4	-0.255	-9.39
Dummy county 5	-0.271	-6.41
Dummy county 7	-0.280	-6.49
Dummy county 9	-0.097	-2.29
Dummy county 14	-0.224	-3.02
Dummy county 18	0.297	4.21
After 97	-0.297	-14.29

Figure 4-35   Estimation results, simplified IRI-model

The magnitude of the ESAL4-parameter for roads with AADT below 5000 is also very modest, especially when compared to the significant variations represented by the county dummies. E.g. 1 million ESALs will only contribute 2 percent to IRI deterioration, whereas the difference between counties 1, 6, 11, and county 2 represents an increase in IRI measurements of around 15 percent. Applying the estimators representing a road with AADT less than 5000, and counties 1, 6 and 11 yields a predicted pavement life of 34 million ESALS. The sample mean daily traffic, measured in ESALs, is 2 121. This means that the predicted pavement life of 34 million ESALs represents a pavement life of 44 years evaluated at the mean traffic level and with a maintenance triggering IRI-value of 4.5. Predictions representing the other counties yield equivalent pavement lives of between 26 and 61 years. The Norwegian handbook for road maintenance (NPRA 1998) contains different maintenance trigger values for roughness, depending on road class. Trigger values range from 5.6 mm/m for the low volume secondary roads, to 3.6 mm/m for high volume (above 5000 AADT) arterial roads. The estimated pavement life for such a high volume arterial road, with 2 121 ESALs/day (sample mean), is then 30 years for counties 1, 6 and 11. *Even for such a high*

volume road, this means that pavement rehabilitation is likely to be triggered by other mechanisms than traffic-related roughness development, according to my model.

The very significant county-dummies may also be a reflection of poor consistency in the measurement of IRI-values, making the reliability of the model questionable. Based on this assessment, I will not take the results of the IRI-model further into the calculation of marginal road wear costs.

4.5.2   Regression of rut-depth

As for the roughness model presented above, I have very briefly discussed some of the factors affecting the development of rutting. My database includes measurements of mean rut-depths on my observation sites, and also quite good information about the traffic loads, maintenance work, and some climatic data (precipitation, temperatures and snow-depth).

The most important shortcoming in my data, with respect to factors affecting rut-depth development on Norwegian roads, is that I do not have direct control over the use of studded tyres on bare roads. My database also still lacks information about the physical structure of the road construction.

Once again, I would like to be able to express the marginal contribution of the different factors in terms of a percentage change in rut depth values, and therefore I choose a log-linear model specification. This model form also seems plausible since most damage mechanisms for roads typically follow a exponential development pattern, rather than a linear one.

The estimation is done subject to the following model:

180

Equation 4-4

$$\ln RUTDEPTH = \beta_0 + \beta_1 SUMAXLES + \beta_2 AVGTEMP + \beta_3 PRECIP + \sum_{i=2}^{14} \beta_{i+2} COUNTY_i + \beta_{12} AFTER99 + \varepsilon$$

Where

β_0 - β_{14}	are coefficients to be estimated
RUTDEPTH	is the observed average rut-depth (mm) for a 200m road section around the observation site
SUMAXLES	represents the accumulated sum of axles that has passed the observation site in the observation period
AVGTEMP	is average monthly temperature at the observation site in the observation period (degrees Celcius)
PRECIP	is the calculated accumulated amount of precipitation, based on monthly figures, in the observation period (mm)
COUNTY$_i$	are dummy variables for counties 2,4,5,6,7,9,11,and 14
AFTER99	is a dummy variable indicating that the measurement was done under the adjusted procedure adopted at January 1999.

I have also tested models substituting the traffic variable SUMAXLES with figures representing the number of standard axles passed, with different exponents (from 2 to 6). A rising exponent seems to give slightly poorer model fits, indicating that there is no rationale for treating heavy axles different to the lighter axles. This may be a reflection of the fact that my sample is mainly based on solid roads, indicating that the only significant impact of traffic seems to be related to the use of studded tyres in the winter season. I have therefore ended up with this formulation where the only traffic-related item is the number of axles (regardless of axle weight)[114]. Similar to the IRI-model presented above, I have also tested a

---

[114] Since it is mainly cars and light vans that use studded tyres, a model specification with car-axles as the sole traffic variable was also tested. This yielded a slightly lower adjusted R^2, and an axle-parameter 28 percent higher than the one estimated for SUMAXLES below.

model specification with a SUMAXLE-parameter differentiated by overall traffic level, possibly reflecting different road strengths. This model does not yield significant traffic parameters, and was also dropped.

Estimator	Estimated value	t-value
Constant	1.880	21.54
Sum axles	1.08E-02	5.54
Precipitation	9.28E-05	5.19
Average temperature	-3.30E-02	-4.52
Dummy county 2	-9.53E-03	-0.16
Dummy county 4	-3.31E-02	-0.42
Dummy county 5	-0.102	-1.12
Dummy county 6	6.496E-02	0.56
Dummy county 7	0.360	5.28
Dummy county 9	0.274	4.51
Dummy county 11	2.43E-02	0.37
Dummy county 14	-0.103	-0.82
After 1999	0.300	3.89

Figure 4-36   Estimation results, first rut-depth model

The model was estimated on 442 observations to estimate 12 coefficients. The overall model fit is illustrated by a R^2 of 0.755, and an adjusted R^2 of 0.559. The point estimates and the corresponding t-values are presented in Figure 4-36.

The estimator for the accumulated number of axles passing, comes out significant and with a positive value, indicating that a positive correlation between accumulated traffic and rut-depth development.

The estimator for average temperature is significant and has a negative relationship to the development in rut-depth. This is coincident with the expectations, since a milder climate would reduce the problems related to freeze-thaw cycles, and also reduce the effective season for the use of studded tyres.

Precipitation levels also come out significant, with a positive value. Higher levels

182

of precipitation therefore seem to increase rut-depth development. This may be due to a higher deterioration rate for the pavement under wet conditions, or it may also be attributed to the severity of freeze-thaw problems, which typically will be more severe in moist environments (for given temperatures).

The change in measurement procedure for ruts in 1999 seems to alter the rut-depth loggings significantly, adding $e^{0.3} = 1.35$ to the previous measurement level on average.

Once again I have included dummies for the counties, and as in the roughness model I have left out county number 1, making the other dummies represent "deviations" from this county. Only counties number 7 and 9 seem to be significantly different from county 1 here, and I leave out the other dummies in my final model presented in Figure 4-37. This model yields an adjusted R^2 of 0,562.

Estimator	Estimated value	t-value
Constant	1.806	35.08
Sum axles	1.14E-02	7.50
Precipitation	9.78E-05	7.32
Average temperature	-2.72E-02	-5.20
Dummy county 7	0.363	5.60
Dummy county 9	0.282	5.30
After 1999	0.316	4.22

Figure 4-37   Estimation results, adjusted rut-depth model

As with the roughness model, the underlying mechanisms behind the significant county-dummies for county number 7 and 9 compared to the rest of the counties may be attributable to systematic measurement errors (vehicles).

4.5.3   *Using the rut-depth model for calculating average road wear costs*
In the following section I am presenting a calculation model for marginal tax

relevant external costs. One of the elements here will be the marginal road wear costs. The Newbery theorem presented in section 2.4, shows that marginal road wear can be represented by average maintenance costs under fairly general and plausible conditions. Assuming that road wear in the form of rutting is the triggering damage mechanisms on Norwegian roads[115], we could apply the model presented above to predict the traffic levels (number of accumulated axles) necessary to reach a maintenance triggering level of rutting, and by combining this figure with typical pavement overlay costs, we have the information necessary to calculate average maintenance costs related to road wear.

The trigger-level for rutting on Norwegian national roads is 25 mm (NPRA 1998). This level was earlier differentiated by speed limit, and has been considered made differentiated by traffic volume as well, but currently there is no differentiation for national roads. However, maintenance actions may be taken at lower rutting levels when other performance indicators[116] have reached their trigger level. This is mainly relevant for low volume roads. The trigger level is also "advisory" and may be adjusted from year to year depending on available funds for maintenance.

In Figure 4-40 predictions of rut-depth are made for alternative climates (assuming that we are not in counties 7 or 9). The maintenance trigger level for rut-depth is also illustrated. By studying the intersection of the curves with this trigger level it is possible to find corresponding pavement lives expressed as the number of cumulative axles the pavement has been exposed to. This could also be done analytically by solving the equation for the critical rut-depth of 25 mm.

The estimated equation for rutting with this trigger level yields:

---

[115] Given the very extensive use of studded tyres on the Norwegian road network, this is not an unlikely assumption.

[116] E.g. cracking, drainage, curvature, longitudinal roughness.

Equation 4-5

$$25 = e^u \cdot e^{\frac{\sigma^2}{2}}$$

$$u = 1.806 + 1.14E - 02 \cdot SUMAXLES - 2.72E - 02 \cdot AVGTEMP + 9.78E$$
$$- 05 PRECIP + 0.363 \cdot County\,7 + 0.282 \cdot County\,9 + 0.316\,AFTER\,99$$

The last term is a "correction" factor necessary because the log-linear specification makes the error terms log-normally distributed. σ^2 is then the variance of the estimate. For my purpose I solve this equation with respect to the SUMAXLES:

Equation 4-6

$$SUMAXLES = \frac{\ln\left[\dfrac{25}{e^v \cdot e^{\frac{\sigma^2}{2}}}\right]}{1.14E - 02}$$

$$v = 1.806 - 2.72E - 02 \cdot AVGTEMP + 9.78E - 05 PRECIP$$
$$+ 0.363 \cdot County\,7 + 0.282 \cdot County\,9 + 0.316\,AFTER\,99$$

For illustration I use the sample means for average temperature and precipitation, and assume that we are not in counties 7 or 9 (i.e. dummies equal 0). The sample mean temperature is 7.50 degrees, and the mean accumulated precipitation level is 2215 mm. The standard deviation of the estimate for this model is 0.3452. Inserting these figures into the equation, yields a point prediction for the accumulated sum of axles necessary to reach the triggering rut depth level of 25 mm of 90 million. This figure varies somewhat with alternative climatic factors. This is illustrated in Figure 4-38 where the "cold & wet" alternative brings the life of the pavement down to 57 mill. axles, whereas a "warm & dry" climate yields a prediction of a life of 111 mill. axles. Evaluated at the sample mean 11 769

AADT[117], these figures implicate estimated pavement lives between 13 and 26 years, respectively.

Environment:	Sample means	Cold	Warm	Dry	Wet	Cold & Dry	Cold & Wet	Warm & Dry	Warm & Wet
Avg. temp (C):	7.50	0.00	12.00	7.50	7.50	0.00	0.00	12.00	12.00
Accumulated precipitation (mm):	2215	2215	2215	1000	4000	1000	4000	1000	4000
Predicted pavement life (mill. axles):	90	71	100	100	75	82	57	111	85
Predicted pavement life (years at sample mean AADT):	20.9	16.6	23.2	23.3	17.3	19.2	13.2	25.8	19.8

Figure 4-38   Predicted pavement lives with alternative figures for the climate (25 mm rut-depth trigger level)

Applying an alternative trigger-level of 15 mm rut-depth

The maintenance triggering rut-depth level of 25 mm may not represent the *average* rut-depth for a road section when maintenance actions are taken. The "rule" is that maintenance should be performed whenever 10% of the road section has a rut-depth of more than 25 mm. To illustrate how an alternative trigger-level of 15 mm would affect pavement lives, in Figure 4-39 I have made similar calculations to the ones presented in Figure 4-38 with the alternative rut-depth value. Comparing the two tables, we can see that such a reduction in trigger-level yields more than a halving of the estimated pavement lives, now ranging from 2.8 years (Cold & Wet) to 15.4 years (Warm & Dry). Applying the sample mean values for precipitation and temperature yields an estimated pavement life of 10.5 years with this 15 mm trigger level. In Senstad (1994) the average predicted pavement life of such a high volume road on the Norwegian network of national roads would be around 6 years. So it seems that the pavement lives estimated with a 15 mm trigger level better coincides with the findings in the BUAB project.

---

[117]   Assuming 2,2 axles per vehicle.

186

Environment:	Sample means	Cold	Warm	Dry	Wet	Cold & Dry	Cold & Wet	Warm & Dry	Warm & Wet
Avg. temp (C):	7.50	0.00	12.00	7.50	7.50	0.00	0.00	12.00	12.00
Accumulated precipitation (mm):	2215	2215	2215	1000	4000	1000	4000	1000	4000
Predicted pavement life (mill. axles):	45	27	56	55	30	38	12	66	40
Predicted pavement life (years at sample mean AADT):	10.5	6.3	13.0	12.9	6.9	8.8	2.8	15.4	9.4

Figure 4-39   Predicted pavement lives with 15 mm rut-depth trigger level
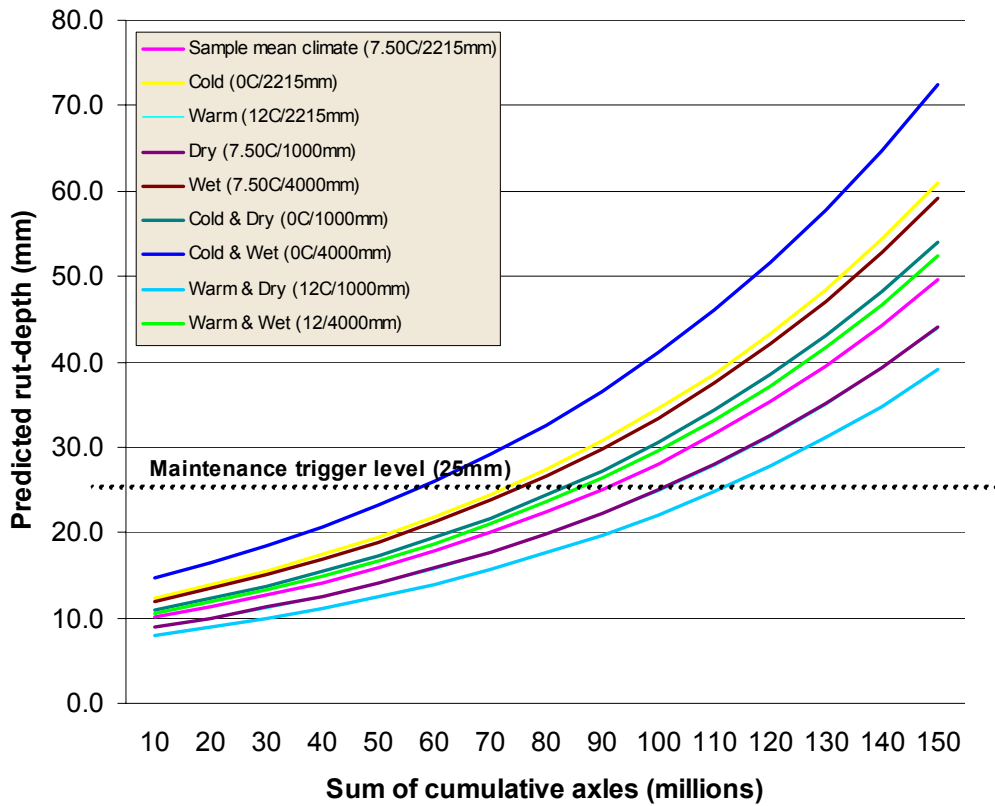
Figure 4-40   Predicted rut-depth for alternative climates

Traffic level (AADT)	Minimum overlay costs, 2001 NOK per meter	Maximum overlay costs, 2001 NOK per meter	Overlay costs with deep ruts, medium IRI, 2001 NOK per meter	Applied overlay costs, 2001 NOK per meter	Applied overlay costs 2002 € per km
Arterial road					
>10000	580	805	765	495	61 959
5001-10000	355	550	495	495	61 959
1501-5000	350	440	390	390	48 816
0-1500	270	360	320	320	40 054
Other national roads					
>10000	465	665	625	460	57 578
5001-10000	330	500	460	460	57 578
1501-5000	320	415	360	360	45 061
301-1500	250	330	290	290	36 299
0-300	200	290	240	240	30 041

Figure 4-41   Calculation of overlay costs, Part 1

Traffic level (AADT)	Applied overlay costs (2002 € per km)	Length of road subnetwork (km)	Subnetwork's share of total network	Subnetwork's average traffic level (AADT)	Network share * average traffic	Subnetwork's share of total traffic	Traffic share * Applied overlay cost (2002 € per km)	Subnetwork's share * Applied overlay cost (2002 € per km)
Arterial roads								
>10000	61 959	711	2.7 %	20695	556.17	23.1 %	14 296	1 665
5001-10000	61 959	899	3.4 %	7260	246.70	10.2 %	6 341	2 105
1501-5000	48 816	2508	9.5 %	2659	252.07	10.5 %	5 105	4 628
0-1500	40 054	3271	12.4 %	893	110.41	4.6 %	1 835	4 952
Other national roads								
>10000	57 578	494	1.9 %	16522	308.51	12.8 %	7 369	1 075
5001-10000	57 578	849	3.2 %	7042	225.98	9.4 %	5 398	1 848
1501-5000	45 061	3994	15.1 %	2541.5	383.68	15.9 %	7 173	6 803
301-1500	36 299	10721	40.5 %	749.5	303.73	12.6 %	4 574	14 710
0-300	30 041	3009	11.4 %	203.5	23.15	1.0 %	288	3 417
Sum		**26456**	100.0 %		**2410.40**	100.0 %	**52 380**	**41 202**

Figure 4-42   Calculation of overlay costs, Part 2

Typical overlay costs for the Norwegian road network

The next element needed for predicting maintenance costs per axle kilometre, is information on typical overlay costs on the Norwegian network. There are no "official figures" for such costs available, but we could develop approximate figures based on a recent SINTEF report (Sund 2001). In Figure 4-41 the cost estimates from this publication is reproduced in the first three columns. The report contains alternative figures for overlays dependent on the current state of the road (average rut-depth and roughness), and for different road classes (mainly based on traffic levels). Columns one and two represent the lower and upper boundary of costs based on different levels of existing rut-depth and roughness prior to the resurfacing. In my context, I want to focus on a responsive maintenance strategy where the trigger level is that rutting has exceeded 25 mm. This means that the relevant category would be cost estimates in the most "severe" category[118] with respect to rutting. I also choose to apply the medium category[119] with respect to roughness.

Because there is a significant proportion of the upper road class that has more than 2 lanes, I choose to use the overlay costs of the 5-10000 AADT class for this class as well. The figures are originally reported in 2001 NOK per meter, and this is converted into 2002 Euros per km. Comparing these figures to the ones reported for Sweden in (Lindberg 2002b), it turns out that the Norwegian figures are in the approximately 2 times the Swedish corresponding figures. The range of the Swedish figures is from k€ 14 per km for the narrow roads in the south, to k€ 71 per km for wide roads in the north. The upper boundary is relevant for multilane roads with an average road width of 18 meters. Swedish roads for AADT volumes between 2000 and 8000 have an estimated maximum overlay cost of k€ 35 per km. The reported average overlay cost in Sweden is k€ 25 per km.

---

[118]   i.e. rut-depth > 15 mm.

[119]   This category represents IRI-values between 1.5 and 5.0 depending on road class (lower values for high volume roads).

189

In Figure 4-42 the cost estimates for Norway are processed further towards average overlay costs for the whole network of national roads. Two alternative averages are produced; the first with weights based on the individual road class' share of the total traffic, the second with weights based on the road class' share of the overall length of the network. The traffic-based average yields an average overlay cost of € 52 380 per km, the distance-based average is € 41 202 per km. The latter turns out with a lower value because the low-traffic, narrow roads (with low overlay costs) are assigned a higher weight when calculating the average this way.

In my setting such averages have only limited interest, because the basic idea behind marginal cost pricing is *differentiation*, and not averaging. This means that one should make the charges reflect realistic differences in the costs related to road use. However, my rutting-model does *not* discriminate between road classes, and it may very well be that rutting development is different for different classes[120]. Applying the different cost estimates for the road classes when calculating average overlay costs is therefore not an ideal solution as this probably might *overstate* the costs related to using the high volume road network.

Using the point prediction based on the climatic sample means, for the number of axle loads necessary to reach a rut-depth of 25 mm, I am now able to calculate some estimates of the average overlay costs per axle-km (Figure 4-43).

Road class	Average cost
Traffic-based average overlay cost:	0.00051 € per axle-km
Distance-based average overlay cost:	0.00040 € per axle-km
Cost arterial road >5000 AADT	0.00060 € per axle-km
Cost other national road 1501-5000 AADT:	0.00044 € per axle-km

Figure 4-43   Estimated average cost per axle-km based on sample means for the climatic factors

---

[120]   E.g. one might expect road wear to be more spread out transversally when road width increases, and one would also typically use more solid pavement types on a high volume road.

190

Doing the same calculations with the most extreme climate figures above, the axle-km cost for the traffic-based overlay average yields € 0.00080 for the "Cold & Wet" climate, and € 0.00041 for the "Warm & Dry". (Note that if we alternatively applied the pavement lives for a 15 mm rut-depth trigger value, given in Figure 4-39, the costs would be two to three times higher than the ones given here).

In Section 5 I will use these figures when calculating the scenarios for tax relevant external costs. To finish off this section, I try to indicate some further research areas related to the FAMAROW-model.

4.6 FURTHER RESEARCH ISSUES RELATED TO THE FAMAROW-MODEL

I have in the previous section presented a first effort to estimate models for the development of roughness and rut-depth on Norwegian in-service roads. New WIM-data, and systematic registrations of road wear entered into the Norwegian Road Data Bank have enabled this new approach to estimating models for road wear.

The most controversial result from my estimation of the rut-depth model is the fact that a "flat" axle variable yielded a better model fit than model formulations including alternative standard axle variables. Exponents from 2 to 6 was tested, and they all yielded poorer fits than the SUMAXLES variable. This result is in direct contradiction to most previous theoretical and empirical work on the subject (e.g. the AASHO Road Test and many other similar results). However, there are many possible explanations for this. My main hypothesis would be that any possible effect on rutting caused by heavy axles might be swamped by the effect of studded tyres. The main rutting effect one would expect from heavy axles is a densification of the subgrade. However, just measuring the resultant rut-depth on the top of the pavement does not allow a separation of this effect from the abrasion effect of studded tyres on the bound asphaltic layer. An important

weakness is also the fact that we do not have direct control over the actual use of studded tyres in our database.

Alternative explanations for the controversial result are:

- There may be too little variation in the traffic mix to pick up the effect of heavier axles. Still, there *is* some variation between the observation sites and periods. This is illustrated by the differences in the shares of passenger cars given in Figure 4-11.

- There may be a correlation between numbers of heavy vehicles and road strength stemming from the prevailing criteria for road design which includes modifications in the recommended road structures when the share of heavy vehicles increases. The lack of data describing the road structure, and hence the strength of the road prevents a further analysis of the magnitude of this problem.

It is evident then, that the available database is far from flawless for our purpose. *Firstly*, there are numerous sources of measurement errors in the variables. It should be possible to control for some of these (e.g. any systematic differences between the test vehicles), and some are more or less inherent to this kind of data (e.g. the link between factual traffic levels and the strength of the road construction and maintenance levels). *Secondly*, I lack important pieces of information in order to control the major damage mechanisms that one may expect to be present. As pointed out earlier, the major shortcomings in this respect would be the lack of information about the physical characteristics of the road constructions, and also reliable information about the factual use of studded tyres on bare roads. *Thirdly*, my observation sites are not chosen subject to any random sampling procedure. The instrumented sites are deliberately chosen to represent road sections with high traffic levels, and with corresponding high accident rates. This means that my sample would generally *not* be representative of the entire

network of national roads in Norway.

Most NPRA county administrations ended their systematic registrations of bearing capacity many years ago. This means that there is no up-to-date registry of deflection measurements available for the Norwegian road network. However, I have extracted the last calculations of actual (not regulated) bearing capacities registered in the Norwegian Road Data Bank (Figure 4-44).

County	Calculated average bearing capacity of county road network of national roads (Axle load, tons)	Period of deflection measurements	Calculated average bearing capacity of FAMAROW road sections (Axle load, tons)	Period of deflection measurements
01 Østfold	16.3	1976-1999	18.2	1976-1992
02 Akershus	14.9	1976-1992	16.9	1976-1992
04 Hedmark	13.9	1976-1997	17.7	1989-1996
05 Oppland	14.4	1976-1998	15.8	1989-1997
06 Buskerud	15.2	1976-1996	6.2	1983-1984
07 Vestfold	14.7	1976-1996	15.4	1987
09 Aust-Agder	15.3	1976-1996	18.0	1989
11 Rogaland	15.3	1976-1997	15.8	1985-1989
14 Sogn og Fjordane	14.3	1976-1992	12.0	1986-1992
18 Nordland	15.1	1976-2002	14.1	1995
Average (unweigted)	*14.9*	*1976-2002*	*15.0*	*1976-1997*

Figure 4-44   Calculated bearing capacity based on deflection measurements (Source: The Norwegian Road Data Bank)

Some of the measurements done dates back to the mid 1970s, but some NPRA county administrations are still testing the bearing capacity of roads. The mean of

the average county bearing capacities (not weighted by the extension of the networks) is 14.9 tons. The corresponding figure for my sample of road sections is 15.0 tons. Judging from these averages, my sample seems to be quite representative of the average bearing capacity of the network of national roads. However, studying this at the county level, the differences are more significant. The largest difference between sample average and network average is for county number 6 (Buskerud). However, the deflection measurements of the sample sections are very old (20 years), and may very well not be representative of the prevailing road conditions. The number of observations from this county is also very limited.

A study of the effective average maintenance-triggering level of the different damage mechanisms would also be worthwhile. I have based my main calculations on the official trigger levels applied whenever a specified proportion of the network exceeds this value. Information about the *average* values when maintenance is performed is not currently available but would be most useful for making realistic calculations of marginal costs.

Despite the obvious shortcomings in my data-sources, I have been able to estimate some crude models that give plausible results that coincide with our expectations with respect to the interactions between the variables concerned. This encourages further research in order to improve the possibility of identifying the factors that determine road wear under different conditions. Such information would be valuable for investment and maintenance policies, as well as for developing pricing policies for different vehicle types.

From this short review of the shortcomings in my database, the agenda for further research seems rather obvious. The major possibilities for improvement would be along the following dimensions:

194

- Better quality assurance of the WIM-data and the data entered into the Norwegian Road Data Bank

- Collection of new information regarding the use of studded tyres, road construction structure and climatic factors at the observation sites

- Collection of WIM-data on low-volume roads and on more roads in West-, Mid- and North Norway

- A study of *average* values of road status indicators (e.g. rut-depth) when maintenance actions are taken. Also a study of which damage mechanisms actually trigger maintenance for different road classes would be worthwhile.

Further, a comparison of the results from this empirical approach and the results of mechanistic models and "laboratory" road tests would be interesting.

My database (comprising pooled cross-section and time-series data) may be suitable for *panel data estimation techniques*. However, the exploration of such methods has been beyond the scope of this study. This means that my analysis has been conducted under the assumption that the effects studied are generic to all units (i.e. observation sites, time periods) in my data set. If this is a plausible assumption, then my approach is the more efficient one. However, if this assumption could not be supported, it means that I may have derived biased estimators. A full exploration of the possible heterogeneity of the observation units would mean a severe loss of degrees of freedom considering the limited number of observations available. However, with a larger data set, exploring these possibilities would be an additional natural development of the analysis, and clearly another area of further research[121].

---

[121]   The fact that I have included some county-specific constant terms (dummies) brings me somewhat closer to exploring the possible heterogeneity between the observation units.

195

5 CALCULATION OF TAX-RELEVANT EXTERNAL EFFECTS

5.1 CATERU – CALCULATING TAX RELEVANT EXTERNAL COSTS OF ROAD USE

Establishing the cost foundation for implementing road user taxes is a complex and demanding task, requiring a stringent and systematic approach, combined with a careful application of the proper economic principles. Recently, comprehensive research programmes (e.g. the U.S. Fedral Highway Cost Allocation Study and in the EU: The UNITE project) have been launched to establish such a cost foundation for many countries. The approaches are quite different, and it is therefore of great interest to compare their outcomes.

I have, in my previous work as a research officer at the Institute of Transport Economics (TØI) in Oslo, gained experience with one particular approach to such calculations. This is a model developed over 25-30 years at TØI, and the last application and development of the model is presented in (Eriksen et al. 1999) (hereafter referred to as the TØI-model). Based on this experience, I have developed a spreadsheet model that inherits the basic structure from the TØI-model, but has been further developed with the primary aim of developing a tool for analysing the response to different input in such calculations. The spreadsheet model is called CATERU[122], and it enables the testing of sensitivities resulting from

- different principal approaches to cost estimation
- alternative estimates of shadow prices

---

[122] CATERU: Model for the Calculation of Tax relevant Road User costs.

Compared to the TØI-model (*op.cit.*), I have also chosen to use different estimates of cost components where these are considered to be more representative of the current Norwegian road network and vehicle park, or based on more recent or comprehensive studies.

An entirely new element is also added to include a marginal element of the vehicle purchase taxes in the calculation of tax to cost ratios (see the rationale for this established in section 5.2.4). The results of the last TØI-model update (op.cit.) are compared to the Base Scenario from the CATERU model in section 5.4.2.

In the following sections the *building blocks* and the contents of the model are presented according to the following structure:

- A short description of the *contents and structure* of the main building blocks of the CATERU model
- A *comprehensive review* of how the calculation of the various cost elements (i.e. global and local emissions, time costs, accident costs and road wear) is done, partly reviewing the foundation of the TØI-model, and partly introducing new – or revised – elements.

Following this description of the model is a section where the *model output* is presented. This part contains:

- A comparison of the Base scenario to the output from the TØI model
- A presentation of alternative CATERU scenarios, where some elements of the Base scenario has been substituted by combinations of alternative inputs
- A sensitivity analysis of the Base scenario where the impacts of alternative values of single input parameters are studied
- A comparison of the CATERU output with recent international studies

198

5.2   THE BUILDING BLOCKS OF THE MODEL

The model is established in an Excel Workbook with 6 spreadsheets, illustrated by the modules in Figure 5-1.
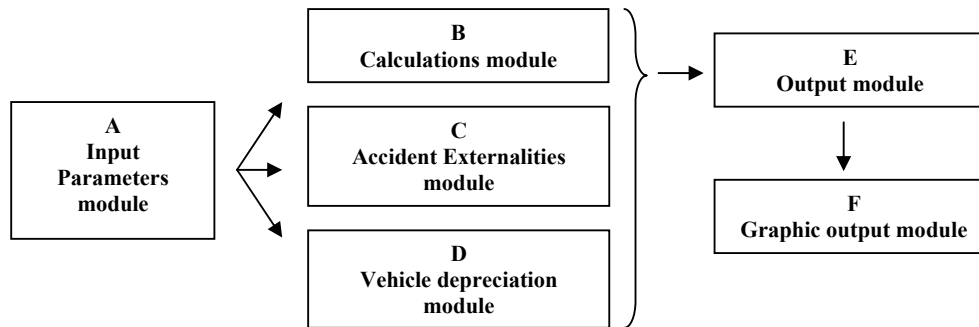
Figure 5-1   Illustration of the building blocks of the CATERU-model

The Base Scenario calculation of external road user costs and taxes can be found in the Appendix. The main elements of each module are described in the subsequent sections.

5.2.1   A: The Input Parameters module

As far as possible, basic input parameters are collected in the Input Parameters module. This is done to enhance the general transparency of the model, but also to ease the process of carrying out sensitivity analysis by altering the model input parameters in this module only. The rest of the model structure is dynamically linked to these input parameters, and therefore relates automatically to changes made in Module A.

Examples of parameters in this module are:

- Value of time
- Value of accidents
- Values of environmental impacts

- Road network maintenance and operation costs
- Fuel consumption figures
- Current taxation rates
- Currency conversion factors
- Time conversion factors (price indexes)

5.2.2 B: The Statistics and Calculations module

Intermediate calculations based on statistical information and the input parameters from Module A are collected in Module B: The Statistics and Calculations Module. The statistical information in this module is collected from many different sources, mainly Norwegian official statistics and Norwegian research reports, the latter often relying on international research results that have been evaluated and adopted to fit Norwegian conditions. The more detailed source references are left to section 5.3.

Elements contained in Module B are:

- Traffic volume statistics by vehicle type on the Norwegian road network
- Specific emission factors by vehicle type
- Estimated road use related noise figures
- Road accident statistics
- ESAL-km calculations for road wear assessment based on AASHO-factors
- Time cost vehicle equivalency units

5.2.3 C: The Accident Externalities module

The Accident Externalities module is mainly based on the model presented in Elvik (1994). This is a comprehensive model based on a very thorough review of international research regarding the economic valuation of reduced risk of traffic accidents. The whole project is documented in Elvik (1993).

200

The accident frequency statistics within the model have to some extent been replaced by more recent statistical figures. The shadow prices for a statistical life and figures derived from these are updated to current official Norwegian figures.

Some of the tables in this module represent necessary intermediate calculations that are not directly documented in Elvik (1994), and the last part of the module represents an extension of Elvik's model, necessary to apply the results of Elvik's work in my marginal external cost setting.

A more detailed description of the accident externality calculations is given in section 5.3.4 below.

5.2.4   D: The Vehicle Depreciation module

Generally, when estimated external costs related to road use are compared to the current level of taxation, the focus is put on taxes and charges that are directly linked to fuel consumption or kilometres driven. In Norway, more than in many other countries, there is a substantial level of taxation put on the registration of vehicles. Currently this taxation varies from 35 to 60 percent of the c.i.f. import price. This kind of taxation is normally not considered relevant when studying *marginal* costs and taxation, because these taxes are not directly linked to the *use* of the vehicles. However, one might expect the depreciation costs of owning a vehicle to be partly dependent on the number of kilometres driven as well as time. To concretise, the second hand price of a car will not only depend on the age of the car, but also the kilometrage (among other factors).

To illustrate this point, I have conducted a simple regression analysis of a small part of the Norwegian second hand car market, in order to investigate this 'hypothesis' empirically. The analysis is not complete, as it simply covers a few typical car makes in each car category. Nor is it very advanced with respect to

controlling for all the factors that may explain differences in second hand prices. My aim here is merely to give an illustration of how this could be entered into the tax-to-cost ratio picture.

I have collected advertised prices for 13 different car and van makes in a database[123] containing second hand cars located at brand-dealers all over the country. All in all, I found more than 3000 vehicles of these models. In addition to price, the database also contained information about kilometrage, model year, engine size, fuel type, car body type, and four wheel drive (among other factors).

I estimated a simple log-linear model with the natural logarithm of the advertised price as the dependent variable, and age and kilometrage as independent variables. For the models where this was relevant, I also included dummies for estate car body, diesel engine and four wheel drive. The regression equation applied is given in Equation 5-1.

Equation 5-1 $\ln(p) = \beta_0 + \beta_1 \cdot AGE + \beta_2 \cdot KMS + \beta_3 \cdot EST + \beta_4 \cdot DIESEL + \beta_5 \cdot FWD$

Here

p	is the advertised second-hand price of the car (NOK)
AGE	is the age of the vehicle (years)
KMS	is the kilometrage of the vehicle (10 000 kms)
EST	is a dummy for estate car body
DIESEL	is a dummy for diesel engine
FWD	is a dummy for four wheel drive

For some of the models one or more of the dummies were not relevant, and therefore left out of the regression equations. The log-linear model specification

---

[123] The database used was Bilnorge.no, and the records were downloaded in week 31 and 32 in 2002.

makes the interpretation of the estimates fairly simple, e.g. the estimate for β_1 (multiplied by 100) would indicate by how many percent the price would decrease by one additional year of age. Our a priori expectations of the signs of the parameters would be that β_0 should be positive, β_1 and β_2 should take a negative value since we expect prices to fall with increasing age and kilometrage. Finally we would expect the estimates of the dummy parameters to be positive, as estate body, diesel engine and four-wheel drive are all factors that normally contribute to a higher second hand price.

Class	Const	AGE	KMS	EST	FWD	DIESEL	R-sq-adj	n
Class A, Small cars	12.122	-0.122	-0.004	ns	na	0.238	0.82	391
		(-35,10)	(-2.28)			(4.68)		
Class B, Lower middle class cars	12.414	-0.106	-0.006	ns	0.166	0.167	0.72	1419
		(-35.78)	(-2.98)		(4.82)	(6.65)		
Class C, Upper middle class cars	12.596	-0.120	-0.016	0.057	0.166	ns	0.87	737
		(-39.82)	(-9.58)	(4.51)	(3.21)			
Class D, Large cars	13.262	-0.162	-0.015	ns	0.164	ns	0.87	509
		(-29.82)	(-5.05)		(5.36)			
Class V, Vans	12.271	-0.099	-0.018	na	na	na	0.83	132
		(-13.14)	(-5.98)					

Figure 5-2   Estimation results, Vehicle depreciation model[124]

I divided the cars into four size categories, and added a fifth category for vans (2-3,5 tons maximum permitted gross weight). The models chosen to represent each category are not necessarily representative of the class. Here, in this illustration, I

---

[124]   The 'ns' means that the factor was included in the regression analysis, but did not come out with a significant estimator (95% test level). The 'na' means that this factor was not entered into the regression equation because it was not relevant for the category. The numbers in brackets below the estimates are the corresponding t-statistics indicating significance level.

have merely chosen a few models that are among the most popular in each class[125].

All significant estimators seem to come out with the sign expected a priori. Both age and kilometrage are highly significant factors for all vehicle classes. The dummy for estate body is only significant for upper middle class cars, where it seems to increase the expected price with 5-6 percent (compared to sedans and hatchbacks). Four-wheel drive is very popular in Norway, and adds approximately 17 percent to the expected price in middle class and large cars. Diesel engines seem to add 18-27 percent to the price of the two smallest car categories.

However, my major focus here is on age and kilometrage. It seems that an extra year of age reduces the expected second hand car price by 10 to 16 percent (cet. par). Accordingly, having an extra 10 000 kilometres on the distance recorder would lower the expected price by between 0 and 2 percent.

One should, of course, interpret the relationship between the factors age and kilometrage with some caution, because they are quite highly correlated. However, it seems that for all vehicle classes age is the major determinant of price, but still kilometrage turns out significant for all classes. This indicates that there may very well be an element of vehicle depreciation that could be called kilometre-dependent, and hence *marginal*.

By calculating typical levels of registration taxes for the five vehicle groups, combined with marginal kilometre-dependent depreciation by vehicle class, a crude estimate of the marginal part of the registration tax is included in the analysis. Since registration taxes increase by motor size/output and vehicle

---

[125]   Class A (small) is represented by VW Polo, Ford Fiesta and Nissan Micra, class B (lower middle class) by VW Golf, Toyota Corolla, Renault 19 & Megane and Citroen ZX & Xara, class C (upper middle class) by Ford Mondeo, Peugeot 405 & 406 and Mazda 626, class D (large) by Opel Omega, Audi 100 & A6 and BMW 5-series, class V (vans) by VW Transporter, Toyota HiAce and Ford Transit.

weight, the marginal registration tax component varies considerably by vehicle class. This is illustrated in Figure 5-3.

Vehicle class	Estimated marginal part of registration tax (€ per vkm)
Class A, Small cars	0.00116
Class B, Lower middle class cars	0.00271
Class C, Upper middle class cars	0.00820
Class D, Large cars	0.01505
Class V, Vans	0.00320
Average for cars (i.e. ex. Vans)	*0.00678*

Figure 5-3   Estimated marginal part of registration tax by vehicle class

Based on this I will, as an illustration, include this small share of registration taxes in my figures for the current marginal taxation level at the end of my analysis.

5.2.5   E: The Output module, and F: The Graphic Output module

The last two modules in CATERU give the output from the model. In module E the output is summarised in tables representing three "dimensions": Vehicle type, externality type and area (city, town, rural). Finally the sums of the marginal externalities are compared to current levels of marginal taxation in Norway.

In Module E the same figures are presented graphically.

5.3   THE CALCULATION OF EXTERNAL ELEMENTS

Section 5.2 has provided an overview of the main modules of the CATERU model, now I will explain in more detail how the calculations of each external element have been carried out. The figures presented in the tables in this section are collected from

5.3.1   Global emissions

The CO_2-emissions contribute to the problem of global warming. The actual location of the emissions does not affect the impact, so this cost component will

be the same per kg emitted for rural, town, or city environment. Unlike many other substances, CO_2-emissions cannot be gotten rid of by a cleansing technology, the CO_2-emissions are therefore closely linked to the fuel consumption of the vehicles which are given in Figure 5-4.

Vehicle type	Avg. fuel consumption Litres/10 km
Cars (< 2t), petrol	0,82
Cars (< 2t), diesel	0,77
Vans (2-3,4 t), p	1,36
Vans (2-3,4 t), d	1,24
Buses d	2,96
Light trucks (3,5-7,4 t) d	2,08
Trucks (7,5-15,9 t) d	2,80
Trucks (16-23 t) d	3,10
Trucks (> 23t) d	4,31

Figure 5-4   Average fuel consumption figures (Holtskog 2001)

In the calculations these fuel consumption figures are differentiated by location (Rural or Town/City) because driving in congested areas generally gives significantly higher fuel consumption due to frequent acceleration-deceleration cycles. This is done by using the same proportions as in Eriksen et al. (1999).

The fuel consumption figures for each vehicle class are converted into grams per kilometre using the true density of petrol and diesel, respectively. Each kilogram of fuel consumed produces more than 3 kilograms of CO_2-emissions (figures collected from Holtskog (2001):

- 1000 grams of petrol give 3130 grams of CO_2-emissions
- 1000 grams of diesel give 3170 grams of CO_2-emissions

The combination of the fuel consumption figures and the CO_2-emissions per unit of fuel gives the average CO_2-emissions per kilometre for each vehicle class.

These figures should then be multiplied by a shadow-price of CO_2-emissions, reflecting the value of the environmental hazard these emissions represent.

There is an extremely high level of uncertainty about the estimated damage that the problem of global warming might impose on different regions of the world. Indeed, many scientists still argue that global warming is not connected to human activities at all. If this is true, we should not have to deal with CO_2-emissions at all in this context. However, the magnitude of the possible damage that global warming might impose, and the rather high probability that human made emissions do have significant effect, call for a precautionary attitude to this problem area. I would therefore argue that a cost element related to CO_2-emissions should be included in the foundation of road user charges.

Still, even if we accept that there is a link between humanly generated CO_2-emissions and global warming, estimating the magnitude of the expected damage that this global warming might cause is an extremely complex task. We have limited understanding of the factors affecting our global and regional climate, and any scenario of the effects of global warming will therefore be very uncertain. We have indeed seen very different scenarios estimating the expected rise of the sea levels (from a few centimetres to several metres), climatic changes (from a more humid environment and more frequent storms to a new ice-age and massive desertification).

With this extremely high level of uncertainty, there are limited opportunities for estimating shadow prices for CO_2-emissions. We cannot base this on a direct estimate of the damages, so the alternatives would be some sort of indirect valuation of the emissions:

- Either we could use a CVM/Stated Preference technique trying to identify the willingness to pay for reduced CO_2-emissions (or required

compensation for a willingness to accept higher levels), or

- We could estimate the implicit shadow price of expressed policies of reductions or limitations of national, regional or global emission levels.

We could hardly expect any respondents in a CVM or Stated Preference study to be able to relate sensibly to a problem of stating their willingness to pay for a reduced level of CO_2-emissions, given the degree of uncertainty. The alternative then, is to utilise econometric models assessing how high a CO_2-tax would have to be to achieve some specified political aims related to the level of CO_2-emissions. This has been done in some Norwegian studies under different scenarios related to the expressed target levels of emissions, and with respect to the sort of international agreement one is operating under. Apart from deciding the target levels, the central question is: Should the expressed targets be reached with absolute national quotas, or should one allow for CO_2-emission quotas to be traded among nations? For an industrialised nation like Norway, the necessary taxation level would have to be significantly higher under the first regime, than under the tradable permits regime. This is very well illustrated by the estimated shadow prices from two different studies representing these alternatives (Jensen 1998 and Holtsmark & Hagem 1998). Emission levels are according to the levels suggested in the Kyoto-protocol:

- Shadow price of CO_2-emissions without tradable permits: 0,050 € per kg
- Shadow price of CO_2-emissions with tradable permits: 0,015 € per kg[126]

The prospects of a global system with tradable permits are not very good, partly due to the reluctance of key nations to commit to such a system (e.g. USA). However, there are still negotiations on an internationally committing system including a selection of nations. If this is the outcome, then the suitable shadow price for CO_2-emissions should lie somewhere between the two alternatives

presented above. I relate to the low alternative (i.e. with tradable permits) in my Base Scenario, and illustrate some of the uncertainty in my estimates related to this in the sensitivity analysis.

Vehicle type	Estimated marginal cost of CO_2-emissions €per vehicle km	
	Rural areas[127]	Towns & Cities
Cars (< 2t), petrol	0.003	0.004
Cars (< 2t), diesel	0.002	0.005
Vans (2-3,4 t), p	0.005	0.005
Vans (2-3,4 t), d	0.005	0.005
Buses d	0.011	0.017
Light trucks (3,5-7,4 t) d	0.008	0.011
Trucks (7,5-15,9 t) d	0.011	0.016
Trucks (16-23 t) d	0.011	0.021
Trucks (> 23t) d	0.016	0.022

Figure 5-5   Estimated marginal cost of CO_2-emissions, Base Scenario (low CO_2-cost alternative)

5.3.2   Local emissions, including noise

Some of the emissions from road use have very different impacts when emitted in rural areas compared to congested and heavily populated areas. This is reflected in the figures presented in Figure 5-6, which are based on ECMT (1998) and Rosendahl (1999).

The average emission of the Norwegian vehicle park with respect to these factors has been changing rapidly over the last decade, due to the fact that a catalytic converter was made mandatory for all new cars in 1989. This cleansing system dramatically reduced the emissions of CO, NOx, NMVOC and PM10 from vehicles fitted with it. However, the replacement of old vehicles without this

---

[126]   Both figures are updated to 2002 price level by use of the Norwegian Consumer Price Index.

[127]   The distinction between Rural areas and Towns & Cities is based on a factor collected from Eriksen et.al. (1999), illustrating a typical driving pattern and the belonging figures for fuel consumption, and not different shadow prices for CO_2.

device takes many years, since Norway has one of the oldest car parks in Europe, with an average age of cars of 9,9 years. This means that 40% of the cars in Norway still do not have a catalytic converter (OFV 2002).

Emission type	Rural areas	Towns	Cities
VOC emissions	4.356	8.712	8.712
NO$_x$ emissions	4.356	8.712	8.712
PM10 emissions	0	26.587	225.988
SO$_2$ emissions	2.500	9.725	9.725

Figure 5-6   Estimated shadow prices of local road use emissions by area (€ per kilogram)

My figures are based on a fairly recent study (Holtskog 2001), and should therefore be quite representative of the current Norwegian vehicle park. Bearing this in mind, one should note that the figures for local air emissions will diminish over the next decade as well, until the entire vehicle park has the converter system.

A similar situation prevails with respect to the emissions of SO$_2$. These are closely linked to the sulphur content of the fuel, and as new fuel is being introduced with lower sulphur content, these figures will also diminish over time. The reductions in sulphur content started recently (2000), and lower levels will gradually be introduced over the next 5-6 years (according to passed and proposed EU regulations).

The local emission levels are quite different for petrol and diesel vehicles. Due to the relatively better fuel efficiency of the diesel engines, CO$_2$-emissions are lower per km for diesel vehicles. Diesel also gives a relatively lower level of NO$_x$, VOC and CO emissions compared to petrol. However, the problems related to SO$_2$ and PM10 emissions are more severe for diesel driven vehicles than for the equivalent petrol driven ones. The latter problem is also gradually reduced as more vehicles are fitted with particle traps, or alternative measures aimed at getting rid of the

210

particles. The average emission factors based on Holtskog (op.cit.) are given in Figure 5-7.

Vehicle type	SO$_2$	NO$_x$	NMVOC	PM10
Cars (< 2t). petrol	0.16	16.2	26.3	0.26
Cars (< 2t). diesel	0.8	8.1	2.6	3.06
Vans (2-3.4 t). p	0.16	14.1	20	0.2
Vans (2-3.4 t). d	0.8	7.9	3	2.7
Buses d	0.8	38.93	2.77	2.34
Light trucks (3.5-7.4 t) d	0.8	31	3.8	1.9
Trucks (7.5-15.9 t) d	0.8	32.4	3.7	2.2
Trucks (16-23 t) d	0.8	31.1	3.6	2.1
Trucks (> 23t) d	0.8	31.1	3.6	2.1

Figure 5-7   Average emission factors by vehicle type (grams per kg fuel)

Vehicle type	SO$_2$	NO$_x$	NMVOC	PM10
Cars (< 2t), petrol	0.0000	0.0039	0.0064	0.0000
Cars (< 2t), diesel	0.0001	0.0015	0.0005	0.0000
Vans (2-3,4 t), p	0.0000	0.0062	0.0088	0.0000
Vans (2-3,4 t), d	0.0002	0.0036	0.0014	0.0000
Buses d	0.0004	0.0379	0.0027	0.0000
Light trucks (3,5-7,4 t) d	0.0003	0.0221	0.0027	0.0000
Trucks (7,5-15,9 t) d	0.0004	0.0313	0.0036	0.0000
Trucks (16-23 t) d	0.0005	0.0310	0.0036	0.0000
Trucks (> 23t) d	0.0007	0.0455	0.0053	0.0000

Figure 5-8   Estimated costs of local emissions, Rural areas
(€ per vehicle km)

These average emission factors are multiplied by the average fuel consumption figures given in Figure 5-4, and combined with the estimated shadow prices for these emissions in different locations (Figure 5-6). These products represent the estimated costs per vehicle kilometre given in Figure 5-8, Figure 5-9 and Figure 5-10.

Vehicle type	SO$_2$	NO$_x$	NMVOC	PM10
Cars (< 2t), petrol	0.0001	0.0107	0.0174	0.0005
Cars (< 2t), diesel	0.0008	0.0074	0.0024	0.0085
Vans (2-3,4 t), p	0.0002	0.0123	0.0174	0.0005
Vans (2-3,4 t), d	0.0008	0.0072	0.0027	0.0075
Buses d	0.0029	0.1244	0.0089	0.0228
Light trucks (3,5-7,4 t) d	0.0018	0.0612	0.0075	0.0115
Trucks (7,5-15,9 t) d	0.0026	0.0954	0.0109	0.0198
Trucks (16-23 t) d	0.0034	0.1193	0.0138	0.0246
Trucks (> 23t) d	0.0036	0.1262	0.0146	0.0260

**Figure 5-9   Estimated costs of local emissions, Towns
(€ per vehicle km)**

Vehicle type	SO$_2$	NO$_x$	NMVOC	PM10
Cars (< 2t), petrol	0.0001	0.0107	0.0174	0.0045
Cars (< 2t), diesel	0.0008	0.0074	0.0024	0.0723
Vans (2-3,4 t), p	0.0002	0.0123	0.0174	0.0045
Vans (2-3,4 t), d	0.0008	0.0072	0.0027	0.0640
Buses d	0.0029	0.1244	0.0089	0.1940
Light trucks (3,5-7,4 t) d	0.0018	0.0612	0.0075	0.0973
Trucks (7,5-15,9 t) d	0.0026	0.0954	0.0109	0.1681
Trucks (16-23 t) d	0.0034	0.1193	0.0138	0.2089
Trucks (> 23t) d	0.0036	0.1262	0.0146	0.2210

**Figure 5-10   Estimated costs of local emissions, Cities
(€ per vehicle km)**

No new calculations of external noise costs are made

The noise calculations are merely updated to 2002 prices from the calculations made in Eriksen et al. (1999). These calculations are based on a survey of the welfare impacts related to noise reduction stemming from a 50% traffic reduction in cities (Sælensminde and Hammer 1994). The figures are multiplied by an estimated number of affected persons, and eventually allocated by vehicle kilometres and noise equivalency units per vehicle type (light trucks: 5 car units, HGVs and buses: 10 car units). These figures can hardly be said to represent *marginal* noise impacts, but are the only figures presently available for Norwegian roads.

212

Vehicle type	Cities	Towns	Rural
Cars (< 2t), petrol	0.017	0.017	0.000
Cars (< 2t), diesel	0.017	0.017	0.000
Vans (2-3,4 t), p	0.017	0.017	0.000
Vans (2-3,4 t), d	0.017	0.017	0.000
Buses d	0.168	0.168	0.000
Light trucks (3,5-7,4 t) d	0.084	0.084	0.000
Trucks (7,5-15,9 t) d	0.168	0.168	0.000
Trucks (16-23 t) d	0.168	0.168	0.000
Trucks (> 23t) d	0.168	0.168	0.000

Figure 5-11   Marginal external noise costs, updated figures based on Eriksen et al. (1999), € per vehicle km

A framework for proper calculation of external noise costs

Generally the marginal external noise disturbance costs may seem to decrease with traffic volume. On one hand traffic noise expressed in decibel is increasing digressively with the traffic volume. On the other hand, the willingness to pay to reduce noise may very well increase progressively with the decibels. The actual marginal external noise costs of one additional vehicle will thus be very dependent on the prevailing "background" noise level. A noise impact assessment model must therefore be able to represent the environment (receptors, buildings), the vehicle technology (car, HGV etc.) and the traffic situation (e.g. speed and traffic volume) adequately (Bickel and Schmid 2002a). Then this model must be combined with models for the assessment of the responses of humans to noise exposure. In the UNITE project such information is based on the state of the art summary provided in De Kluizenaar et al. (2001), containing detailed assessments of impacts of noise levels on infarction, angina etc. Finally monetary evaluation of the physical impacts is called for. The cost components could be expressed as (Metronomica 2001):

a) Resource costs, i.e. medical costs paid by the health service

b) Opportunity costs, i.e. mainly the costs in terms of productivity lossess

c) Disutility, i.e. other social and economic costs of the individual or others

Such models have been further developed and applied for some case studies under the UNITE project. All the studies emphasise the limited transferability of the case studies to other situations, as the outcome of the studies are very dependent on local circumstances. Still, to illustrate the magnitude of the estimates carried out, I have compiled some tables representing examples of inter-urban and city estimates respectively. The tables are based on two Finnish studies (Tervonen et al. 2002a and Tervonen et al. 2002b), two German studies (Bickel and Schmid 2002b and Bickel and Schmid 2002b), and one Italian study (Enei et al. 2002).

Vehicle type	Basel - Karlsruhe	Strasbourg-Neubrandenburg	Milano-Chiasso	Bologna-Brennero
Car	0.02	0.12	0.01	0.001
LGV	0.06	0.89	0.03	0.001
HGV	0.11	3.04	0.09	0.006
Coach	0.05	0.70	na	na

Figure 5-12   Marginal external noise cost estimates from the UNITE case studies, *Inter-Urban*, Average *daytime* values, €cent per vkm

Vehicle type	Basel - Karlsruhe	Strasbourg-Neubrandenburg	Milano-Chiasso	Bologna-Brennero
Car	0.03	0.19	0.04	0.002
LGV	0.10	1.49	0.12	0.005
HGV	0.18	5.06	0.35	0.021
Coach	0.08	1.17	na	na

Figure 5-13   Marginal external noise cost estimates from the UNITE case studies, *Inter-Urban*, Average *at night* values, €cent per vkm

Vehicle type	Helsinki	Stuttgart	Berlin
Car, daytime	0.22	1.50	0.47
Car, at night	0.53	4.50	1.45
HGV, daytime	1.58	25.75	7.67
HGV, at night	3.86	78.25	23.33

Figure 5-14   Marginal external noise cost estimates from the UNITE case studies, *Urban*, Average values, €cent per vkm

214

Though one should be very careful to make conclusions on the relative magnitudes of external noise costs based on an arbitrary sample of case studies like this, these figures suggest that marginal noise costs are significantly higher at night compared to daytime, and that urban marginal costs may be higher than inter-urban ones. The high marginal costs at night is probably related to the low background noise level, and the higher urban marginal costs is probably due to higher number of persons exposed to the noise in densely populated areas.

However, the most striking fact is the extreme differences between the estimates within the vehicle categories. An HGV driving between Bologna and Brennero imposes a very low cost of 0.006 €cents per vkm, wheas the "same" HGV driving through Stuttgart at night imposes 78.25 €cents per vkm according to these figures.

Comparing the updated figures from Eriksen et al. (1999), presented in Figure 5-11, with the recent European case studies presented above, it seems that the Norwegian estimates for cars are only exceeded by the Stuttgart at night example. The HGV figures, however seems more in line with the German daytime estimates.

Again, it is perfectly clear that the Norwegian estimates are not based on a proper impact pathway approach as outlined above. Adding to this, the examples given above indicate that there is really no rationale for calculating "average" noise externalities for towns and cities, as the actual costs will vary very much with the prevailing situation. I have therefore chosen not to include noise costs in my Recommended Scenario (presented later). This does *not* mean that I generally regard marginal noise costs as being insignificant, but merely that these costs would need to be evaluated under the time and place specific conditions at hand.

5.3.3   Time costs

External time costs are a reflection of the delays imposed on other road users by an extra vehicle entering the road system. These costs are only relevant in congested road networks, and are consequently only included in the calculation of external costs for "Cities".

Calculating the external time costs is not possible without advanced network traffic flow models containing "volume delay functions" making travel time within the network dependent on total traffic volumes. By simulating the effects of extra traffic volumes put out on congested networks, it is possible to calculate the system time loss resulting from the "marginal" increase in traffic load. Based on this system time loss, an average marginal external time cost per vehicle entering the system could be calculated.

Time of day	Oslo	Trondheim
Morning – the hour before peak hour	0.97	0.42
Morning – peak hour	2.50	1.11
Morning – the hour after peak hour	1.11	0.69
Average hour 9.a.m.-3.p.m.[128]	0.69	
Afternoon – the hour before peak	2.78	0.42
Afternoon – peak hour	3.33	1.25
Afternoon – the hour after peak	2.36	0.97
An average hour with low traffic	0.14	

Figure 5-15   Average external congestion cost per vehicle trip in the Oslo and the Trondheim area. 2002 € per vehicle trip[129]

This has been done for the Norwegian cities Trondheim and Oslo[130] in a study applying the EMME/2 network model, and based on networks of 133 and 440 zones respectively (Grue et al. 1997). Some of the results of these simulations are summarised in Figure 5-15, indicating that the external congestion costs at the

---

[128]   Unit costs are increased by 25% due to higher share of commercial traffic (including goods vehicles).

[129]   Figures from Grue, Larsen et al. (1997), inflated with CPI from 1997 to 2002 figures, and transformed from NOK to €.

[130]   Oslo has a population of approximately 500 000 inhabitants, Trondheim 150 000.

216

peak hour is in the area of € 3 in Oslo, and somewhat above € 1 in Trondheim. These are average figures for all routes, and also averages representing a full peak hour. Specific trips may have significantly higher, or significantly lower external congestion costs within the city network[131].

Based on these studies, Eriksen et. al. have calculated average external time costs per vehicle kilometre for cities by taking the average over the whole 24 hours, and averaging over all trips within the networks. Then it is assumed that trucks and buses have twice the impact of cars, and trailer trucks three times the impact of the cars (Eriksen et al. 1999). Adjusted for price increases and converting it into European currency, the estimate is € 0.12 per car-kilometre in cities, € 0.23 per bus/truck-kilometre, and € 0.34 per trailer-truck kilometre.

However, these figures are not very interesting when discussing a pricing regime, because they do not reflect the externalities of the given trip in question. Any practical pricing scheme will (in the foreseeable future) have to rely upon some degree of averaging when it comes to estimating the true external time costs, but figures averaged over the full 24 hours do not represent a good reflection of true externalities related to the individual trip. Pricing according to such figures will mean that there is still far too much traffic in the most congested periods and areas, and too little traffic in periods and areas of the network without congestion. For the sake of comparison with Eriksen et al. (1999), I will still keep this estimate of congestion costs in the scenarios.

5.3.4   Accident costs

As indicated above, the accident cost model is to a large extent based on the outcome of a project conducted by the Institute of Transport Economics in Oslo in 1993. I have replaced some accident frequency figures by more recent statistics, and put in a value of a statistical life (VOSL) according to current Norwegian

---

[131]   Ranging from € 0 to about € 24 in the Oslo morning rush hour (op.cit.)

official figures. Figures have been transformed from NOK to EUR, using the January 2002 exchange rate. The VOSL applied in the base scenario is then close to 2.4 million Euros. In addition to this, some amendments have been made to obtain the accident externality estimates per vehicle km needed in the present context.

I will present the conceptual model applied here in a significantly abbreviated version. The documentation report from the project is very comprehensive, and interested readers are referred to Elvik (1993)[132] for further details.

The aim of the study was to elicit figures for the economic valuation of reduced risk of road accidents. This is based on a comprehensive meta-analysis of willingness to pay (WTP) studies related to reduced fatality risk, and an assessment of different approaches to so-called Quality Ajusted Life Years (QUALY) estimation. The 1993 study also relies heavily on earlier studies (Haukeland 1991a and Haukeland 1991b) where the impacts of traffic accidents on human welfare have been surveyed.

```
┌──────────────┐       ┌──────────────┐       ┌──────────────┐
│ Estimated    │       │ Estimated loss of │   │ Estimated WTP │
│ Willingness to │  →   │ Quality Adjusted │ → │ for reduced risk │
│ Pay (WTP) for │      │ Life Years by    │   │ by accident   │
│ reduced fatality │   │ accident severity │  │ severity group. │
│ risk.        │       │ group.           │   │              │
└──────────────┘       └──────────────┘       └──────────────┘
```
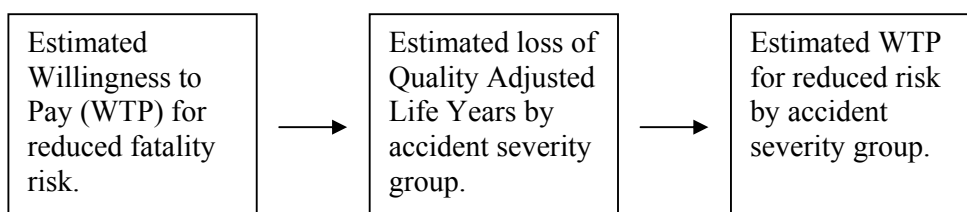
Figure 5-16   Conceptual model for the calculation of WTP for reduced accident risk by severity group

The combination of these studies enables the calculation of WTP figures for different severity groups (Figure 5-16). This is considered to be a representation of the costs related to the *reduced health state* effects of accidents. Additionally

---

[132]   The report is in Norwegian language, but includes an English summary.

the accident effects related to the components of *material well being* are based on the theory of human capital and resource cost.

The meta-analysis of the WTA studies collected 191 different estimates on the value of a statistical life, i.e. a risk reduction that statistically corresponds to avoiding one death. There are considerable differences between the estimates (see Figure 5-17), and Elvik assumes they are due to at least four potential sources:

1. Flaws in research design and data
2. Absence of rational trade-offs in choice situations
3. Systematic variation in WTP related to demand factors
4. Systematic variation in WTP related to contextual factors
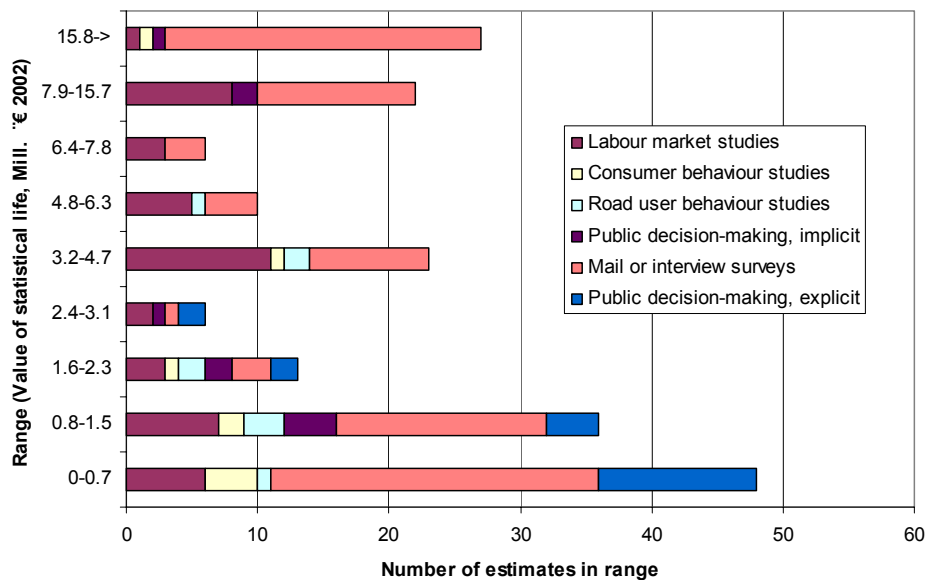
Figure 5-17   Illustration of the dispersion in estimates of WTP for reduced fatality risk in Elvik (1993). Converted to 2002 €[133]

---

[133]   The value ranges in the original Table 7.11 in *(op. cit.)* were given in 1991 NOK. Currency conversion and inflation to 2002 €-figures make the range specification look somewhat strange.

Each study was evaluated subject to a long list of detailed rating criteria, focusing on the scientific validity and reliability of the studies. Then a weight was assigned to each estimate based on an assigned "score value" for each rating criterion. This way the estimates based on the most reliable and appropriate studies were given a more significant impact on the resulting "average" values.

Based on this procedure Elvik calculates a "validity-weighted" average median value of a statistical life corresponding to € 2.75 mill. (after currency and inflation conversion). However, Elvik chooses to rely on a smaller subset of the studies, based on the following considerations:

- Estimates based on labour market studies and consumer behaviour studies are omitted because they are based on other forms of risk than the risk road users are exposed to.
- Likewise, the calculations made to elicit the implicit valuation made by public decision-makers are omitted primarily due to two reasons mainly: Firstly, most of these studies may not have adequate control over the many dimensions present in the prioritisation of alternatives. Secondly, this method would lead to some form of circularity when applying the estimates for future decision-making.
- This leaves studies of road user behaviour, interview surveys and public decision-makers' explicit valuation as potentially relevant for the study.
- Among the first two of these groups of studies, only studies
  - with a fairly large sample size (N>500),
  - which have a representative sample of road users,
  - which are conducted in countries with similar traffic risk levels to Norway,
  - where knowledge and abilities to relate rationally to risk evaluation have been tested,

are selected for further consideration. The selection of studies should also represent different approaches to the estimation problem, allowing an analysis of the possible impact of different methodologies to be evaluated.

- The explicit valuation by public decision-makers, varies considerably between countries, but is still taken into account for comparison with the other WTP studies.

Based on a qualitative review of the remaining subset of studies, Elvik ends up recommending a value of a statistical life (VOSL) of NOK 10 million (1991-prices). This corresponds to about € 1.6 million at 2002 price level[134].

Comparing this to the totality of the studies, as represented in Figure 5-17, this seems like a fairly conservative estimate. However, compared to the majority of the explicit public decision-maker values at that time, the estimate seems quite high.

This VOSL figure is the starting point for the calculation of the external road accident cost by vehicle type, and it is assumed to express the total cost to road users of a fatal accident. This total cost of an injury comprises the costs related to:

- Lost quality of life
- Travel time delay
- Medical treatment
- Lost output
- Property damage, and
- Administrative costs

---

[134] Note that the VOSL-figure used in my further calculations is *not* based on this figure, but on the official Norwegian VOSL, as stated above. However, the rest of the figures for allocating the costs to different affected parties and accident severity groups are all based on the proportions of the original study established by Elvik.

Apart from road users, household members, private third parties and the public sector are also affected by accidents, and thus assumed to carry some of the cost elements. The calculation of cost can therefore be based on the type of information exemplified in Figure 5-18. The figures in this table represent a *fatal* accident, which is one of four different severity groups in Elvik's model. The other severity groups are: Very severe, severe and slight, defined according to Norwegian road accident statistics. I will relate my following review of the calculation procedure to the severity group example used in Figure 5-18.

Type of cost	Road users	Household members	Private third parties	Public sector	Total costs
Lost quality of life	2 126	300	-	-	2 426
Travel time delay	-	-	1	-	1
Medical treatment	-	-	-	1	1
Lost output	259	260	-	200	720
Property damage	7	1	1	-	9
Administrative costs	4	0	0	2	7
Total costs	2 397	561	3	203	3 163

Figure 5-18   Estimated costs of one *fatal* accident (1000 €)[135]

The point of departure is the figure for "Total costs" for Road users (€ 2.397 million), which is based on Elvik's review as described above. Travel time delay is a cost component imposed on private third parties, and is of a rather modest magnitude due to a largely uncongested Norwegian road network. The costs of medical treatment are, to a large extent, paid by the public sector in Norway, but any private costs are assumed to be distributed on a 50/50 basis among road users and their household members. The estimated costs of lost output are assumed to be € 259 000[136] assigned to the road users, an equal amount is assigned to the household, and the rest to the public sector which loses income tax payments. The

---

[135] Based on Table 3 in Elvik, R. (1994). converted to Euros and adjusted to 2002 price level using CPI.

[136] The figures for Medical treatment, Lost output, Property damage and Administrative costs are all based on a survey reported in Hagen, K.-E. (1993). which is a comprehensive review of more than 36 000 reported injuries related to road traffic accidents.

222

calculation of lost output includes both paid and unpaid work, and is based on a discounted average gross wage rate over the lost years of labour force participation (80% of the average wage rate assigned to unpaid household work). The details of this procedure are explained in Haukeland (1991b). The calculation of property damage figures is mainly based on information provided by insurance companies. Most of the property damage is paid for by the road users through their insurance, so the bulk of these costs is already internalised. Administrative costs (insurance administration, police, courts etc.) are calculated in Hagen (1993) based on public accounts and statistics.

The estimated cost assigned to "Lost quality of life" is then calculated as the residual value after the deduction of all these other cost items from the estimated € 2.4 million total costs. For a fatal accident this item is estimated to be € 2.1 million. This far I have focused on the accounts for the road users themselves.

The costs imposed on the household members of the injured or killed road user are estimated as follows: Based on the survey of WTP-studies, the "Lost quality of life"-figure for household members is assumed to be 12,5% of the total costs assigned to road users, i.e. about € 0.3 million.

The figures for "lost quality of life" for the other severity groups are all based on the estimated € 2.4 million figure representing a reduction of risk corresponding to one fatal injury (the value of a statistical life). It is assumed that these figures are proportional to the corresponding estimated number of lost years of living with perfect health. These estimates are based on a QUALY-technique applied in a survey of the daily life of traffic injury victims Haukeland (1991b). Applying these figures, the estimated economic values for lost quality of life for road users by injury severity as follows:

- Very severe: About € 0.60 million per injury

- Severe: About € 0.19 million per injury

- Slight: About € 0.02 million per injury

As for the fatalities, the assumed costs of household members are calculated to be 12,5% of these figures representing the cost of road users themselves. I have not explained every detail underlying the calculations of injury costs in this model, so interested readers are referred to Elvik (1994) and Elvik (1993).

Type of cost	Road users	Household members	Private third parties	Public sector	Total costs
Lost quality of life	599	75	-	-	674
Travel time delay	-	-	1	-	1
Medical treatment	6	6	-	17	30
Lost output	53	53	23	213	342
Property damage	5	1	1	-	7
Administrative costs	3	0	0	6	10
Total costs	667	135	25	236	1 062

Figure 5-19   Estimated costs of one *very severe* accident (1000 €)[137]

Type of cost	Road users	Household members	Private third parties	Public sector	Total costs
Lost quality of life	192	24	-	-	216
Travel time delay	-	-	0	-	0
Medical treatment	2	2	-	13	17
Lost output	9	9	15	71	104
Property damage	4	0	1	-	5
Administrative costs	2	0	0	2	5
Total costs	208	35	16	86	346

Figure 5-20   Estimated costs of one *severe* accident (1000 €)[138]

---

[137]   Cf. footnote 135.

[138]   Cf. footnote 135.

224

Type of cost	Road users	Household members	Private third parties	Public sector	Total costs
Lost quality of life	24	3	-	-	**27**
Travel time delay	-	-	0	-	**0**
Medical treatment	0	0	-	1	**2**
Lost output	-	-	1	1	**2**
Property damage	3	0	0	-	**3**
Administrative costs	2	0	0	0	**3**
Total costs	**29**	**4**	**2**	**3**	**37**

Figure 5-21   Estimated costs of one *slight* accident (1000 €)[139]

Not all the costs estimated in Figure 5-18, Figure 5-19, Figure 5-20, and Figure 5-21 are tax relevant because some of the cost items should be considered internal. In addition to the cost items in the column "Road users", the consumption related figures under "Household members" are also considered to be internal to the road users due to a common economy in the household. The rest of the cost items, covering the "Lost quality of life" for household members, and all the items under "Private third parties" and "Public sector" are considered to be system-external costs to the road users.

Calculating the ratio of external costs to total costs in this way for each severity group gives the corresponding system-external shares expressed as a percentage of total costs per injury. Fatal accidents get an externality share of 18,0%, very severe injuries 37,3%, severe injuries 43,5%, and slight injuries 23,2 %. To obtain an average externality percentage for system-external costs, these figures are multiplied by the estimated number of annual injuries in each category[140], and the external and total costs per injury respectively. Then the sums of annual external and total costs for all injury types are calculated, and the weighted average system-externality figure is found to be 29,3%.

---

[139]   Cf. footnote 135.

[140]   This number is *estimated* due to the expected underreporting of accidents in the official figures. The estimates are based on Hagen, K.-E. (1993)..

Mode of travel	Counterpart in accident										
	Car	Van	Truck	Bus	Moped	MC	Other motor	Bicycle	Pedes-trian	None	Total
Car	11206	460	606	146	-	63	412	21	10	2698	**15622**
Van	151	50	20	20	-	-	-	-	-	60	**301**
Truck	28	-	75	-	-	-	10	-	-	19	**132**
Bus	46	-	11	23	-	-	118	-	-	23	**221**
Moped	742	19	9	38	388	-	49	10	-	1 061	**2 316**
MC	384	31	10	21	-	103	11	-	-	694	**1254**
Other motor	210	-	42	19	-	63	65	-	-	188	**587**
Bicycle	1606	29	48	134	58	39	201	1490	39	9272	**12916**
Pedestrian	1961	77	48		86	105	110	382	19	326	**3114**
Total	**16334**	**666**	**869**	**401**	**532**	**373**	**976**	**1903**	**68**	**14341**	**36463**

Figure 5-22   Traffic injuries by mode of travel and injured group of road users by mode of infliction (Sources: Elvik 1994 and Hagen 1993)

Costs that are system-internal to road users as a group, may be assigned to different categories of road users. In Elvik's study the categories representing the different modes of travel are: Cars, vans, trucks, buses, mopeds, motor cycles, other motorised vehicles, bicycles and pedestrians. Based on accident reports, it is possible to estimate figures for the number of injuries by mode of infliction: That is for each mode of travel one can calculate the number of injuries inflicted by other user groups, the number of self inflicted injuries and the number of accidents inflicted on others. The results are shown in Figure 5-22. Here, the diagonal represents the injuries that are internal to each road user group. The rest of the table illustrates the externalities among the road users (i.e the physical externalities). An example of how the table should be read could be: Among the accidents including a car and a van, 151 persons in vans are injured annually, and 460 persons in cars. This table makes the calculation of physical externalities for each road user group possible. In Figure 5-23, the column "Injuries inflicted on others" represents the physical system-internal externality for each group of road

users[141]. From the last column in the table we can see that the externality expressed as a percentage varies from 4% for bicyclists to 89% for the HGVs and buses. Cars seem to have an average externality of 27% based on this analysis.

Mode of travel	Self inflicted injuries	Injuries inflicted on others (physical externality)	Physical externality percentage
Car	13 904	5 128	27 %
Van	110	616	85 %
Truck	94	794	89 %
Bus	46	378	89 %
Moped	1 449	144	9 %
Motor cycle	797	270	25 %
Other motor	253	911	78 %
Bicycle	10 762	413	4 %
Pedestrian	345	49	12 %
Sum all categories	27 760	8 703	
Percent of total	76.1 %	23.9 %	

Figure 5-23   Physical externalities, based on Elvik (1994)

We have now obtained the physical externality shares for each road user group, and this should be combined with the relevant cost items to produce estimates of external costs per injury. Referring to the cost items in Figure 5-18, Figure 5-19, Figure 5-20 and Figure 5-21 above, the costs represented in the columns for household members, private third parties and the public sector are already accounted for as system externalities.

---

[141]   This figure is termed "gross" physical externality in Elvik, R. (1994). because he also calculates a "net" physical externality figure by deducting the number of injuries inflicted *by* others on each group. In my context this figure does not seem meaningful, and my subsequent calculations are therefore based on the so-called gross physical externalities only.

Injury severity	Internal costs	System external costs	Physical external costs	Total external costs	Total costs
Fatal	532	505	2 126	2 631	3 163
Very severe	115	342	605	947	1 062
Severe	24	128	194	322	346
Slight	5	8	24	32	37
Mean for all injuries	12	25	63	88	100

Figure 5-24   Internal and external costs per physical injury (1000 €)[142]

This leaves us with the items in the column road users as potentially relevant to the physical externalities. In Norway, third party insurance is compulsory, and this means that the cost items "Administrative costs" and "Property damage" are already internalised. To a large extent this also goes for the "Lost output"-figure since compensation for lost income is partially covered by insurance (the part that is not compensated by public social security arrangements). We are then left with the items "Lost quality of life" and "Medical treatment" as relevant items to include in the valuation of physical externalities. Summing up, the value of the total external effects per injury could be calculated as illustrated in Figure 5-24.

Combining these figures with the estimated externality shares by road user group given in Figure 5-23, we are able to calculate figures for the average external cost per injury, by injury severity and road user group (Figure 5-25). The weighted average externality shares are ca. 42% for cars, and in the area of 78-81% for heavier vehicles.

The next step is to multiply these average external costs per injury by the annual number of injuries in each category. I have obtained such figures from Statistics Norway, representing police reported accidents for the year 2000 (the most recent statistics available). As commented on earlier, there is a significant under-

---

[142]   Source: Based on Ibid. Table 8, with some minor alterations due to rounding of figures, plus conversion to Euros at 2002 price level.

228

reporting of accidents (Hagen 1993), and I have therefore used the mark-up factors found in (op.cit.) to inflate the police reported accident numbers into estimated annual numbers, as illustrated in the fourth column of Figure 5-26.

Mode of travel	Injury severity	Physical externality	System externality	Total average external cost per injury	Total costs	Externality %
Car	Fatal	573	505	1 078	3 163	34.1 %
	Very severe	163	336	499	1 062	46.9 %
	Severe	52	127	179	346	51.7 %
	Slight	7	7	14	37	38.0 %
	Mean for all injuries[143]	**17**	**25**	**42**	**100**	**41.7 %**
Van	Fatal	1 804	505	2 309	3 163	73.0 %
	Very severe	514	336	849	1 062	79.9 %
	Severe	164	127	291	346	84.1 %
	Slight	21	7	28	37	76.3 %
	Mean for all injuries	**53**	**25**	**78**	**100**	**77.9 %**
Truck	Fatal	1 901	505	2 406	3 163	76.1 %
	Very severe	541	336	877	1 062	82.5 %
	Severe	173	127	300	346	86.6 %
	Slight	22	7	29	37	79.3 %
	Mean for all injuries	**56**	**25**	**81**	**100**	**80.8 %**
Bus	Fatal	1 896	505	2 401	3 163	75.9 %
	Very severe	540	336	875	1 062	82.4 %
	Severe	173	127	299	346	86.5 %
	Slight	22	7	29	37	79.1 %
	Mean for all injuries	**56**	**25**	**81**	**100**	**80.6 %**

Figure 5-25   Average external costs by mode of travel and injury severity (1000 €)

---

[143]   The mean for all injuries is a weighted mean, with the number of injured persons in each category used as weights.

In column five, the total annual external costs are calculated and summarized for each vehicle type. Finally this figure is divided by the total annual vehicle kilometres on the Norwegian road network to obtain the estimates for the average external accident costs per vehicle kilometre.

Mode of travel	Injury severity	Registered number of injuries	Markup factor	Estimated annual number of injuries	Tot. ann. ext. acc. cost (mill. €)	Total # of annual vehicle kms (mill. km)	Average external accident cost (€ per veh. km)
Car	Fatal	199	1	199	214		
	Very severe	77	1.92	148	74		
	Severe	647	1.92	1 242	222		
	Slight	7 074	3.25	22 991	320		
	Total			**24 580**	**831**	**25 461**	**0.033**
Van	Fatal	6	1	6	14		
	Very severe	6	1.92	12	10		
	Severe	43	1.92	83	24		
	Slight	511	3.25	1 661	46		
	Total			**1 761**	**94**	**5 026**	**0.019**
Truck	Fatal	13	1	13	31		
	Very severe	1	1.92	2	2		
	Severe	19	1.92	36	11		
	Slight	175	3.25	569	17		
	Total			**620**	**60**	**2 457**	**0.025**
Bus	Fatal	6	1	6	14		
	Very severe	-	1.92	-	-		
	Severe	6	1.92	12	3		
	Slight	124	3.25	403	12		
	Total			**421**	**30**	**627**	**0.047**

Figure 5-26   Calculation of average external accident cost per vehicle km

Step 1:

Identify relevant cost items and stakeholders affected by road accidents

Cost items:
- Lost quality of life
- Travel time delay
- Medical treatment
- Lost output
- Property damage
- Administrative costs

Stakeholders:
- Road users
- Household members
- Private third parties
- Public sector

Step 2:

Quantify cost items per injury for the stakeholders

Meta-study of WTP for reduced fatality risk (Elvik 1994)

Recommended figure for the WTP for a risk reduction equivalent to saving a statistical life

Study of the welfare of accident victims and families (Haukeland 1991a)

Estimated loss of Quality Adjusted Life Years in different accident severity groups

Study of impacts of traffic accidents on the costs of social security systems (Hagen 1993)

Assessment of Resource costs (medical treatment, lost output, property damage and administrative cost)

Total costs per injury by severity type

Step 3:

Identify system and physical externalities

All cost items for private third parties and the public sector are considered external

Lost quality of life costs for household members are considered external

System external costs per injury

Calculate physical externality shares for each road user group based on the number of injuries imposed on other road user groups

Multiply externality shares by the system internal cost components

Physical external costs per injury

Figure 5-27   Step 1 to 3 of the accident externalities calculation procedure

231
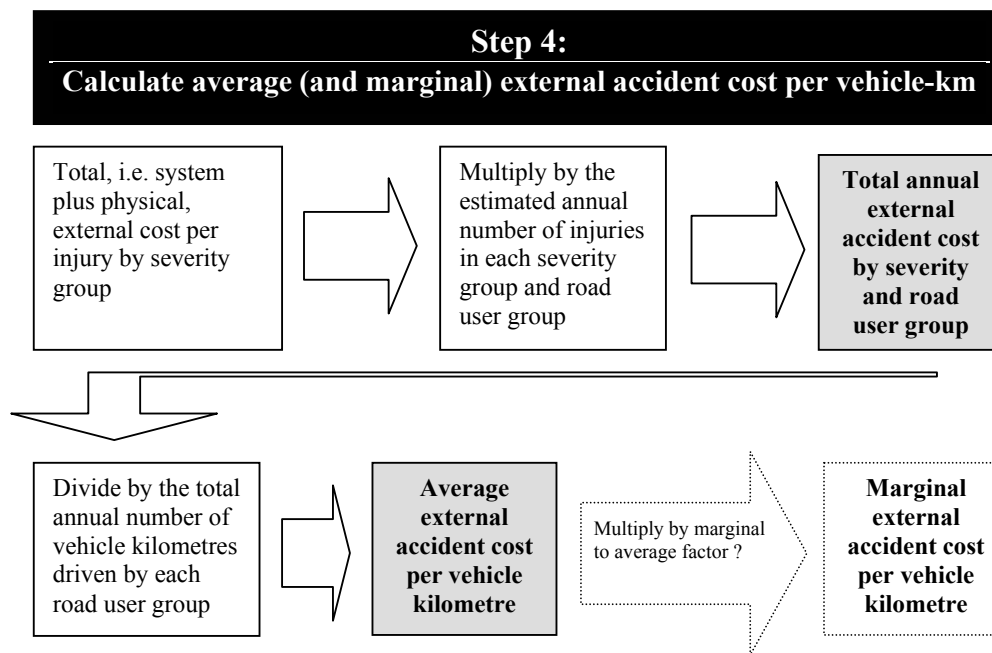
Figure 5-28   Step 4 of the accident externalities calculation procedure

Externalities within the road user groups are not included

Based on the model established by Elvik, I have now calculated average externalities accounting for accident costs related to road use. However, there is one important element that I have not included: The externalities that occur *within* the road user groups (e.g. in accidents involving two cars). These externalities are just as relevant as the system externalities and the physical externalities that I have accounted for, but my statistical information does not allow for the calculation of this cost element.

About the relationship between marginal and average accident cost

I have now presented a framework for calculating *average* external accident cost per vehicle kilometre. For pricing purposes the relevant cost notion would be *marginal* external accident costs. In many studies, e.g. the US Federal Highway Cost Allocation Study marginal accident costs are assumed to equal the average costs. The seminal works of Walters (1968) and Newbery (1988c) both raise this issue, and both conclude that there is scarce empirical evidence for such a

232

relationship. Jansson (1994) proposes a modelling approach to investigate this relationship, and Lindberg (2001a) and Lindberg (2001b) are examples of efforts made to establish an empirical foundation of marginal accident costs, based on the theoretical developments made by Newbery and Jansson. These estimates made for Sweden are reproduced in Figure 5-29. Comparing these estimates to the ones made based on the Elvik model shows that the estimates for car are in the same area when comparing the Swedish figures for Urban car, and the general Norwegian figures for car, whereas the non-urban car value for Sweden is significantly lower than the general Norwegian estimate. For the heavy vehicles the Norwegian figures for buses are much in line with the Swedish estimates, whereas the Norwegian HGV figures are lower than the Swedish figures for heavy vehicles.

Vehicle category	Location	Average cost		Total external costs
		$r(a+b)$	Rc	$r(a+b+c)[(1-\theta)+E]+\theta rc$
Car	Non-urban	79	10	11
	Urban	76	10	28
Heavy Vehicle	Non-urban	86	10	32
	Urban	100	13	45
Unprotected	Non-urban	42	4	-19
	Urban	42	5	-18

Figure 5-29   Average accident cost and marginal external accident cost in Sweden, €/1000 vehicle-km (Source: Lindberg 2001b)

The average cost figures presented in the Lindberg study are not comparable to the figures that I have estimated based on the Elvik model. These figures represent the *total accident costs* divided by the number of vehicle kilometres driven by each vehicle category, whereas the Elvik model yields *average external costs*[144]. Still, the fact that the figures I have calculated does not properly represent true *marginal* external costs, remains. In order to obtain such figures one needs to

---

[144]   Still excluding the intra-mode element.

consider the relevant *risk elasticities*, describing the relative relationship between accident risk and traffic volume. Empirical estimates of such elasticities are scarce, especially with respect to the Norwegian road network. However, a comprehensive Norwegian econometric accident model documented in (Fridstrøm 1999a) and (Fridstrøm 1999b), and this study suggests that the marginal accident costs for cars may very well be close to zero (or even negative), possibly due to the calming effect of traffic congestion. The same study indicates at the same time that marginal costs for heavy vehicles may be substantial (Fridstrøm 2000).

Obviously, having an accident externality model directly estimating the marginal, rather than the average, accident cost would have been more satisfactory. However, given the limited time and resources for this doctoral work, I have chosen not to develop such a model, but merely to update the model originally developed by Elvik. By including average, rather than marginal, external accident costs in my scenarios, I implicitly assume these to be equal. There is really no theoretical or empirical foundation for this assumption (Lindberg 2001b). Theoretically the risk elasticity may assume values less than, equal to, or greater than unity. The true relationship between marginal and average figures will probably be very different from case to case, varying with accident types (car-car, car-pedestrian, car-truck etc.), traffic speeds, traffic volumes and traffic densities.

The theoretical approach used by *op. cit.* is also adopted by the UNITE programme as the preferred methodology (Bossche et al. 2001) for estimating marginal external accident costs. An outline of the methodology is presented in the Appendix.

5.3.5   Road wear

The estimation of marginal road wear in the TØI-model has to a large extent been based on the same principles since the early 1970s. Short-term marginal costs are assumed to be approximated by average traffic-volume dependent maintenance

costs per vehicle kilometre. These traffic-volume dependent maintenance costs are allocated to different vehicle classes based on a moderated[145] AASHO 4th power rule for calculating equivalent standard axle loads (ESALs).

	# of axles	Avg. Veh. weight	% axle #1	% axle #2	% axle #3	% axle #4	% axle #5	ESAL per veh.	Total # of ESAL-km (mill.)
Cars (< 2t), p	2	1	50 %	50 %	0 %	0 %	0 %	0.0019	44
Cars (< 2t), d	2	1	50 %	50 %	0 %	0 %	0 %	0.0019	4
Vans (2-3,4 t), p	2	2	43 %	58 %	0 %	0 %	0 %	0.0109	23
Vans (2-3,4 t), d	2	2	43 %	58 %	0 %	0 %	0 %	0.0109	32
Buses d	2	11	42 %	58 %	0 %	0 %	0 %	0.7806	489
Light trucks (3,5-7,4 t) d	2	3.3	41 %	59 %	0 %	0 %	0 %	0.0389	33
Trucks 7,5-15,9 t) d	2	8.2	41 %	59 %	0 %	0 %	0 %	0.3791	100
Trucks (16-23 t) d	2	9-15.75	42 %	58 %	0 %	0 %	0 %	1.6128	1 018
Trucks (> 23t) d[146]	3	12-40.33	-	-	-	-	-	2.7698	2 006
Other road vehicles	2	2.5	41 %	59 %	0 %	0 %	0 %	0.0195	1.62
Total									**3 749**

Figure 5-30   Calculation of total number of ESAL-kilometres

In Figure 5-30 the average weight of each vehicle category is distributed on the axles based on their relative share of maximum permitted weight. In order to obtain the total number of ESAL kilometres, the calculated ESAL per vehicle is multiplied by the number of annual vehicle kilometres driven by each vehicle group. Summing up for all vehicle groups, a grand total of 3 749 million ESAL kilometres are carried by the Norwegian road network annually. The axle load distribution of the heaviest vehicles is calculated in Figure 5-31, and then inserted

---

[145]   The transferability of the AASHO Road Test results from the early 1960s, was evaluated in a Nordic project called STINA in the mid 1970s. Based on this assessment a 2.5 power rule was applied instead of the 4th power rule.

[146]   See separate calculation procedure for axle weights in Figure 5-31.

into Figure 5-30. The average axle load distribution is based on available statistics on load factors and share of tours with empty vehicles, and driving with a trailer, respectively.

	Number of axles	Veh. weight	% axle #1	% axle #2	% axle #3	% axle #4	% axle #5	ESALs per vehicle	Comment	% of veh.km within group
Trucks (16-23 t) d	2	9	42 %	58 %	0 %	0 %	0 %	0.4727	Empty	20.95 %
Trucks (16-23 t) d	2	15.75	42 %	58 %	0 %	0 %	0 %	1.9150	Loaded	79.05 %
Trucks (> 23t) d	3	12	33 %	33 %	33 %	0 %	0 %	0.5027	Empty, no trailer	16.17 %
Trucks (> 23t) d	3	19.95	33 %	33 %	33 %	0 %	0 %	1.7916	Loaded, no trailer	31.28 %
Trucks (> 23t) d	5	20	20 %	20 %	20 %	20 %	20 %	0.8400	Empty, with trailer	10.49 %
Trucks (> 23t) d	5	40.33	20 %	20 %	20 %	20 %	20 %	4.8503	Loaded, with trailer	42.06 %

Figure 5-31   Calculation of ESALs per vehicle kilometre and the percentage of the vehicle kilometres carried out by each sub-category of the heaviest vehicles

The total annual expenditure on road operation and maintenance is estimated to be approximately € 600 million[147] for the entire Norwegian road network. 35% of these costs are assumed to be dependent on traffic *volume* (Eriksen et al. 1999), so the relevant costs for calculating these "average marginal" cost figures are approximately € 210 million. This figure is then corrected for average VAT on road maintenance (15%, according to Sund (2001), and divided by the total number of ESAL-kilometres driven, and an average marginal cost of € 0.05 per ESAL-kilometre is then derived. Multiplying this figure by the number of ESALs per vehicle in all vehicle groups gives the estimated road wear per vehicle

---

[147]   This estimate is based on the average allocation over the years 1998-2002 for national roads. A mark-up of 25% is applied for covering other public roads. Based on a survey of the accounts performed in Eriksen et. al (1999), this figure represents 2/3 of total o&m costs, thus leaving out non-relevant infrastructure like ferries etc.

236

kilometre given in Figure 5-32.

Vehicle group	€ per vehicle kilometre
Cars (< 2t), p	0.0001
Cars (< 2t), d	0.0001
Vans (2-3,4 t), p	0.0005
Vans (2-3,4 t), d	0.0005
Buses d	0.0380
Light trucks (3,5-7,4 t) d	0.0019
Trucks (7,5-15,9 t) d	0.0185
Trucks (16-23 t) d	0.0785
Trucks (> 23t) d	0.1348

**Figure 5-32   Estimated average traffic volume dependent road wear costs
per vehicle kilometre**

This approach, a typical top-down procedure for estimating marginal road wear costs, has some rather obvious weaknesses. *Firstly*, these costs are based on a rather crude analysis of accounts covering the whole Norwegian network of national roads. The allocation of costs to the various account categories may not give a true picture of the basic cost structures as costs may be posted somewhat arbitrary, or according to available allocations. *Secondly*, The chosen proportion of traffic volume-dependent costs (35%) is based on a survey of international figures that may very well *not* be representative of Norwegian roads. *Thirdly*, the costs estimated are not really *marginal* costs, but average traffic volume dependent costs.

For the sake of comparison with earlier Norwegian studies, I have still chosen to present this model with updated input figures in the base scenario. In another scenario I have presented a bottom-up approach based on the FAMAROW-model for rut-depth development presented in the previous section.

I have now presented the major framework of the modules constituting the CATERU-model. In the following section I will apply the model to illustrate the

237

sensitivity of the marginal cost estimates to some alternative inputs.

5.4   MODEL OUTPUT

To illustrate the sensitivity of the marginal cost estimates to alternative inputs and, to some extent, methodological approaches, I have defined 4 different scenarios and an additional sensitivity analysis on the base scenario.

5.4.1   The scenarios

The Base Scenario

The Base scenario actually represents a reconstruction of the TØI model as it is presented in Eriksen et al. (1999). The reconstruction is partly based on the information given in *op.cit.*, and partly by gathering additional information from authors and written sources that the TØI-model builds on. The last calculations made by TØI used figures available in 1999. I have gathered updated information where available, and inserted it into the model. Whereever new cost figures are not available, the figures in *op.cit*. have been inflated by the Norwegian CPI to represent the 2002 price level consistently applied in this thesis.

The main updates are:

- The *traffic production figures* are updated from 1997 to 2000 based on Rideng (2001). The *distribution* of traffic figures between cities, towns and rural areas, and between vehicle classes, are assumed to be the same as in Eriksen et al. (1999).
- New figures for *fuel consumption* have been collected from Holtskog (2001) and Bang et al. (1999). The differentiation between Town/Cities and Rural areas is assumed to be the same as in Eriksen et al. (1999).
- Average *air-emission factors* per fuel unit for the different vehicle classes are now based on Holtskog (2001).

- The figures for *police reported accidents* have been updated to year 2000 figures based on data commissioned from Statistics Norway (SSB). The assumed level of under-reporting is the same as in Eriksen et al. (1999).

- The *proportion of the traffic production performed by empty and loaded HGVs* has been updated based on SSB (2000).

- *Fuel taxes and registration taxes* are updated to January 2002 figures based on information from the Norwegian Ministry of Finance.

- The *value of a statistical life* (VOSL) is updated based on Norwegian official figures used for road investment appraisal.

- Annual *operation and maintenance costs* for the Norwegian national road network have been updated, and are now based on a price-level adjusted average of the allocations given for 1998-2002.

- To convert accounted operations and maintenance costs to *factor prices*, an estimated average VAT of 15% is deducted Sund (2001).

All in all these updates represent some small and some considerable alterations to the figures applied in Eriksen et al. (1999). Rather than commenting on each change in the input variables, I will comment on the major differences when comparing the output of the Base scenario to the figures presented by *op.cit.*.

The FAMAROW Scenario

The next scenario is identical to the base scenario with one exception: The top-down approach of the TØI-model with respect to marginal road wear costs has been substituted by a bottom-up approach based on the FAMAROW rut-depth development model presented above. This is done by applying the model to predict the number of axle passages necessary to reach the maintenance-triggering level of rut-depth (25mm). The model yields different predictions depending on precipitation and temperature levels, and also different predictions for counties 7 and 9. As an illustration of the principle, I have used the sample mean figures for the climatic factors and not included the correction factors for the counties

mentioned.

The model predicts, under these conditions, a pavement life corresponding to 90 million axle passages. With an estimated average overlay factor cost of € 45 500, this yields an average road wear cost of € 0.51 per 1000 axle kilometres, or approximately € 1 per 1000 vehicle kilometres for a 2-axle vehicle.

Since the most plausible explanation for this traffic volume related road wear is the extensive use of studded tyres in Norway, this cost should mainly be allocated to cars and light vans (HGVs rarely use studded tyres, rather snow-chains under extreme conditions). However, since cars and vans are responsible for more than 90 percent of the axle-kilometres driven on Norwegian roads, this would not make much difference to the estimated cost per axle. In this scenario these costs are allocated to all vehicle classes according to the typical number of axles for each class.

As can be seen when the results are presented below, this bottom-up calculation of marginal road wear results in much lower estimates than the traditional top-down approach in the TØI-model. Multiplying the bottom-up estimate by the total traffic production figures yields a total of € 36 million, compared to the € 182 million allocated through the top-down approach. The distribution of the costs between vehicle classes is also very different, as the TØI-model assumed these costs to be proportional to the number of ESALs, and not to the number of axles in the FAMAROW scenario.

The European Scenario
This scenario contains generic factor prices from recent European research programmes, mainly from the UNITE programme. I have substituted the average value of time (VOT) estimate of approximately €8 per car hour used in Eriksen et al. (1999) to a new value of € 15. This is done with reference to Nellthorp et al. (2001) where generic VOT values are proposed. This value is applied solely to the

City figures in my framework, and the travellers in the rush-hours are predominantly commuters. Nellthorp et al. (2001) suggest a VOT of € 6 (1998) for commuters, but a much higher VOT for business trips. I have therefore used a VOT of € 8, which amounts to € 15 when converted into 2002 price level, and multiplied by 1.75 which is the average car occupancy rate on the Norwegian road network (according to Holtsmark and Hagem 1998).

Further, I have also applied a value of statistical life (VOSL) from the same UNITE-publication, where the recommended figure for Norway is € 1.93 million at 1998 price level. This corresponds to € 2.1 million at 2002 price level, which is 13 % lower than the figure applied in the Base scenario.

In Capros and Mantzos (2000) an average factor price of CO_2-emissions in Europe is proposed to be €20 per tonne. This figure is also used in the UNITE Pilot Accounts for Germany and Switzerland, and is now entered into the European scenario (adjusted for inflation, € 21 per tonne). In my Base scenario the equivalent figure was founded on the Norwegian estimate of a shadow price on the Kyoto agreement (alternative *with* tradable permits in Holtsmark and Hagem (1998). This figure equals € 15 per tonne when updated to 2002 price level. The figure applied in the European scenario is therefore somewhat higher than this Norwegian estimate, but still significantly lower than the Norwegian estimate *without* tradable permits (€ 50 per tonne). One should also note that both these studies date back a few years, and that CO_2-emission levels have not been reduced according to Kyoto aims over the recent years. This means that the shadow prices for achieving the goals in the Kyoto protocol probably should be even higher now.

Instead of the marginal road wear costs based on the Norwegian top-down procedure, I have based these figures on a recent Swedish study conducted under the UNITE programme (Lindberg 2002b). These estimates are partly based on data collected under the Swedish Long Term Pavement Performance (LTPP)

programme where road deterioration has been studied since 1985. The theoretical approach is based on a development of the models developed by David Newbery in the late 1980s (see section 2.4 for a closer presentation of Newberys Fundamental Theorem of Road User Charges). Lindberg's marginal road wear costs per ESAL are differentiated by the strength of the road (measured by the Surface Curvature Index) and by traffic volume (ESALs/day). This results in marginal cost estimates, measured in €/100ESAL-km, between 0.07 (strong road/low volume) and 1.62 (weak road/high volume). The CATERU model does not discriminate between road classes, so I have merely applied the sample mean marginal cost of 0.8 €/100ESAL-km from Lindberg's study. This is a much lower figure than the one applied in the Base scenario (4.9 €/100ESAL-km). Applying the average figure presented by Lindberg, and multiplying this with the total annual traffic production measured in ESALs, yields an estimated marginal road wear of € 30[148] million a year for the Norwegian road network. The equivalent figure under the FAMAROW Scenario was € 36 million. Hence, it seems that these bottom-up approaches yield much lower estimates for total traffic volume dependent road wear costs than the top-down approach applied this far by TØI (yielding an estimated € 182 million/year).

The Recommended Scenario

Finally, I have constructed what I have called a "Recommended Scenario". Actually this is a slightly misleading name, because I do actually *not* recommend a charging system based on average figures for the whole road network, regardless of time and location. I will return to this when discussing a new pricing policy for Norway in the following section. However, for comparison with earlier estimates, I have included this fourth scenario where I combine some of the alternative inputs used in the three other scenarios into a recommended scenario for describing an overall marginal cost level. These estimates would be relevant if the

---

[148] As stated above, the Swedish overlay cost estimates seem to be much lower than the Norwegian equivalents. Applying Norwegian figures for overlay costs would therefore almost double this amount.

current pricing regime, where fuel excise duties are almost the only taxing instrument for marginal road use, is prolonged.

Compared to the Base scenario I have deleted the cost element related to *marginal noise costs*. The empirical foundation for the costs included in the TØI-model is very weak. This estimate was based on Sælensminde and Hammer (1994) where a stated preference survey was used for investigating the willingness to pay for a noise reduction of 20%. This is by no means a marginal change, as such a reduction would necessitate a 52.3% reduction in traffic levels. Noise is predominantly a problem in towns and cities with a high number of affected persons. When traffic volumes are high, *marginal* changes in traffic volumes in these areas will typically *not* result in very significant changes in the magnitude of the noise problem. However, the combination of high population densities and low traffic volumes at night, may cause considerable marginal noise costs. The fact that I have removed the noise cost estimates stemming from the Base scenario does not mean that I generally consider noise costs to be insignificant, but that the actual costs should be calculated under the specific situation at hand. The main cost drivers of marginal noise costs are the level of "background" noise level (e.g. stemming from current traffic volumes), vehicle characteristics, speed, traffic flow, physical surroundings, and the number of affected persons.

The second change relative to the Base scenario is that I have included the *FAMAROW-based road wear* estimates instead of the top-down estimates of the TØI-model. I am aware of the fact that there are several weaknesses in my modelling approach, and that the empirical foundation for choosing these estimates should be corroborated by further research. However, since my results seem to match the outcome of the most recent Swedish study as well[149], I choose this approach for my Recommended scenario.

---

[149]   Regarding the accumulated level of marginal road wear.

In the light of the discussion of the *shadow prices for CO_2-emmissions* above, I also find the figures applied in the TØI-model to be rather on the low side. There is of course considerable uncertainty about the damaging effect of greenhouse gas emissions. However, having chosen to base the shadow price on political agreements (i.e. the Kyoto protocol), I have argued above that the € 15 per tonne applied in the Base scenario is probably too low, even if a system with tradable permits should be implemented. I have therefore included the € 21 per tonne also applied in the European scenario. This figure may very well be on the conservative side, as much higher values have been suggested in other studies (e.g. Suter et al. 2002).

I have now described the different inputs in the four scenarios, and in the following sections I present the output from the scenarios and also conduct a small sensitivity analysis for the Base scenario on some central parameters.

Vehicle class	CO2	SO2	NOx	NM-VOC	PM10	Noise	Time	Acci-dents	Road wear	Total
Cars (< 2t), petrol	0.003	0.000	0.004	0.006	0.000	0.000	0.000	0.033	0.000	*0.046*
Cars (< 2t), diesel	0.002	0.000	0.001	0.000	0.000	0.000	0.000	0.033	0.000	*0.037*
Vans (2-3,4 t), p	0.005	0.000	0.006	0.009	0.000	0.000	0.000	0.019	0.001	*0.039*
Vans (2-3,4 t), d	0.005	0.000	0.004	0.001	0.000	0.000	0.000	0.019	0.001	*0.029*
Buses d	0.011	0.000	0.038	0.003	0.000	0.000	0.000	0.047	0.038	*0.137*
Light trucks (3,5-7,4 t) d	0.008	0.000	0.022	0.003	0.000	0.000	0.000	0.025	0.002	*0.059*
Trucks (7,5-15,9 t) d	0.011	0.000	0.031	0.004	0.000	0.000	0.000	0.025	0.018	*0.089*
Trucks (16-23 t) d	0.011	0.000	0.031	0.004	0.000	0.000	0.000	0.025	0.078	*0.149*
Trucks (> 23t) d	0.016	0.001	0.045	0.005	0.000	0.000	0.000	0.025	0.135	*0.227*

Figure 5-33   Base scenario: Marginal costs (2002 €) per vkm, Rural areas

5.4.2 *Comparing the results of the TØI-model with the CATERU Base Scenario*

The output from the CATERU Base scenario is given in Figure 5-33, Figure 5-34 and in Figure 5-35, representing cost estimates for rural areas, towns and cities

244

respectively.

The totals for towns are about twice as high as the corresponding figures for the rural areas. This is mainly due to higher fuel consumption figures per kilometre, and higher factor prices for local emissions in towns and cities. External time costs are assumed to be zero for both the rural and the town case. Marginal accident costs and road wear costs are assumed to be equal for all areas.

Vehicle class	CO2	SO2	NOx	NM-VOC	PM10	Noise	Time	Acci-dents	Road wear	Total
Cars (< 2t), petrol	0.004	0.000	0.011	0.017	0.001	0.017	0.000	0.033	0.000	*0.082*
Cars (< 2t), diesel	0.005	0.001	0.007	0.002	0.009	0.017	0.000	0.033	0.000	*0.074*
Vans (2-3,4 t), p	0.005	0.000	0.012	0.017	0.001	0.017	0.000	0.019	0.001	*0.071*
Vans (2-3,4 t), d	0.005	0.001	0.007	0.003	0.008	0.017	0.000	0.019	0.001	*0.059*
Buses d	0.017	0.003	0.124	0.009	0.023	0.168	0.000	0.047	0.038	*0.429*
Light trucks (3,5-7,4 t) d	0.011	0.002	0.061	0.008	0.011	0.084	0.000	0.025	0.002	*0.203*
Trucks (7,5-15,9 t) d	0.016	0.003	0.095	0.011	0.020	0.168	0.000	0.025	0.018	*0.356*
Trucks (16-23 t) d	0.021	0.003	0.119	0.014	0.025	0.168	0.000	0.025	0.078	*0.453*
Trucks (> 23t) d	0.022	0.004	0.126	0.015	0.026	0.168	0.000	0.025	0.135	*0.520*

Figure 5-34   Base scenario: Marginal costs (2002 €) per vkm, Towns

The totals for cities are significantly higher than the ones for towns, due to the added external time cost element.

The graphic presentation of the figures enables easy comparison of the magnitude of the different cost components for the various vehicle classes. In Figure 5-36 the cost estimates for the rural case are given, and in Figure 5-37 the equivalent figures for the city case are illustrated.

For cars and vans the CO_2-component is very dominant, especially for the diesel-driven vehicles where the significance of other air emissions is small. For buses and heavy goods vehicles (HGVs), the allocated marginal road wear becomes very significant (allocated by ESALs in the Base scenario), as do the NO_x-

emissions.

Vehicle class	CO2	SO2	NOx	NM-VOC	PM10	Noise	Time	Acci-dents	Road wear	Total
Cars (< 2t), petrol	0.004	0.000	0.011	0.017	0.004	0.017	0.113	0.033	0.000	*0.199*
Cars (< 2t), diesel	0.005	0.001	0.007	0.002	0.072	0.017	0.113	0.033	0.000	*0.250*
Vans (2-3,4 t), p	0.005	0.000	0.012	0.017	0.005	0.017	0.113	0.019	0.001	*0.188*
Vans (2-3,4 t), d	0.005	0.001	0.007	0.003	0.064	0.017	0.113	0.019	0.001	*0.229*
Buses d	0.017	0.003	0.124	0.009	0.194	0.168	0.226	0.047	0.038	*0.827*
Light trucks (3,5-7,4 t) d	0.011	0.002	0.061	0.008	0.097	0.084	0.226	0.025	0.002	*0.515*
Trucks (7,5-15,9 t) d	0.016	0.003	0.095	0.011	0.168	0.168	0.226	0.025	0.018	*0.730*
Trucks (16-23 t) d	0.021	0.003	0.119	0.014	0.209	0.168	0.226	0.025	0.078	*0.863*
Trucks (> 23t) d	0.022	0.004	0.126	0.015	0.221	0.168	0.339	0.025	0.135	*1.054*

Figure 5-35   Base scenario: Marginal costs (2002 €) per vkm, Cities
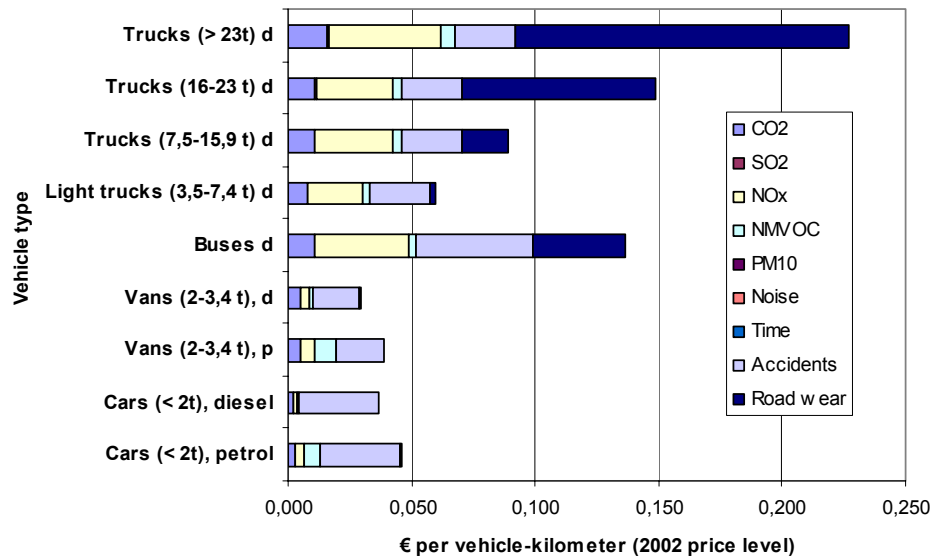

Figure 5-36   Base scenario: Marginal costs, 2002 € per vkm, Rural areas

For the cities the picture becomes quite different[150]. Now, local emissions, noise and external time costs become dominant. For petrol driven cars time costs now

---

[150]   Note that the scale of the two figures are not the same!

constitute more than 50% of the estimated cost, whereas for the HGVs time costs represent about 30% due to the more significant local pollution and noise costs. Road wear has the same absolute magnitude as in the rural case, but becomes less dominant in the city case due to the increase in other cost components.
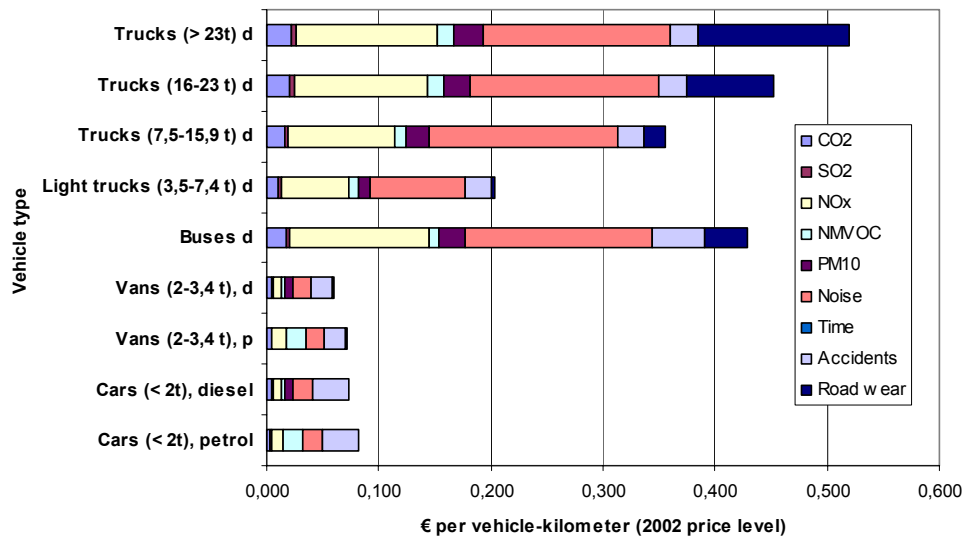
Figure 5-37   Base scenario: Marginal costs, 2002 € per vkm, Cities

The Base scenario is built on the traditional Norwegian model for estimating the marginal costs of road use (the TØI-model). However, in the Base scenario a lot of input figures have been updated, and a comparison between the Base scenario and the results presented in the last published version of the TØI-model (Eriksen et al. 1999) should be of some interest. The totals for all three locations are presented in Figure 5-38. Here the Eriksen figures are inflated by the use of the Norwegian consumer price index, and the original NOK figures are converted into Euros using the January 2002 exchange rate.

In the rural case the differences are quite small. Apart from diesel cars, all CATERU estimates are lower than the ones based on (op.cit.). The relative picture is the same for the town and the city case as well, but the changes are more

significant for these areas. The dominant effects are stemming from:

- Better fuel efficiency for heavy vehicles (buses and HGVs), thus reducing the costs related to emissions
- Low sulphur diesel having been introduced
- A reduction in accident costs for the same vehicle groups due to lower accident frequencies
- Lower local emissions from cars and vans due to a higher proportion of the vehicle fleet being fitted with catalytic converters.

Vehicle class	Rural		Towns		Cities	
	CATERU Base sc.	TØI-model	CATERU Base sc.	TØI-model	CATERU Base sc.	TØI-model
Cars (< 2t), petrol	0.046	0.045	0.082	0.070	0.199	0.186
Cars (< 2t), diesel	0.037	0.032	0.074	0.072	0.250	0.242
Vans (2-3,4 t), p	0.039	0.040	0.071	0.069	0.188	0.185
Vans (2-3,4 t), d	0.029	0.028	0.059	0.061	0.229	0.231
Buses d	0.137	0.165	0.429	0.586	0.827	1.143
Light trucks (3,5-7,4 t) d	0.059	0.069	0.203	0.263	0.515	0.619
Trucks (7,5-15,9 t) d	0.089	0.101	0.356	0.433	0.730	0.872
Trucks (16-23 t) d	0.149	0.169	0.453	0.589	0.863	1.263
Trucks (> 23t) d	0.227	0.247	0.520	0.650	1.054	1.324

Figure 5-38   Comparison of estimates from (Eriksen et al. 1999)[151] and the CATERU Base scenario, 2002 € per vkm

To illustrate how different the relative importance of the different cost components is, I have presented pie charts for a petrol driven car and the heaviest truck in Figure 5-39 and Figure 5-40.

---

[151]   Figures are updated to 2002 Euros applying the Norwegian consumer price index, and the NOK/Euro exchange rate of Januar 2nd 2002.

Cost components car (p), Rural areas

Cost components Truck>23 t (d), Rural areas
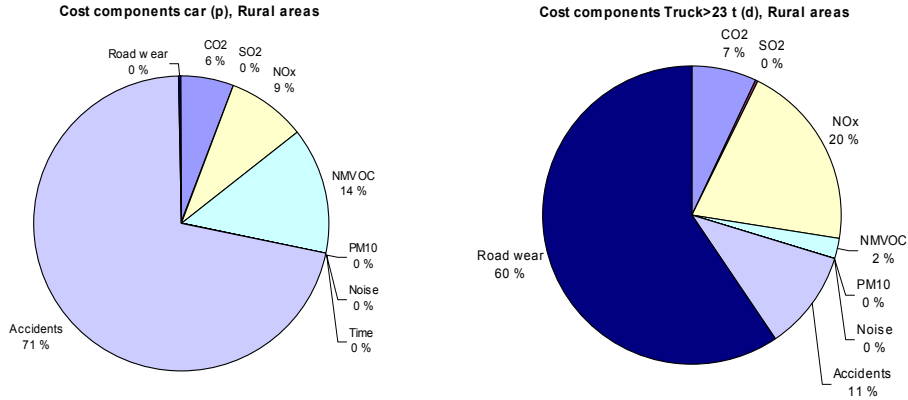
Figure 5-39   Cost components, Base scenario, Rural areas

This concludes the review of the Base scenario, but before I turn to the output
from the three other scenarios constructed, I will briefly present a small sensitivity
analysis on the Base scenario, trying to illustrate how dependent the output is on
some central individual input parameters.

Cost components Car (p), Cities

Cost components Truck>23 t (d), Cities
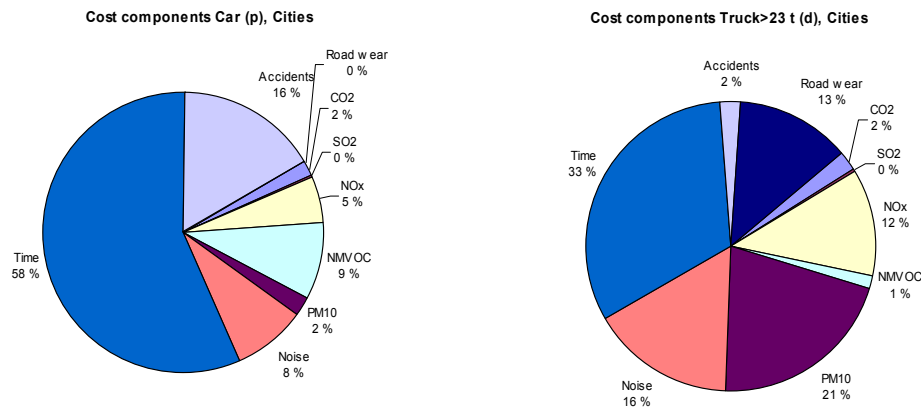
Figure 5-40   Cost components, Base scenario, Cities

5.4.3   *Sensitivity analysis of the Base Scenario*

Since the estimation of shadow prices for local emissions (VOE), statistical life
(VOSL) and the value of time (VOT) is complicated, there is a high level of

249

uncertainty about these figures.
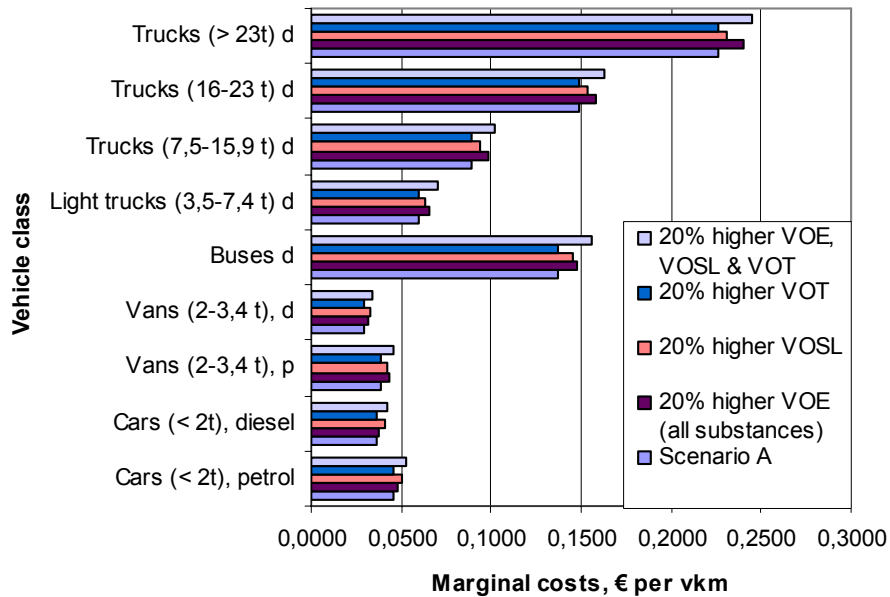
Figure 5-41 Sensitivities of VOT, VOSL and VOE, Rural areas

In Figure 5-41 and Figure 5-43, the impact of a 20% increase in these factor prices is analysed, first by adding 20% to each of the shadow price groups, and then by adding 20% to all these factor prices. The increase in VOE seems to give the highest increase for heavy vehicles, whereas an equivalent increase in VOSL seems to yield the most significant impact for cars (both for the rural and the city case).

In Figure 5-42 and Figure 5-44 the sensitivity to changes in fuel consumption, the magnitude of the maintenance costs and the response to a high CO_2 shadow price are illustrated. Naturally a 20% reduction in the fuel consumption figures yields the highest impact for heavy vehicles, since the reductions in air emissions would be biggest for these groups. This fuel consumption reduction causes a 4% drop in marginal costs for cars, and a 9% drop for the heaviest goods vehicles.

250
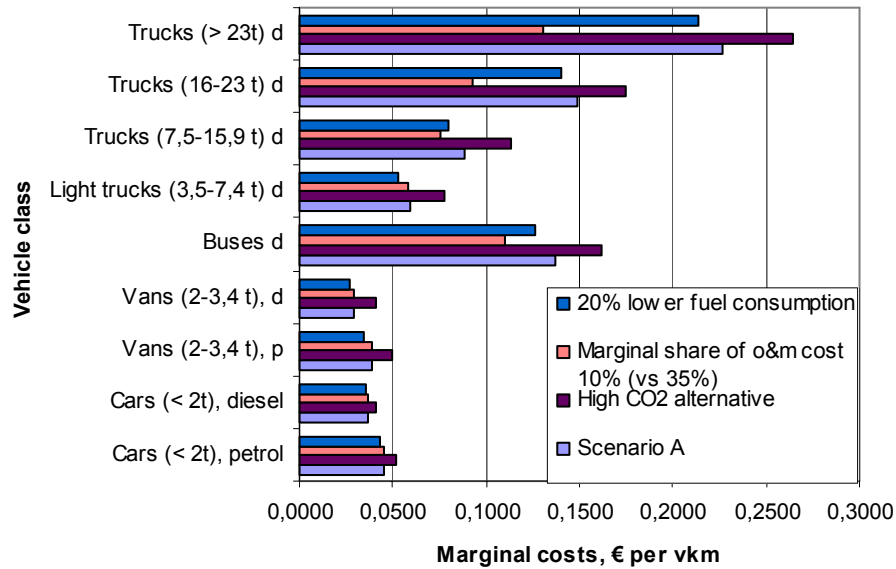
Figure 5-42   Sensitivities of fuel consumption, marginal o&m costs and shadow price of CO$_2$, Rural areas

In Eriksen et al. (1999) it is assumed that 35% of the total operation and maintenance costs could be considered marginal, i.e. traffic volume dependent. Since road wear costs are distributed by ESALs in the base scenario, a reduction of this share to 10% has an impact on the heavier vehicles only, yielding almost the same reduction in marginal costs for the heaviest vehicles as the 20% fuel reduction.

The relative impact of applying the Norwegian high CO$_2$ shadow price alternative (i.e. € 50 per tonne, instead of the low alternative € 15 per tonne), is biggest for the heavy vehicles in rural areas, because other emissions have a modest magnitude here. The heaviest trucks will get a marginal cost estimate that is
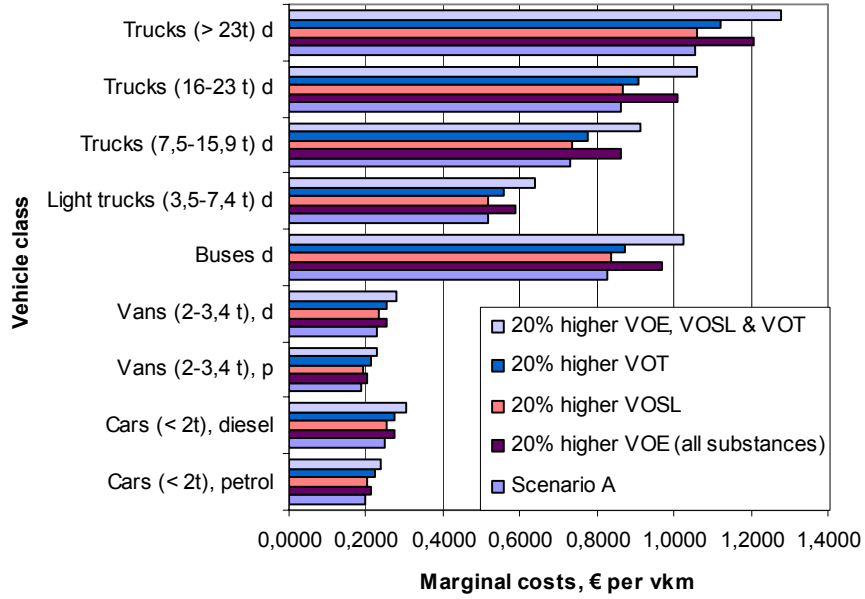
251

Figure 5-43   Sensitivities of VOE, VOSL and VOE, Cities
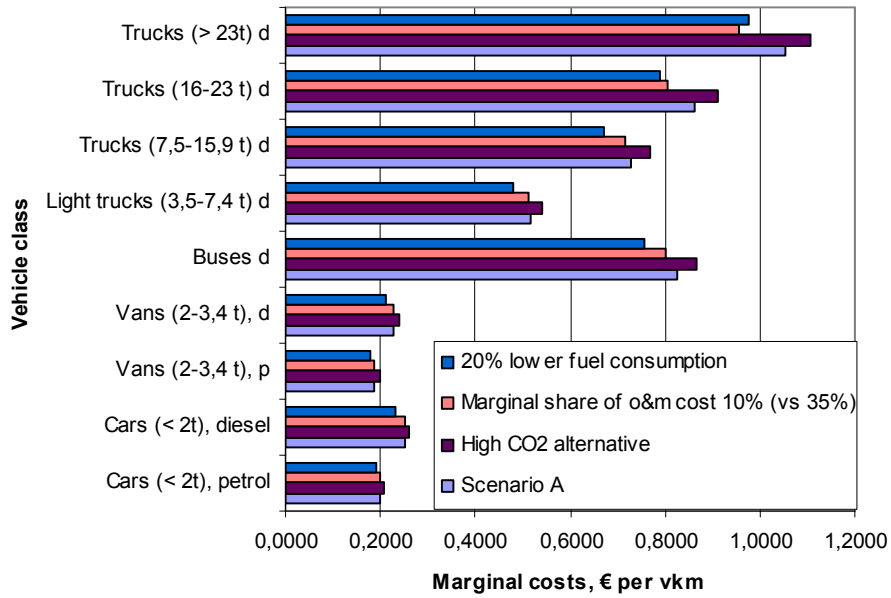
Figure 5-44   Sensitivities of fuel consumption, marginal o&m costs and shadow price of CO₂, Cities

17.5% higher with this price alternative in the rural case, whereas the same proportion for the city case is only 4.9%.

The alterations made in some of the most important input figures, do not necessarily illustrate the magnitude of uncertainty very well. It might be argued that changing the shadow prices of time, statistical life and environmental factors by 20% represents a very small change compared to the actual uncertainty connected to these figures.

Reducing overall fuel consumption figures by 20% may be feasible as energy efficiency keeps getting better. However, this far much of the improved engine technology has been offset by increasing average vehicle weight (mainly due to safety and comfort equipment added). Average consumption figures may therefore not fall as much as 20% in the near future unless more radical changes are made technologically (e.g. by introducing fuel cells) or behaviourally (e.g. by people buying smaller cars).

Having given some illustrations of sensitivities to changed input in the Base scenario, I now turn to presenting the output of the other scenarios.

5.4.4   Output from the FAMAROW Scenario

In the previous main section of this thesis, I have developed models that could be used for predicting average maintenance cost per vehicle kilometre. This is illustrated for the FAMAROW rut-depth model in section 4.5.3. The FAMAROW Scenario is identical to the Base scenario apart from the fact that I have substituted the road wear cost element originally based on a top-down approach to the bottom up approach based on the FAMAROW model. Not only does this yield a much lower cost allocated to overall traffic volume dependent costs, but the *distribution* of the costs is based on a flat per axle rate, rather than the original 2.5 power ESAL distribution applied in the TØI-model. As illustrated in Figure 5-45

this yields a quite significant effect, especially for the heavy vehicles, which now get significantly lower marginal road wear costs allocated.

Vehicle class	Rural		Towns		Cities	
	Base scen.	FAMAROW scen.	Base scen.	FAMAROW scen.	Base scen.	FAMAROW scen.
Cars (< 2t), petrol	0.046	0.047	0.082	0.083	0.199	0.200
Cars (< 2t), diesel	0.037	0.038	0.074	0.074	0.250	0.251
Vans (2-3,4 t), p	0.039	0.039	0.071	0.072	0.188	0.188
Vans (2-3,4 t), d	0.029	0.030	0.059	0.060	0.229	0.229
Buses d	0.137	0.100	0.429	0.392	0.827	0.790
Light trucks (3,5-7,4 t) d	0.059	0.058	0.203	0.202	0.515	0.514
Trucks (7,5-15,9 t) d	0.089	0.071	0.356	0.338	0.730	0.712
Trucks (16-23 t) d	0.149	0.071	0.453	0.375	0.863	0.786
Trucks (> 23t) d	0.227	0.094	0.520	0.387	1.054	0.921

Figure 5-45   Comparison of estimates in the Base Scenario and the FAMAROW scenario (2002 € per vkm)

Once again the most dramatic relative effect is for the heavy vehicles in rural areas, where overall marginal costs are reduced by 59%. In cities the corresponding effect is a reduction by 13%, due to the smaller relative importance of road wear in cities.

5.4.5   Output from the European Scenario

The European scenario contains alternative calculation prices for time, statistical life and CO_2-emissions. It is also based on a bottom-up approach to the estimation of road wear from a recent Swedish study. For the rural case, this yields slightly higher costs for cars, vans and light trucks, but significantly lower estimates for buses and HGVs. This is mainly due to a much lower road wear component than the ones used in the Norwegian top-down approach. For towns and cities the European scenario values almost consistently exceed those from the Base scenario. This is because the higher CO_2-cost and the higher value of time compensate the lower road wear costs. Output from the European scenario is presented in Figure 5-46.

254

Vehicle class	Rural		Towns		Cities	
	Base scen.	European scen.	Base scen.	European scen.	Base scen.	European scen.
Cars (< 2t), petrol	0.046	0.049	0.082	0.085	0.199	0.298
Cars (< 2t), diesel	0.037	0.040	0.074	0.078	0.250	0.350
Vans (2-3,4 t), p	0.039	0.042	0.071	0.074	0.188	0.287
Vans (2-3,4 t), d	0.029	0.032	0.059	0.062	0.229	0.328
Buses d	0.137	0.114	0.429	0.409	0.827	0.998
Light trucks (3,5-7,4 t) d	0.059	0.063	0.203	0.208	0.515	0.712
Trucks (7,5-15,9 t) d	0.089	0.080	0.356	0.349	0.730	0.915
Trucks (16-23 t) d	0.149	0.090	0.453	0.398	0.863	1.000
Trucks (> 23t) d	0.227	0.122	0.520	0.418	1.054	1.240

Figure 5-46   Comparison of estimates in the Base Scenario and the European scenario (2002 € per vkm)

5.4.6   Output from the Recommended Scenario

In Figure 5-47 I have finally presented the output from the Recommended scenario. Here I have combined some of the input alternatives from the previous scenarios based on the argumentation above. Basically this means that I have dropped the noise component, applied bottom-up road wear estimates based on the FAMAROW model, and applied the "European" value for CO_2-emissions.

Vehicle class	Rural		Towns		Cities	
	Base scen.	Recomm. scen.	Base scen.	Recomm. scen.	Base scen.	Recomm. scen.
Cars (< 2t), petrol	0.046	0.047	0.082	0.083	0.199	0.200
Cars (< 2t), diesel	0.037	0.038	0.074	0.074	0.250	0.251
Vans (2-3,4 t), p	0.039	0.039	0.071	0.072	0.188	0.188
Vans (2-3,4 t), d	0.029	0.030	0.059	0.060	0.229	0.229
Buses d	0.137	0.100	0.429	0.392	0.827	0.790
Light trucks (3,5-7,4 t) d	0.059	0.058	0.203	0.202	0.515	0.514
Trucks (7,5-15,9 t) d	0.089	0.071	0.356	0.338	0.730	0.712
Trucks (16-23 t) d	0.149	0.071	0.453	0.375	0.863	0.786
Trucks (> 23t) d	0.227	0.094	0.520	0.387	1.054	0.921

Figure 5-47   Comparison of estimates in the Base Scenario and the Recommended scenario (2002 € per vkm)

255

For cars this alternative is very similar to the Base scenario in all areas. The major differences arise for the buses and the heavy trucks, mainly because of lower allocated road wear costs, and (for towns and cities) the exclusion of noise costs. Noise costs is the major component in the Base scenario (see Figure 5-40).
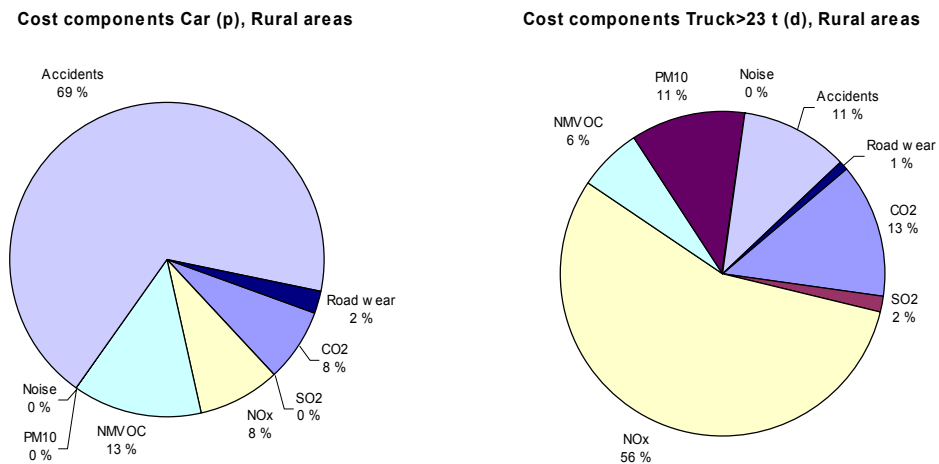
Figure 5-48   Cost components, Recommended scenario, Rural areas

The relative composition of the different cost components under the Recommended scenario is illustrated in Figure 5-48 and in Figure 5-49. In the Rural case, accident costs become very dominant for petrol-driven cars, constituting 69% of total marginal costs. The picture is very different for the heaviest goods vehicle. Here NO_x-emissions is the dominant cost component (56%).

In the city case, external time costs take over the dominating role for both example vehicles, constituting 61% and 45% of total marginal costs, respectively. For the heavy goods vehicle, particle emissions have also become a major cost component, representing almost one third of the total marginal costs.

256

Cost components Car (p), Cities

Cost components Truck>23 t (d), Cities
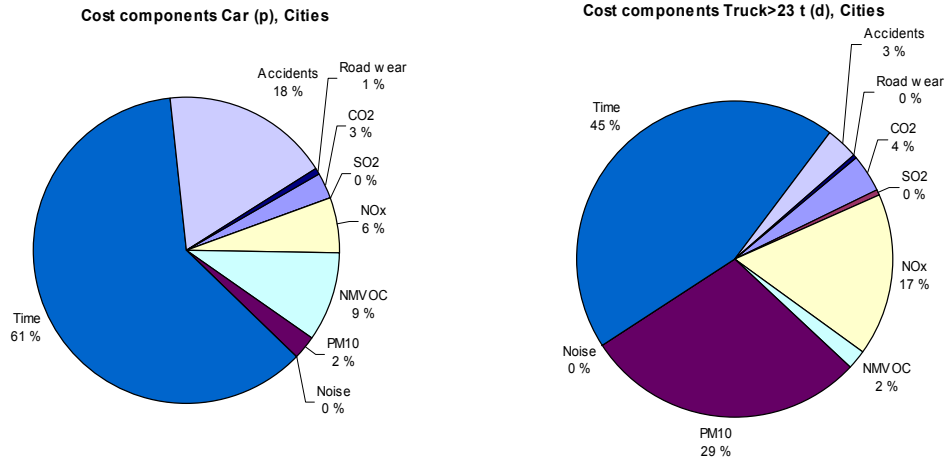
Figure 5-49   Cost components, Recommended scenario, Cities

5.4.7   *Comparing the alternative scenarios*

I have now presented four different scenarios, and related each of the latter three to the Base scenario. To complete the picture of how these scenarios compare to one another, I have produced three figures representing all scenarios in rural areas, towns, and cities respectively (Figure 5-50, Figure 5-51, and Figure 5-52).
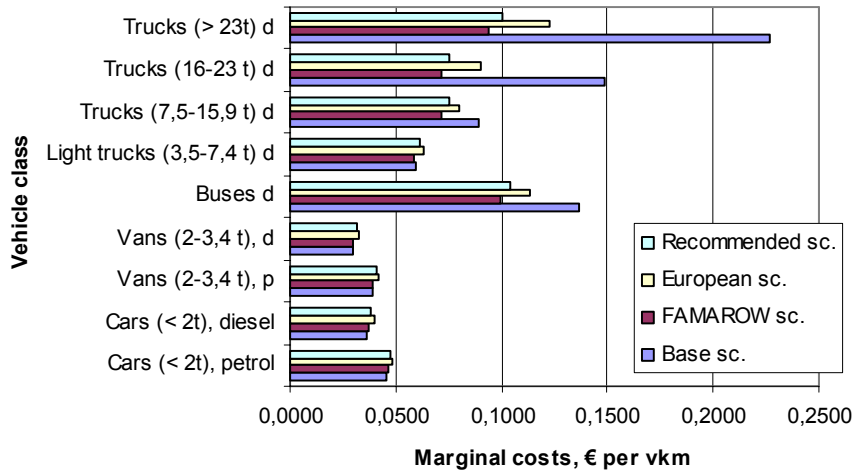
Figure 5-50   Comparison of the scenarios, Rural areas

The most striking difference between the scenarios is the magnitude and the distribution of the marginal road wear costs. The Base scenario has a top-down approach to the estimation of total marginal road wear costs, whereas the other three are based on bottom-up procedures. This means that the overall sum to be distributed is 6-7 times greater in the Base scenario, than in the other cases.
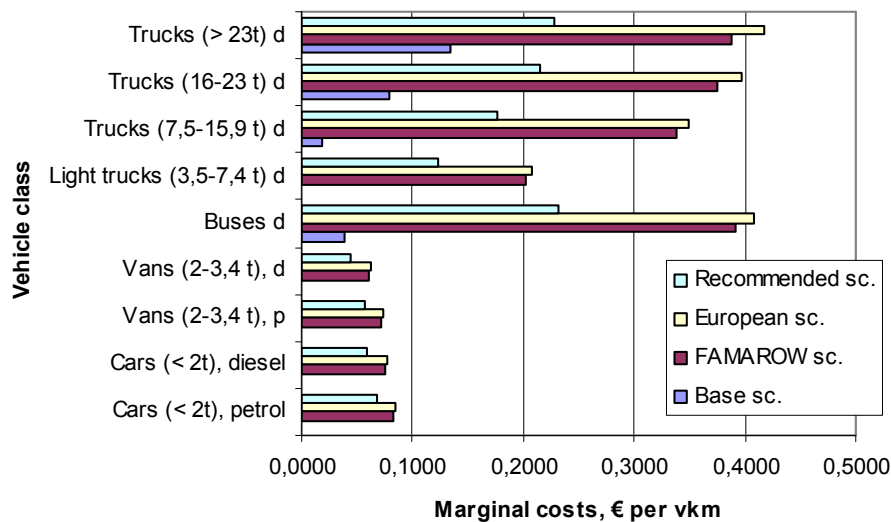
Figure 5-51   Comparison of the scenarios, Towns

The Base case and the European scenario distribute the road wear costs by the number of ESALs each vehicle class represents, whereas the other two scenarios are based on a flat per axle cost estimate. This means that road wear costs are allocated mainly to the heavy vehicles in the Base and European scenario, but are more evenly distributed in the FAMAROW-based scenarios.

The differences between the four scenarios are not the same for the three "locations". In the Rural case the Base scenario stands out as yielding significantly higher estimates. In the Towns case the FAMAROW and the European scenarios yield the highest estimates, especially for heavy vehicles. In

the city case, however, it is the European scenario that yields the highest estimates. Here the Recommended scenario is the one that represents the lowest estimates, mainly due to the exclusion of the noise costs.
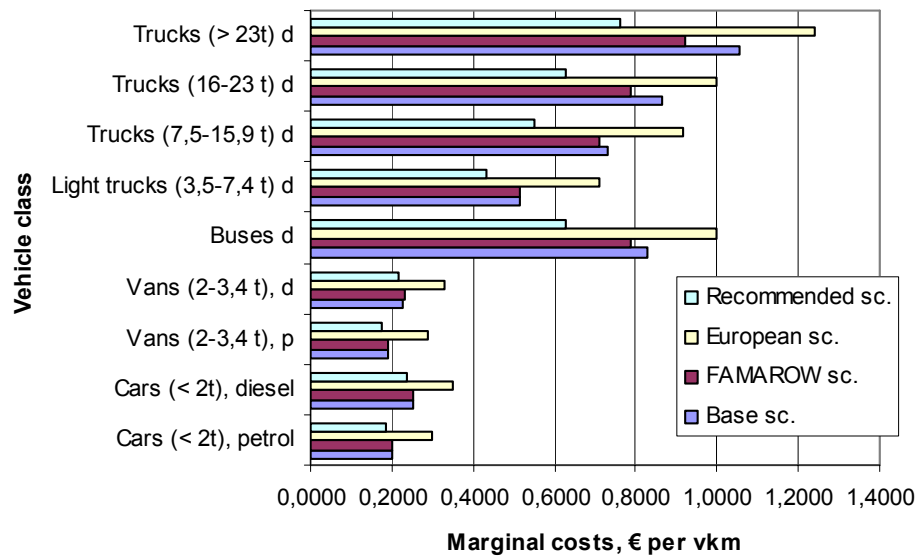
Figure 5-52   Comparison of the scenarios, Cities

5.4.8   Comparing the Recommended scenario to other empirical research results

Since marginal cost estimates are highly dependent on prevailing traffic and population characteristics, there is no reason to expect that comparisons between countries should coincide. However, such comparisons may still be of some interest because they could trigger important research questions stemming from a need to explain any observed discrepancies between the estimates.

Many countries have conducted studies that have ended up with estimates of the marginal costs of road use. I have chosen three examples here: The USA, Ireland and Sweden. The first one (USA) is chosen merely because the methodology

applied in the Federal Highway Cost Allocation Study (FHCAS) is somewhat different from the typical European studies (see section 3.2.3 for a close description of this study). The latter two (Ireland and Sweden) are chosen because both these countries have some similarities to Norway. These are all countries with fairly extensive rural road networks covering sparsely populated areas, and towns and cities of quite moderate sizes.

Vehicle class / Highway class	Pave-ment	Con-gestion	Crash	Air Pollution	Noise	Total
Autos / Rural interstate	0.000	0.006	0.007	0.008	0.000	0.021
Autos / Urban interstate	0.001	0.055	0.008	0.009	0.001	0.074
40 kip 4-axle S.U. truck / Rural Interstate	0.007	0.017	0.003	0.027	0.001	0.056
40 kip 4-axle S.U. truck / Urban Interstate	0.022	0.174	0.006	0.032	0.011	0.245
60 kip 4-axle S.U. truck / Rural Interstate	0.040	0.023	0.003	0.027	0.001	0.094
60 kip 4-axle S.U. truck / Urban Interstate	0.129	0.232	0.006	0.032	0.012	0.410
60 kip 5-axle Comb. truck / Rural Interstate	0.023	0.013	0.006	0.027	0.001	0.072
60 kip 5-axle Comb. truck / Urban Interstate	0.075	0.131	0.008	0.032	0.020	0.265
80 kip 5-axle Comb. truck / Rural Interstate	0.090	0.016	0.006	0.027	0.001	0.141
80 kip 5-axle Comb. truck / Urban Interstate	0.291	0.142	0.008	0.032	0.022	0.495

Figure 5-53   Marginal cost estimates from the US Federal Highway Cost Allocation Study (US Department of Transportation 2000), converted into 2002 € per vkm by the author

In Figure 5-53 I have converted the figures from the FHCAS into 2002 Euros. Vehicle classes are not directly comparable to the ones used in my study since average car size will typically be bigger in the USA, and the truck categories are not directly comparable to the ones I have chosen. The figures presented are also representative of rural and urban *interstate highways*, which means that there are no figures representing city traffic in general. The major discrepancy compared to the other studies is, however, the fact that the US emission figures do not contain

260

CO_2-emissions. This is a very significant part of the other studies (e.g. between 38% for HGVs and 65% for cars in the Swedish figures).

In the same manner, I have also converted the figures from Ireland in Figure 5-54. Here only a few vehicle classes are presented, but prices are differentiated by time of day (i.e. peak and off peak).

Vehicle class / area / time of day	Con-gestion	Air pollution	Noise	Acci-dents	Road damage	Total
Car (p), Ireland, Off peak	0.033	0.008	0.000	0.053	0.000	0.094
Car (p), Dublin, Peak hour	0.455	0.009	0.001	0.033	0.000	0.498
Car (p), Dublin, Off peak	0.045	0.008	0.001	0.033	0.000	0.087
Bus (d), Dublin, Peak hour	0.912	0.025	0.005	0.024	0.000	0.966
Truck (d), Ireland, Off peak	0.067	0.121	0.000	0.064	0.016	0.269

Figure 5-54   Marginal cost estimates for Ireland and Dublin (De Borger and Proost 2001), converted into 2002 Euros per vkm by the author

Finally I have also converted the Swedish figures into 2002 Euros in Figure 5-55.

Vehicle class / area	Road wear	Local emissions	Noise	Acci-dents	CO2 (150SEK/t)	Total
Car (p) with cat., Rural areas	0.001	0.002	0.001	0.010	0.025	0.039
Car (p) wo cat., Rural areas	0.001	0.023	0.001	0.010	0.027	0.062
Car (d) with cat., Rural areas	0.001	0.002	0.001	0.010	0.021	0.035
Car (d) wo cat., Rural areas	0.001	0.005	0.005	0.010	0.031	0.052
Truck (d), 3.5-16 t, Rural areas	0.003	0.030	0.005	0.026	0.077	0.140
Truck (d), >16t, Rural areas	0.006	0.061	0.011	0.026	0.152	0.256
Car (p) with cat., Town/City	0.001	0.009	0.007	0.019	0.038	0.073
Car (p) wo cat., Town/City	0.001	0.056	0.007	0.019	0.040	0.122
Car (d) with cat., Town/City	0.001	0.018	0.007	0.019	0.029	0.073
Car (d) wo cat., Town/City	0.001	0.098	0.007	0.019	0.036	0.160
Truck (d), 3.5-16 t, Town/City	0.003	0.089	0.045	0.046	0.071	0.254
Truck (d), >16t, Town/City	0.006	0.140	0.102	0.046	0.176	0.469

Figure 5-55   Marginal cost estimates for Sweden (Hesselborn 2001), converted into 2002 Euros per vkm by the author

As vehicle classes and operation area and traffic conditions are not directly comparable in these four studies, collecting the figures into one table is rather complicated. There are, however, a few estimates that match each other fairly well, and I have tried to summarise these in Figure 5-56. Comparisons should of course be made bearing in mind that the categories are not strictly identical.

Vehicle class	Rural				Cities			
	CATERU Recomm.	US FHCAS	Ireland & Dublin	Sweden [152]	CATERU Recomm.	US FHCAS [153]	Ireland & Dublin	Sweden
Cars (< 2t), petrol	0.047	0.021	0.094	0.051	0.200	0.074	0.498	0.098
Cars (< 2t), diesel	0.038	na	na	0.044	0.251	na	na	0.117
Vans (2-3,4 t), p	0.039	na	na	na	0.188	na	na	na
Vans (2-3,4 t), d	0.030	na	na	na	0.229	na	na	na
Buses d	0.100	na	na	na	0.790	na	0.966	na
Light trucks (3,5-7,4 t) d	0.058	na	na	0.140	0.514	na	na	0.254
Trucks (7,5-15,9 t) d	0.071	na	na		0.712	na	na	
Trucks (16-23 t) d	0.071	0.056	0.269	0.256	0.786	0.245	na	0.469
Trucks (> 23t) d	0.094	0.072	na		0.921	0.265	na	

Figure 5-56   Comparison of estimates in the Base Scenario and the Recommended scenario (2002 € per vkm)[154]

The only category that contains information from all studies is the petrol-driven *cars*. For the rural case, it seems that there is a fairly good match between the Norwegian and the Swedish figures, whereas the US figures are significantly lower, and the Irish figures are significantly higher than the Nordic estimates. The picture is somewhat different for cities or urban areas. Here the CATERU figures are more than twice as big as the figures from Sweden and the USA. Once again the Irish figures are by far the highest ones. The Norwegian figures presented here are for major cities (i.e. Oslo, Bergen and Trondheim), whereas the Swedish

---

[152]   For Sweden the average of cat/ no cat figures are applied.

[153]   Note that City figures from the US FHCAS actually represent Urban interstates.

[154]   Note that vehicle classes are not the same for the different studies. Figures from the foreign studies are located according to the judgement of the author. See the tables above for correct vehicle classes for each individual study.

figures represent both towns and cities. The USA figures also represent urban Interstates, and not city traffic in general. Possibly, the Norwegian figures for Towns would have been more relevant here. Then the marginal cost for petrol-driven cars would be € 0.083 instead of € 0.200 per vkm, which complies much better with the American and the Swedish figures.

The figures for *buses* in cities in Norway and Ireland are not very different, bearing in mind that the figures for Dublin concern peak traffic, whereas the Norwegian ones are averages.

Only Norwegian and Swedish figures are comparable for the lighter trucks (<16t). Here the Swedish figures exceed the Norwegian ones significantly for the rural case, whereas the situation is the opposite for the city case. Once again figures for Norwegian *towns* match the Swedish ones very well.

For the *heaviest vehicles* (>16t) it seems that Norwegian and American figures match very well in the rural case, whereas Irish and Swedish figures are dramatically higher. This difference between the Norwegian and the Swedish figures may be partly explained by the fact that the Swedish figures contain road wear costs that are allocated according to ESALs, which means that heavy vehicles are allocated a higher proportion of these costs. For the city case, Norwegian figures are much higher than the American ones (which still are not truly city traffic figures), and also somewhat higher than the Swedish estimates. The Norwegian figures for *towns* are almost halfway between the American and the Swedish estimates.

Summing up, it seems that the American figures lie at the lower end of the scale, which is not surprising, considering that these figures are exclusive of CO_2-emissions. Adding 50% to the American figures would in many cases bring them into the area of the Nordic estimates. The Irish estimates seem, however, to be

almost consistently higher than the Nordic figures. The presentation of the Irish underlying studies is not thorough enough in (De Borger and Proost 2001) to enable a closer analysis of the reasons for the Irish figures being so comparatively high.

Preliminary results from the UNITE project

The final results of the UNITE project (see section 3.3.3) have not been published yet (February 2003). However some preliminary results have been presented in Nash and Johnson (2002), and these are reproduced in Figure 5-57 and Figure 5-58.

Overview of MC for Car Travel (Euro per vkm)
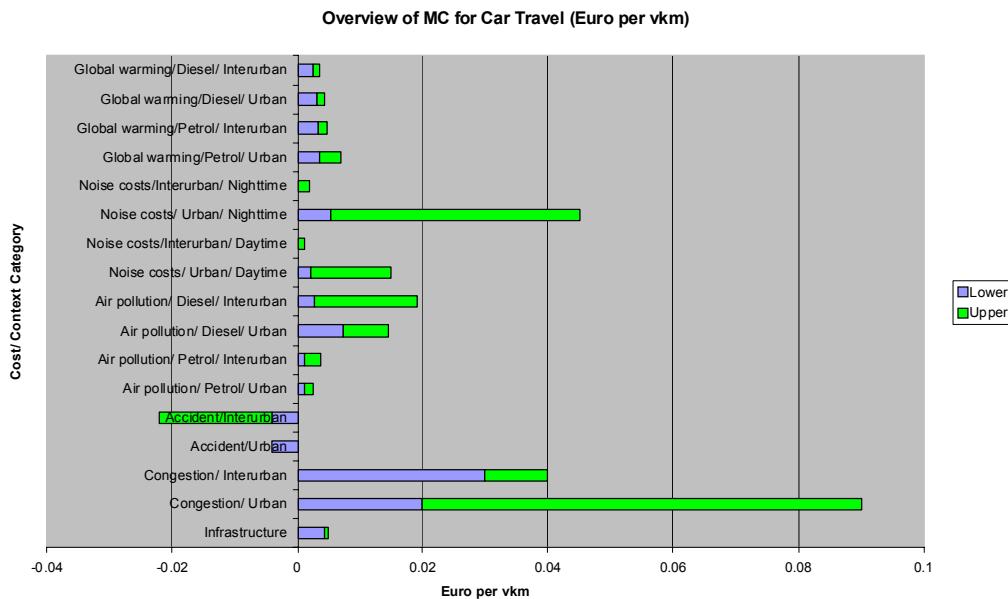
Figure 5-57   Overview of marginal costs for car travel (€ per vkm). Source: Nash and Johnson (2002)

Not knowing anything about the background for the various estimates, it is hard to relate them to my findings. Such a comparison will have to wait until the final results are well reported[155]. However, it is possible to give some general

---

[155]   All figures in the subsequent paragraph are based on interpretations of the graphs presented in op.cit., as I do not have access to the actual figures.

264

comments based on the figures. Compared to my findings, where infrastructure costs (at least for cars) seem to be very low, it seems that the numbers reported here are significantly higher than my findings. Congestion estimates do, not surprisingly, vary considerably. However, none of the estimates seem to be as high as my estimates for HGVs in cities, which are in the area of € 0.2 and € 0.3 per vkm.
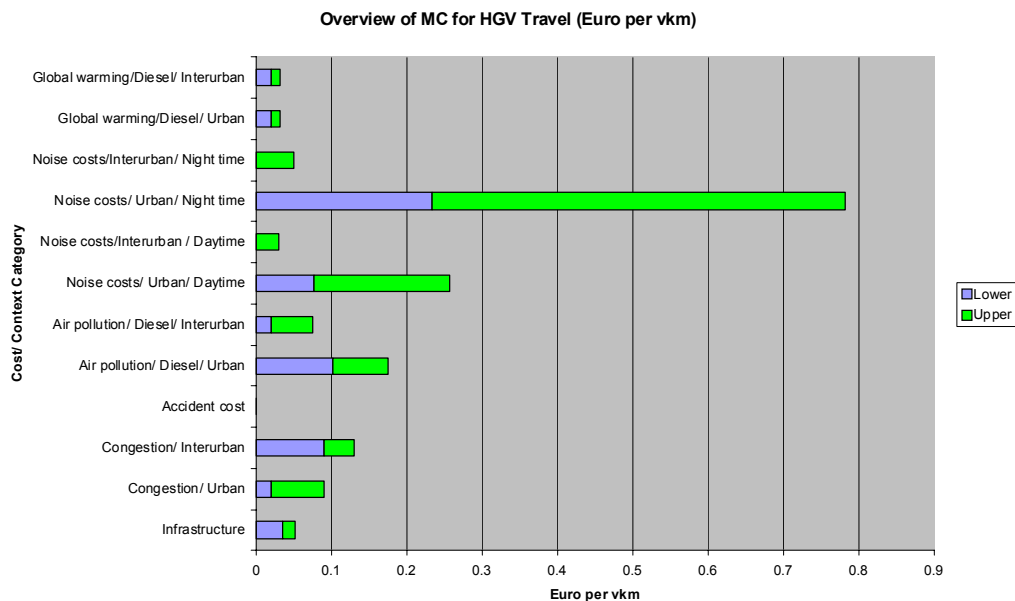
Figure 5-58   Overview of marginal costs for HGV travel (€ per vkm). Source: Nash and Johnson (2002)

Estimated external accident costs are reported to be zero, or even negative. This is of course very much in contrast to my estimates, which are really not marginal, but average figures. The costs of local air pollution vary between € 0.002 per vkm (petrol car/urban) to € 0.17 per vkm (HGV/urban). This compares to the range of € 0.029 per vkm (petrol car/urban) to € 0.366 per vkm (truck>23t/urban) in my findings. Especially, noise externalities at nighttime seem to have substantial magnitudes in the UNITE figures. In my calculations noise costs are left out because *marginal* noise costs probably are low on already heavily trafficked roads. However, marginal levels could be very significant if the initial level of

265

traffic is low (as it may very well be at night), and when population densities are high.. This will, however be dependent on the prevailing local conditions. Finally, the figures for global warming effects range from € 0.03 per vkm (interurban car) to € 0.3 per vkm (HGVs) in the UNITE figures. My corresponding estimates are € 0.004 per vkm for the interurban car, and € 0.031 per vkm for the urban HGV. UNITE figures thus seem to be ten times higher than my estimates. Since there is little argument about the amount of CO_2-emissions per vkm, this very significant discrepancy must be due to the use of other shadow prices for such emissions.

Consensus on principles and estimates of marginal costs in Europe?

Gunnar Lindberg (Lindberg 2002a) has reviewed the possibility of reaching consensus on the marginal cost estimates on the European level. His conclusions are summarised in Figure 5-59. According to Lindberg there is still a need for more studies of infrastructure costs, accident costs and on the costs of greenhouse gas emissions (possibly also on local air pollution), whereas there seems to be both principal and operational consensus concerning the marginal external costs of congestion.

Category	Costs/valuations	External	Marginal
Infrastructure	**No** Discussion on expenditure/costs.	**Yes**. All costs external	**No** Studies not conclusive
Accidents	**Yes** WTP approach used. Actual level is uncertain.	**No** Too few studies.	**No** Too few studies.
Congestion	**Yes**	**Yes**	**Yes**
Air pollution	**Yes** Same uncertainty as for accidents.	**Yes** All costs external.	**?** Too few studies.
Greenhouse gases	**No**	**Yes**	**No**

Figure 5-59   Consensus on marginal cost estimates (Lindberg 2002a)

266

Having presented different scenarios for the CATERU-model, and comparing the output from the Recommended scenario to other international studies, I will now turn to a comparison of the results from the Recommended scenario to the current marginal taxation level in Norway.

5.5 COMPARING ESTIMATED EXTERNALITIES TO THE CURRENT TAXATION LEVEL

The purpose of estimating the marginal costs of road use in this thesis is ultimately to provide an improved empirical foundation for Norwegian road user charges. Marginal taxes connected to road use in Norway are almost equivalent to *fuel taxes*. Some years ago HGVs had a kilometre-tax instead of the fuel tax, but this was substituted by diesel-taxes and a system of annual weight differentiated taxes (which are not truly marginal). Extensions of the Norwegian road network are also partially financed by tolls, but the tolls are not charged by the kilometre, and hence these tolls are not truly marginal either.

	€ per litre
Fuel taxes:	
Petrol tax (unleaded)	0.48
Diesel tax, Low sulphur	0.35
Diesel tax, High sulphur	0.39
General marginal charges:	
CO2-tax petrol	0.09
CO2-tax diesel	0.06

Figure 5-60   Norwegian taxes on fuel 2002, € per litre fuel[156]

In Figure 5-60 the current fuel excise duties are summarised (converted into Euros). For petrol there is a basic fuel tax, and on top of that there is a so-called CO_2-tax, totalling € 0.57 per litre petrol. For diesel the tax is differentiated by the sulphur content, where a content of less than 50ppm yields a reduction in the tax.

---

[156]   Source: Norwegian Ministry of Finance

There is also a lower CO_2-tax on diesel. Taxation on low sulphur diesel therefore totals € 0.41 per litre, and high sulphur diesel € 0.45 per litre.

In section 5.2.4 I have argued that some part of the very significant vehicle registration taxes should be considered marginal as well, because a proportion of the vehicle depreciation tends to be dependent on the kilometrage of the vehicle. When calculating the marginal tax level, I have therefore included a small marginal part of vehicle registration taxes (relevant for cars and vans only) in the figures for marginal taxation levels, based on the illustrative analysis presented in 5.2.4. For cars this represents between 11% and 15% of marginal taxes.
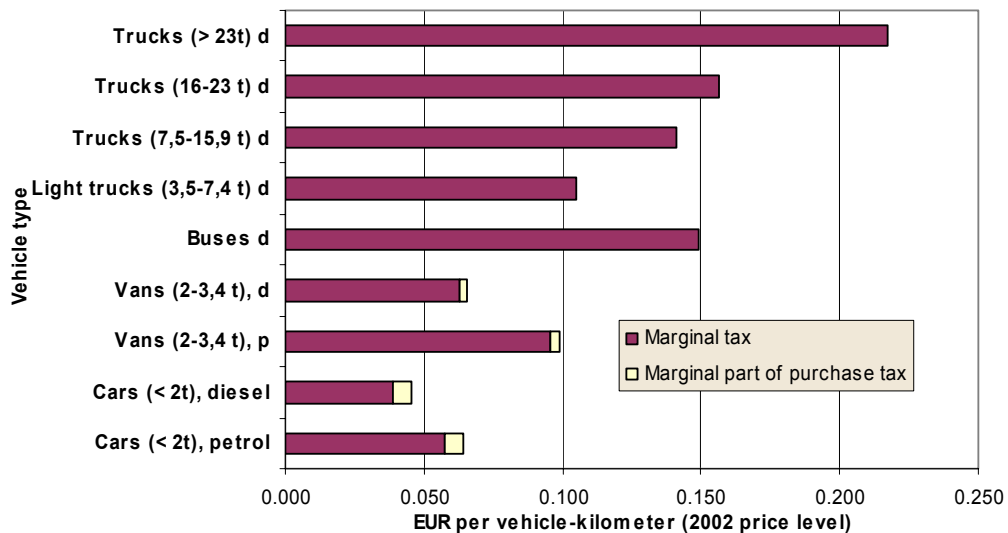
Figure 5-61   Current (2002) marginal taxation levels in Norway

By applying average figures for fuel consumption by vehicle class, marginal taxation levels for Norway are calculated and presented in Figure 5-61. For cars and vans, diesel driven vehicles are taxed much less than the corresponding petrol driven ones. This is partly due to lower taxes on diesel fuels, and partly due to a comparatively lower fuel consumption of diesel vehicles.

268

In Figure 5-62 the marginal cost estimates from the CATERU Recommended scenario is plotted against the current marginal taxation level for all vehicle types and locations. In Figure 5-63 the same figures are used for calculating the corresponding cost to tax ratios.
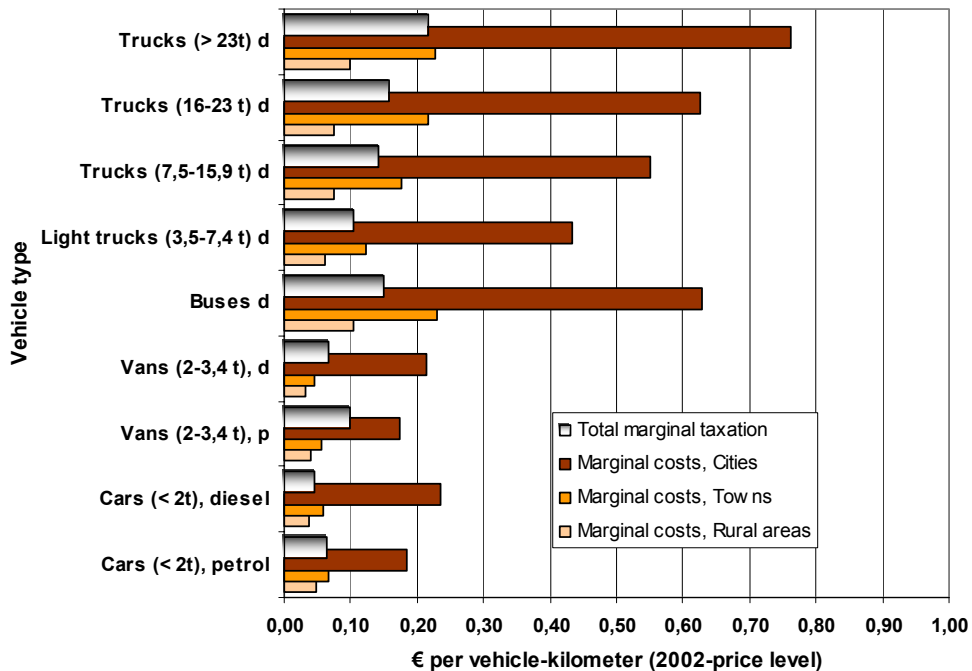
Figure 5-62   Marginal taxation compared to CATERU Recommended scenario output figures

For rural areas it seems that the current taxation level exceeds the calculated marginal costs for all vehicle classes. For towns, petrol driven cars and the heaviest vehicles seem to pay approximately according to estimated marginal costs. Diesel cars, buses and light and medium trucks seem to pay too little in this case, whereas it seems that vans pay too much. For cities, none of the vehicle categories get close to paying the estimated marginal costs. The relative discrepancy is highest for diesel cars and lowest for petrol-driven vans.

	Rural	Towns	Cities
Cars (< 2t), petrol	0.74	1.05	2.86
Cars (< 2t), diesel	0.84	1.31	5.18
Vans (2-3,4 t), p	0.42	0.57	1.76
Vans (2-3,4 t), d	0.48	0.68	3.26
Buses d	0.70	1.55	4.21
Light trucks (3,5-7,4 t) d	0.59	1.17	4.14
Trucks (7,5-15,9 t) d	0.53	1.25	3.90
Trucks (16-23 t) d	0.48	1.38	4.00
Trucks (> 23t) d	0.46	1.05	3.50

**Figure 5-63   Calculated marginal cost to tax ratios, CATERU
Recommended scenario**

6 CONCLUSIONS: A NEW PRICING REGIME FOR ROAD USE IN NORWAY?

In the preceding chapter I presented new estimates of the marginal costs of road use for Norway. Based on this new empirical foundation, I will in this chapter suggest a new pricing regime for road use in Norway.

6.1 CHOOSING THE RIGHT INSTRUMENTS FOR CHARGING

The choice of appropriate charging instruments for the different cost components is discussed in section 1.2.3, and my recommendations are summarized in Figure 1-3. The system used today, with one-sided focus on fuel taxes as the only truly marginal taxing instrument, is not really appropriate because it does not reflect true marginal costs well enough. Only the costs related to CO_2-emissions could be well reflected through a fuel tax.

6.1.1 Charging instruments for road wear

My empirical results based on the FAMAROW rut-depth model may indicate that the major maintenance triggering damage mechanism could be related to the use of studded tyres. At least, this is a plausible assumption on high volume roads. The FAMAROW roughness model does, however, give an indication that there may be some impact from heavy axles on the low volume (i.e. lower standard) roads. The estimated traffic volume depended wear was however very small. My sample of roads does probably not represent the weaker parts of the Norwegian road network very well, so there may still be a case for charging heavy vehicles on weaker roads. This is supported by the estimated models for Sweden (see (Lindberg 2002b), where marginal costs are significantly higher for the weaker roads.

The TØI-model used a top-down approach to estimating marginal road wear. I

have shown that both my bottom-up results and the equivalent Swedish estimates yield a much lower level of marginal road wear costs. I have also shown that if these costs are distributed by vehicle axles and not ESALs, road wear costs become rather insignificant compared to the other cost components for all vehicle classes. The Swedish study (op.cit.) does not provide any explicit empirical argumentation for applying ESALs as the criterion for allocating road wear costs, but there may be very good theoretical arguments for doing that when studying cracking (rather than rut-depths and rutting as I have done). A central question may therefore be what kind of damaging mechanism is the most important one for triggering maintenance actions. This is an issue that I have not been able to address properly in my study. The situation may very well be different in Norway and Sweden on this issue, because there is a much more extensive use of studded tyres in Norway (Jacobsen and Hornwall 1999). Rutting is therefore considered to be the main overlay-triggering mechanism in Norway, at least on high volume roads[157].

Currently, a charging system that differentiates between road classes is not feasible. In the meantime road wear should be charged by a *kilometre tax* rather than a fuel tax. If charges should be based on the CATERU Recommended scenario, a flat per axle charge is appropriate, possibly only levied on cars and vans because heavier vehicles rarely use studded tyres. However, if the Swedish test results are representative of (weak) Norwegian roads, there may also be a case for an ESAL-based component which will increase the road wear component for heavy vehicles.

Could levying a purchase charge on studded tyres be an alternative?
Since attributing a significant part of marginal road wear to the use of studded tyres is a plausible interpretation of my model development, a natural suggestion

---

[157]   In the past 2-3 years a fee for driving with studded tyres has been introduced in Oslo and Trondheim. This has resulted in a very low percentage of cars using studded tyres in these areas, hence one may expect rutting to become a less important factor here.

would be to link the charge closer to the actual use of such tyres. The closest feasible option would then be to levy *a charge on the purchase of studded tyres*. Such a charge could be calculated by multiplying the estimated marginal road wear cost per studded tyre-km by the expected life of the tyre. However, if the use of studded tyres reduces accident risks, then there will also be *positive* externalities connected to the studded tyre use. Studded tyres also contribute to some health problems connected to the fine graded road dust particles produced from such road wear. The net, tax-relevant, externality from the use of studded tyres is therefore not easy to estimate.

Carlsson et al. (1995) make an effort to evaluate the net social benefits of banning studded tyres. Fosser and Sætermo (1995) deal with differences in the probability of having an accident with and without studded tyres. The first study concludes (with a high level of uncertainty) that the positive effects related to less road wear, lower operating costs and lower environmental costs exceed the value of the calculated risk increase. The latter study, however, concludes that there is no significant difference in the risk of having an accident depending on the use of studded tyres. Both these studies are some years old, and may not be representative of the current situation. Technologically there have been significant changes both in tyre and road constructions over the past decade. The weight of the studs has been reduced, and the alternative winter-tyres have been improved. The first factor reduces road wear, the latter reduces the potential risk differences of having an accident. Furthermore, road pavements have been developed, and applied, that better withstand the use of studded tyres. It seems, therefore, that there is a need for further research before the rationale for a purchase tax could be established. Anyway, a purchase tax will always be a second best option, as the first best solution would be to charge according to actual *use*.

6.1.2   *Charging instruments for external time costs*
In the TØI model, and in the CATERU Recommended scenario, time costs are

entered into the City case. These estimates are based on *average* external time costs on the road networks of some major Norwegian cities. Proper marginal cost pricing of congestion costs must include a differentiation by time of day. A fuel charge is therefore *not* appropriate for covering these costs.

A full *electronic road pricing* scheme that could capture the differences by time and region would probably be the best solution in the long run. Awaiting the suitable technological solutions, a time differentiated cordon pricing scheme (like the one just introduced in London) could function as an intermediate crude form of road pricing (Grue et al. 1997).

However, though marginal time costs actually will vary minute by minute and kilometre by kilometre, a road pricing scheme would never be designed to vary charges this much. The rationale behind the use of Pigouvian taxes is that they should represent incentives for adjusting behaviour according to what is socially desirable. This means that the price information must be available and transparent at the time of decision. In the case of road pricing this would mean that price of choosing to use the car alternative to public transport must be clear when this decision is made. Another factor limiting the level of detail in the design of congestion charges, is of course the costs connected to designing and operating such a system. This means that even a future more advanced system (GPS based?) must be rather simple and transparent to the user in order to achieve its goal of adjusting behaviour. The development of such a system is technically feasible already, but there are many political, legal and practical obstacles to be overcome before such a system can be put to work. In the Netherlands a system for paying by the kilometre is under development, and this includes a so-called *mobimeter*, an advanced kilometre counter with many possible applications. The Dutch system will probably be implemented by 2006 (Teule 2002).

274

6.1.3   Charging instruments for air-pollution

The widely used *fuel tax* is only considered suitable for internalising *global*, and to some extent (if differentiated) *regional emissions*. An alternative to the latter could be an area specific kilometre tax.

Local emissions and noise have to be dealt with by applying some form of an area-specific and vehicle specific tax instrument. A *kilometre-tax* dependent on these factors is recommended for the future, possibly integrated with an advanced congestion charging scheme, as these effects are most significant in urban areas. A first step towards this future area-dependent solution, could be to levy this part of the cost responsibility on a kilometre-tax only differentiated by vehicle type (fuel type, with/without catalytic converter, with/without particle trap, engine size/fuel consumption).

6.1.4   Charging instruments for external accident costs

Expected accident costs probably vary considerably by vehicle types, but there is generally no reason for assuming that these costs vary according to fuel consumption, hence neither this cost component seems very suitable for a fuel excise duty.

The CATERU estimates, based on the updated Elvik model, actually allocate lower marginal accident cost to heavy vehicles than to cars. This is contradictory to the previous TØI-model estimates, and also to the prevailing Swedish figures (Hesselborn 2001). However, a recent study (Lindberg 2001a) of HGV accidents in Sweden gives a more complicated picture, where some vehicle classes even get negative marginal accident cost estimates. This study concludes with average marginal costs that are substantially lower than the last official Swedish figures. There seems to be a very strong need for further research into *how* marginal external accident costs vary. This need is accentuated by the very significant magnitude of the prevailing estimates of external accident costs compared to other cost components.

275

The insurance market already has a system for pricing motorists according to risk differences, which might be a suitable instrument for charging for external accident costs. Lindberg (2001b) has - in a study of Swedish accident costs - proposed a system for charging the external accident costs through a more detailed insurance system. If insurance systems become more directly dependent on factual risks and how these vary with distance, location and vehicle types, these systems will probably be well fit for including an external accident cost element as well.

Making insurance premiums dependent on factual risk difference by vehicle type and driver characteristics should not be very difficult. However, such a differentiation will only have limited effects on behaviour. The only effect one might achieve is that people would appreciate safety equipment more when purchasing the vehicle. To obtain a higher awareness of risk differences dependent on time and location, will require a fairly sophisticated technological system based on the actual positioning of the vehicle.

6.2 A NEW PRICING SCENARIO FOR NORWAY

6.2.1 Recommended new structure

Having reviewed appropriate charging instruments, my recommendation is that a new *short term*[158] *pricing scenario* for Norway is based on

1. A weight-distance tax (i.e. a vehicle differentiated km-tax) to cover road wear costs.
2. A time-differentiated cordon tax to cover rush-hour congestion costs in the major cities.
3. Maintaining the CO_2-tax element on fuels, but harmonizing the tax

---

[158] By "short term" I mean a scenario that should be feasible to implement within a 4-5 year time horizon.

276

element making the price per kg CO_2 equal for petrol and diesel.

4. Including local air emission externalities in the weight-distance tax, also differentiating it by vehicle specific emission characteristics (i.e. fuel type, catalytic converters, particle traps etc.).

5. Charging for external accident costs through a mandatory insurance system, differentiated by vehicle type (safety equipment etc.).

The most important change may also be among the easiest one to implement technologically: Making the existing cordon tax systems differentiated by time of day. The highest discrepancy between marginal costs and current taxes occurs during the city rush hours. Harmonizing the CO_2 fuel tax is technologically very easy. However, introducing a weight-distance tax also differentiated by vehicle emission characteristics, *does* require new technology being installed in the vehicles. The EU has decided that all new HGVs must have an electronic tachometer installed by 2004. However, this does not include lighter vehicles, and it is not mandatory for older vehicles to install one. This means that requiring such instruments installed for all vehicles will take some time. However, it should be feasible to introduce some sort of weight-distance tax for HGVs in the near future. Awaiting such instrumentation for cars and vans, it may be necessary to retain a fuel excise duty for covering local air emissions and road wear for some time yet. Charging vehicle type differentiated external accident costs through the insurance system is not technically difficult, but the empirical evidence necessary to discriminate between vehicle types may not be good enough for implementation. Awaiting better empirical evidence, the current charging regime, in which average accident costs are included in the fuel excise duty, may be the only feasible alternative. However, this does not provide any other incentive than reducing the overall demand for road use, and *not* an incentive for reducing risky behaviour.

Summing up, this means that the *immediate new pricing scenario* will be based on:

1. Keeping fuel excise duties that include the costs related to road wear, local air emissions and accidents, awaiting more suitable technology for charging by the kilometre and better empirical evidence for accident risk variations.

2. A time-differentiated cordon tax to cover rush-hour congestion costs in the major cities.

3. Maintaining the CO_2-tax element on fuels, but harmonizing the tax element making the price per kg CO_2 equal for petrol and diesel.

Looking further ahead, the major step towards "full" marginal cost pricing can only be taken when it is fully possible to charge for road use also dependent on location and time. Then charges could be made dependent on the true costs arising on different road classes (road wear element), under different driving conditions (accident externality and time cost element) and in different areas (local air pollution costs and noise).

6.2.2   Cost recovery and earmarking – important political issues

The cost recovery issue has been much focused when discussing the implementation of marginal cost pricing. In many transport sectors there are substantial economies of scale in infrastructure provision which generally tend to create a financial deficit when charging by marginal rather than average costs. The issue has also been much focused in the road sector (see e.g. Newbery (1988a), Newbery (1988c) or Nash and Matthews (2001). The general picture tends to be that marginal costs are likely to fall short of average costs in uncongested road networks (thus representing a potential financial deficit), whereas many congested networks will have marginal costs well above average costs. This picture seems to be supported by my findings in the Norwegian case. If the estimated marginal costs per vehicle km in the CATERU Recommended scenario were actually charged, the city traffic alone would create a revenue approximately 3 times the

278

total budget for national roads in 2001 (€ 1.6 billion). The figures presented in Figure 6-1 are estimated revenues if there were no change in traffic volumes in response to the new pricing regime. This is not a likely scenario as charges would be significantly lower for the rural and the town cases, and much higher for cities, compared to the current level of fuel charges. The estimated revenues for cities would therefore be somewhat lower, and probably the reduction will be much bigger than the corresponding increase in revenues that could be expected for the towns and the rural areas. Still, allowing for traffic volume responses to the new regime, it seems that charging by the new estimates would produce revenues far above total road network expenditure (and cost recovery should not be a problem even including the costs of the other public roads, outside the network of national roads).

However, politically it may be a problem that there will be a lot of cross-subsidizing in such a case, as the users of congested city networks will pay for the deficits in other parts of the network. Will this be regarded as "fair" pricing?

	Rural	Town	City	Total
Cars (< 2t), petrol	92	265	3 244	3 601
Cars (< 2t), diesel	9	30	298	337
Vans (2-3,4 t), p	7	20	268	295
Vans (2-3,4 t), d	8	22	476	506
Buses d	4	19	316	339
Light trucks (3,5-7,4 t) d	3	14	292	309
Trucks (7,5-15,9 t) d	1	6	116	124
Trucks (16-23 t) d	3	16	326	344
Trucks (> 23t) d	5	24	430	460
Total	**133**	**416**	**5 765**	**6 314**

Figure 6-1   Estimated tax revenue if CATERU Recommended Scenario is implemented. (Assuming constant traffic volumes). Million Euros.

Another issue frequently raised, is the issue of *earmarking* the revenues raised by marginal cost pricing for transport purposes only. Several studies have shown that

public acceptability is much higher for such pricing schemes if the use of revenues is somehow restricted to transport investments (Herry 2001, Sulkjær 2002, Begg 2002 and Glazer et al. 2001). Generally, earmarking of tax funds has not been seen as a good idea from the perspective of general economic efficiency. Public funds (from any source) should be allocated to the sector where the marginal benefits are highest. However, if earmarking is the only politically feasible way of implementing a marginal cost pricing scheme, the potential losses from allocating the revenue to the "wrong" sector may be more than offset by the efficiency gains from the improved pricing regime.

6.3   FUTURE PERSPECTIVES AND NEED FOR FURTHER RESEARCH

6.3.1   Possible developments in the level of road user charges

Is it possible to give an opinion on the possible developments in the level of road user charges? If we assume that the principle of marginal cost based charges is followed, then this is a question of the expected development in the marginal costs related to road use. Let us analyse this by having a closer look at each of the main external cost components related to road use.

The possible development of external road wear costs

As roads are getting gradually more solid, the marginal traffic volume dependent road wear may decrease. This development is also reinforced by better suspension systems etc. for the HGVs and buses (diminishing dynamic loads). On the other hand there is always a driving force towards allowing higher vehicle and axle weights, because there are economies of scale in the trucking business. If higher weights are permitted, this will of course increase marginal road wear attributable to HGVs[159].

---

[159]   Actually, the marginal benefits from increased axle loads should equal marginal costs from increased road wear (and other external effects) if axle load restrictions were optimally designed.

280

The possible development of external accident costs

The relationship between vehicle densities and external accident costs has been the focus area of several studies. There are multiple, and some times counteracting mechanisms affecting this relationship. One would expect higher vehicle densities to imply higher accident risks in an iso-speed setting. However, when approaching the capacity of the road networks, increased vehicle densities will also mean reduced average speed. Taking this effect into account makes the outcome more uncertain. Adding to this is also the fact that lower speed levels will contribute to a reduction in the average severity of the accidents, hence it is perfectly possible that marginal accident costs may *decrease* with increasing traffic levels in congested areas.

Traffic safety is also highly focused by car manufacturers, and there is no doubt that future vehicles will be safer than the present ones. New safety devices are developed that both have the potential of reducing accident frequencies (active systems) and the consequences of accidents (passive systems). This will also contribute to lower marginal accident costs.

The possible development of external environmental costs

Some of the environmental externalities have been reduced over the past decade due to better cleansing of some exhaust gases. There are a number of similar developments in motor technology (e.g. more efficient engines reducing overall emissions, hybrid-cars, fuel-cell technology etc.), fuel technology (e.g. lower sulphur content) and cleansing technology (e.g. particle traps) that will reduce hazardous emissions from road vehicles in the years to come. All other things being equal, these developments will also contribute to lower external environmental costs related to road use. However, there are at least two limiting factors that may very well counteract this development:

Firstly, if the external environmental cost of road use as a function of total emission levels has a convex shape (Figure 6-2), then the marginal external

281

environmental cost will increase with total emission levels. For many kinds of environmental hazards such a shape of the cost function may be plausible. This means that the effect of the reduction in emissions from the individual vehicle may be (more than) balanced by the effect of a continued growth in traffic volumes, hence the overall effect on marginal cost levels will be uncertain.

Secondly, there is a very strong trend towards centralisation in European countries. Among other things, this also means that the majority of the growth in traffic volumes will come in densely populated areas[160] where the impact of a given concentration of hazardous gases will be higher than an equivalent level in rural areas. This is also a trend that will contribute to higher marginal external environmental costs in densely populated areas in the future.
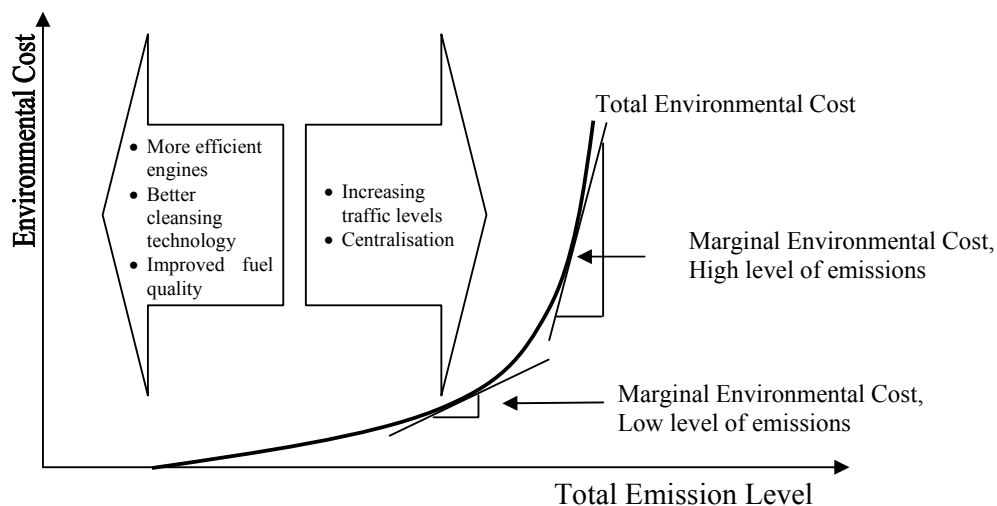
Figure 6-2   Changing marginal environmental cost with total emission levels, and factors affecting the level of emissions from road use

The major uncertainty when it comes to possible developments in environmental costs, is the possible consequences of the greenhouse effect. Using the shadow

---

[160] However, some of the most congested road networks are approaching an absolute capacity limit that will hinder the growth in traffic levels to some extent.

prices of implementing the Kyoto protocol, CO_2-emissions are an almost dominating cost component connected to vehicle use. Still, it is quite possible that the aims in the Kyoto protocol will not be enough to stabilise the effect. At present, however, the uncertainty about the possible impacts of this effect is quite substantial.

The possible development of external time costs
Whereas there is considerable uncertainty about the future development in many marginal cost components, there is little doubt that the magnitude of external time costs will continue to grow with ever rising traffic levels on arterial roads and most of all, in cities. Naturally, this development will, to some extent, be dependent on the investment level in new road capacity. It is possible of course, that new technologies (vehicle routing, adaptive cruise controls etc.) and better demand management systems may improve traffic flows, but this is probably not enough to prevent rising external time costs.

6.3.2 Considerations related to the implementation of new pricing policies
Introducing a new pricing scheme along the lines suggested here may be controversial from a political point of view. Firstly, it would mean introducing some form of an advanced kilometre counter, which also will keep track of the location of vehicles. This may raise serious concern about privacy issues for the road users. However, this issue may not be insurmountable, as such issues have been resolved in other contexts (e.g. the photographing of drivers and passengers within the Norwegian Automatic Traffic Control units).

Secondly, although it would mean a reduction in inter-urban charges, the magnitude of the congestion charges necessary in the major cities will be very controversial. Compared to other cities in Europe, however, this may be slightly easier in Norway as motorists already have been paying some tolls in all the major cities for more than a decade. The acceptance of the current tolling systems has been fairly good, probably due to the earmarking of the funds raised, for transport

purposes.

Norway is not a member of the EU, still a lot of EU regulations have also been implemented in Norwegian legislation due to our membership in the European Economic Area (EEA). EU legislation limits the scope for proper marginal cost pricing (Lindberg 2002a), but it is still unclear how this will affect Norwegian possibilities of implementing a new policy more closely linked to marginal cost principles, should Norway decide to stay outside the EU (the EEA agreement is currently under renegotiation).

6.3.3   The need for more research

Although great effort has been put into research related to establishing an empirical foundation for marginal cost pricing, there is still a substantial need for more research. Lindberg (see Figure 5-59) has evaluated the needs for further research at the European level, and concluded that there is a need for knowing more about all relevant cost components, possibly excluding the congestion cost element. For Norway especially, there is substantial need for more research as we have not participated in the recent major European marginal cost studies. Applying the principles established through the CAPRI and UNITE programmes to studies relevant for the Norwegian road network would be interesting. My contribution to the field of marginal road wear consists in having established the FAMAROW-model. On the other cost components I have merely updated earlier figures using the most recent statistics and estimates from available literature and statistics.

Further research needs on marginal road wear

I have only studied models for rutting and roughness in this thesis. Corresponding research on Swedish roads is based on cracking. Assuming that rutting is the maintenance triggering damage mechanism on most (high volume) Norwegian roads may be plausible, but this should be investigated further. There is a special need to investigate prevailing traffic volume dependent road wear on the weaker

(low volume) parts of the Norwegian road network. The FAMAROW models should also be improved by including data on the solidness of the roads & factual number of axles with studded tyres, by investigating possible measurement errors, and by exploring more advanced statistical models.

Further research needs on accident costs

My estimates of accident costs are average figures based on an updated model established by TØI in the early 1990s. Lindberg concludes in his study that there is a substantial need for more research in this area, mainly on the so-called risk elasticities (i.e. the relationships between accident risk and traffic volume). The UNITE results with negative marginal accident cost estimates are very different from figures used in the Norwegian models, and this calls for new Norwegian studies in this field as well.

Further research needs on environmental costs

Most shadow prices for the emission of greenhouse gases are based on the Kyoto protocol. The political destiny of this protocol has been different from what was initially hoped, and there may be a need for new estimates of shadow-prices based on the prevailing situation.

In central European research programmes (e.g. ExternE), the so-called Impact Pathway Approach has been established as the best way of estimating the external costs of air pollution. The figures included for local air emission costs in my study do not fully follow this approach, as it is based on estimated emissions and because shadow prices from general international and national sources have been applied. More sophisticated models are called for, in order to account for transport and chemical conversion, concentration and deposition, and responses of receptors.

The TØI model included quite substantial noise costs based on a single stated preference study. I have found this evidence too weak to include them in my

recommended figures. However, the preliminary findings in the UNITE programme, indicate that there may exist severe marginal noise costs under special (night-time) conditions. The empirical evidence on noise costs in Norway is very weak, and should be strengthened.

Further research needs on congestion costs
The congestion costs included in my recommended model are based on average (24 hour) congestion costs on the city road networks in Oslo and Trondheim. However, fairly good models exist for providing the necessary information to make cordon tolls for the major cities time differentiated, and thus more in line with factual marginal costs. Congestion is probably a negligible element on the rest of the Norwegian road network.

Summing up, my aim has been to improve the foundation for introducing a charging regime more closely in line with marginal cost pricing principles in Norway. I have introduced some new empirical evidence on the road wear element. These calculations have been made based on a bottom-up approach where the basic building block has been a model (FAMAROW) for rut development based on observations made from a set of measurement sites on Norwegian in-service national roads. The results of this model development may be controversial due to two reasons mainly:

- Firstly, the model suggests that the main cost driver for maintenance-triggering rut development is the use of studded tyres. If this is true for the whole network of Norwegian national roads, then the high proportions of road wear costs traditionally allocated to heavy vehicles (through ESAL based cost allocation) should be severely lowered.
- Secondly, the total magnitude of *marginal* road wear costs calculated by this bottom-up approach is significantly smaller than the one calculated in

286

earlier studies using a top-down approach.

However, these are findings based on a first attempt to develop such models in Norway. There are a number of critical remarks that could be made connected to these findings. Some of them are:

- There are obvious shortcomings in the database used for the estimation of the FAMAROW model: The majority of observation sites represent high volume roads which a high bearing capacity, there is no information about road structure or road strength in the data and there is no direct control over the use of studded tyres.
- Is rutting the dominant maintenance-triggering damage mechanism on Norwegian national roads? The prevailing assumption has been that this may only be true for high volume roads.

These are all factors that are included in the review of further research needs identified in this thesis.

Apart from the new efforts made in the field of estimating marginal road wear costs, I have also updated the estimates of other elements based on recent research and statistics, and tried to identify the remaining research needs for providing a sufficient empirical foundation. I have also tried to provide a good overview of recent international research in the area, thus hopefully also providing a good point of departure for further Norwegian research efforts in this field. Compared to other European countries, Norway has not been very active contributing to new knowledge in this field over the past decades. My hope is that this thesis could trigger more such activity in the years to come.

APPENDICES

Six appendices are added to this thesis. The first one is a review of the state-of-the art of Weigh-In-Motion technology, the second is a brief overview of the data preparations done for the FAMAROW-model, and the third one contains the recommendations from the CAPRI-project on road user charging.

In two separate publications Hjelle (2003a) and Hjelle (2003b) two more appendices (D and E) can be found. The first one containing the Visual Basic macros applied for the FAMAROW data preparation, and the second one the output from the CATERU model (four scenarios).

Finally, in Appendix F I have added a brief presentation of the recommended methodology for estimating marginal external accident costs in the UNITE-project

APPENDIX A: WEIGHING VEHICLES IN MOTION

The efforts to develop methods for weighing vehicles in motion, so-called WIM-systems, has a rather long history, commencing with the first US trials in the early 1950s (Hjelle 2000). The currently available WIM systems may generally be divided into:

- **Permanent.** Sensors and the data acquisition systems collect data at the same location
- **Semi-permanent.** Sensors are built into the pavement while the data collection system is moved from site to site.
- **Portable.** Sensors and equipment are moved freely from site to site.

Both the permanent and the portable systems may be based on different technological approaches to weighing vehicles in motion. In this section I will give a brief description of the current range of systems available, and briefly about primary uses and performance. I concentrate this review on the most common concepts for WIM-systems, i.e. bending plates, bar censors (piezoelectric cables and film), capacitive mats and bridge WIM-systems[161]. This means that I leave out the more traditional systems most suitable for static weighing or very slow motion weighing like the load cell[162] and traditional static scales commonly deployed by the roadside for enforcement purposes.

## Approaches to Weighing Vehicles in Motion

Bending plates
Bending plates are mounted into the road, and the bending in the plate is then

---

[161] There are a number of other systems that have been developed as well (instrumented vehicles, optical fibres etc.), but these are the most used ones.

[162] A device, which can be electrical, hydraulic which produces a signal proportional to the load applied to it

measured as the wheel(s) are passing over the plate. Strain gauges are mounted under the plate to register the load via changes in the measured resistance of the sensor. Some systems are portable and then they are applied to the road surface instead. Mounting bending plates into the road is a rather costly operation, demanding excavation and building a solid base for the plate. Appendix Figure 1 and Appendix Figure 2 illustrate bending plate installations.

Bending plate systems may be of an ASTM Type I, II, III or IV standard depending on the intended use of the device and the number of scales placed in the lane (US Department of Transportation 1997b). Such systems normally consist of either one or two scales. The scale(s) is placed in the travel lane perpendicular to the direction of travel. When two scales are used in a lane, one scale is placed in each wheelpath so that the left and right wheels can be weighed individually. In low traffic volume deployments only one scale is usually used.
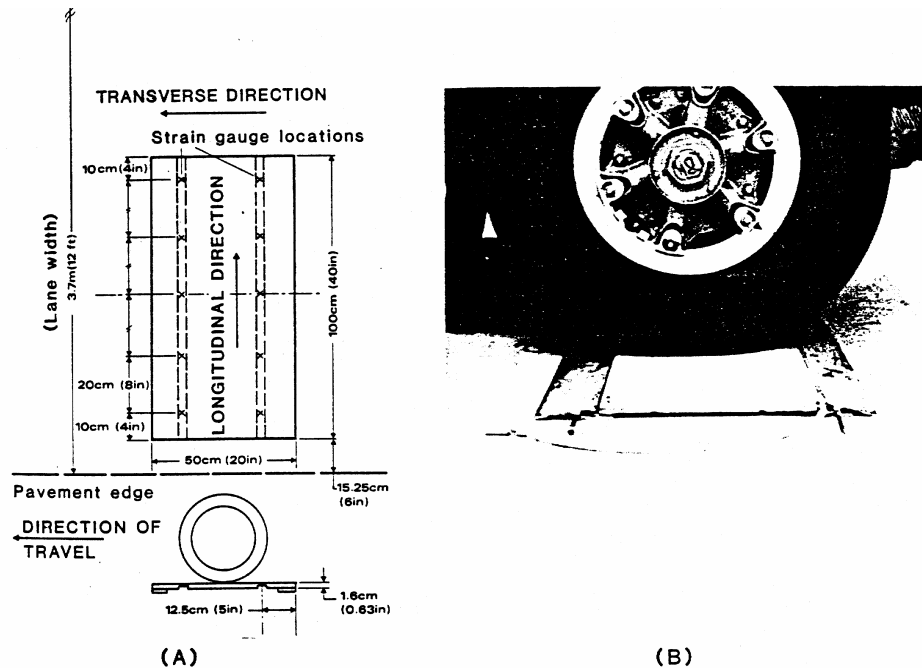
Appendix Figure 1   Illustration of the principle of a Bending Plate[163]

There are both permanent and portable bending plate systems available, however the portable systems are not capable of high (i.e. full) speed measurements.
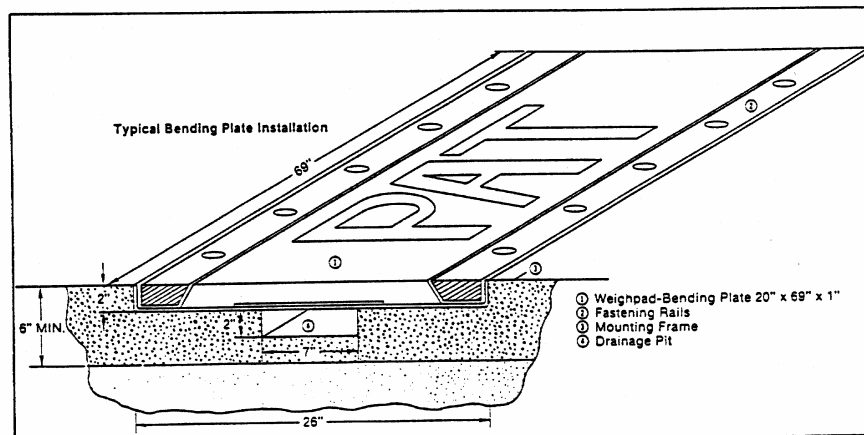
Appendix Figure 2   Illustration of a typical Bending Plate installation [164]

The bending plates are typically combined with one or two inductive loops. An upstream loop is used to detect an approaching vehicle and thereby "alert" the system. Speed and axle spacing can then be established by a second (downstream) loop, or a second weighpad or a cable.

Bar sensors (piezo-sensors)
Bar sensors are based on coaxial cables containing either piezoelectric or piezo-quartz material. The outer and inner conductor are typically made from copper separated with the piezoelectric material. The cables are designed to make the charge independent of the distribution of the longitudinal distribution of the applied load[165]. The cable is thinner than a tyre imprint length and therefore the

---

[163]   Source: Freight vehicle overloading and load measurements, OECD.

[164]   Source: Permanent Automatic Weight and Classification system DAW200, PAT Equipment Corp.

[165]   The firm Thermocoax has developed such a cable, called Vibracoax. Another example is the Vibatek-cable developed

signal generated only expresses the tyre pressure times the width of the tyre. This signal must subsequently be adjusted with respect to the tyre imprint length to obtain the wheel load.

The way the cable is mounted is critical for its performance as a weight sensor. Because the cable is sensitive to stretching, the road should have a high bearing capacity (preferably concrete paving). This may be overcome by moulding the cable in a metal bar, which prevents stretching of the cable. The ideal installation of the cable would be to neutralise stretching of the cable and at the same time to minimise friction towards the metal bar or the pavement. Appendix Figure 3 and Appendix Figure 4 illustrate different approaches to mounting the piezoelectric cable into the road pavement.
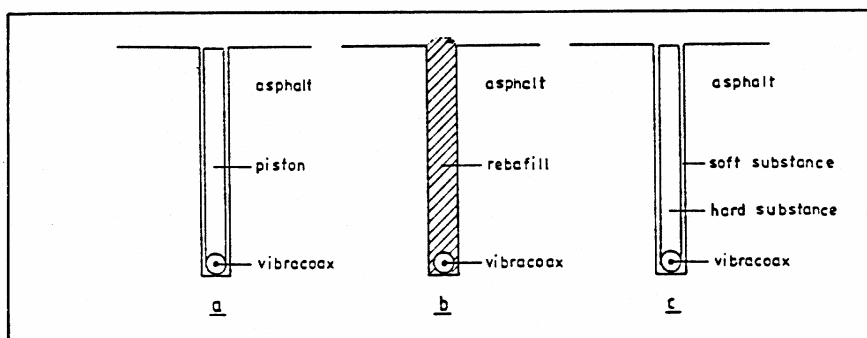
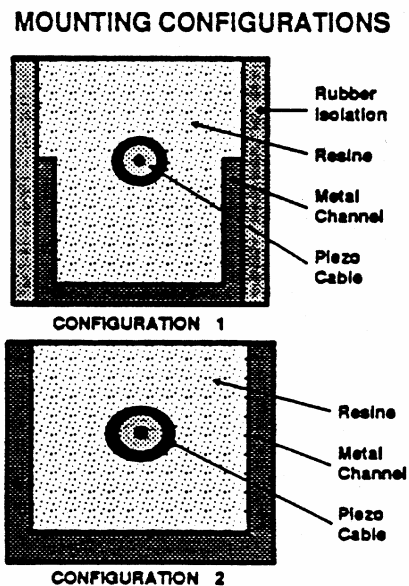**Appendix Figure 3   Examples of different ways of mounting piezoelectric cables[166]**

Another problem may be that the cables are sensitive to changes in temperature. Measurements should therefore be adjusted for observed temperature. High temperatures soften the road surface, which may cause a stretching of the cables.

by Raychem.

[166] Source: Load Measurement under Rough Climatic Conditions in Norway, Datainstrument A/S

Single bar systems do not have sufficient possibilities of controlling the dynamic forces. Systems which use multiple sensors have been developed to cope with this problem. Such systems may be classified as ASTM Type I or Type II, partly depending on the number of sensors used. A more comprehensive review of multiple sensor systems is given in the section called "New approaches to isolating static load: Multiple sensor systems".

MOUNTING CONFIGURATIONS

Rubber
isolation

Resine

Metal
Channel

Piezo
Cable

CONFIGURATION 1

Resine

Metal
Channel

Piezo
Cable

CONFIGURATION 2

Appendix Figure 4   Illustration of two mounting configurations for piezoelectric cables[167]

An even cheaper variety to the piezoelectric sensors is the *piezoelectric film*. In Appendix Figure 5, Appendix Figure 6, and Appendix Figure 7 different concepts of WIM-sensors using piezoelectric film are presented. The latter is a mobile unit attached to the pavement surface. The film-based systems may also be combined in various patterns to enable measurement of contact area and tire pressure (see

---

[167] Source: Development in Piezoelectric Weigh-In-Motion Systems, Ontario Ministry of Transportation, Canada 1987

Appendix Figure 8).

The experiences with these sensors are somewhat mixed. They tend to be rather short-lived and not as manageable as the more traditional piezoelectric bars described above.
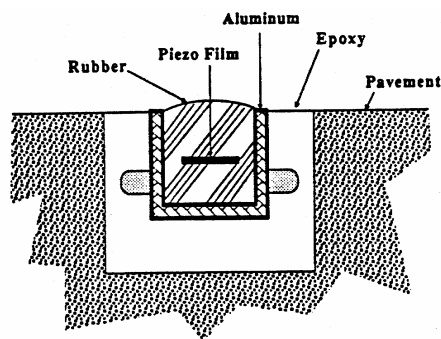
Appendix Figure 5   Piezoelectric film moulded into an aluminium rail[168]
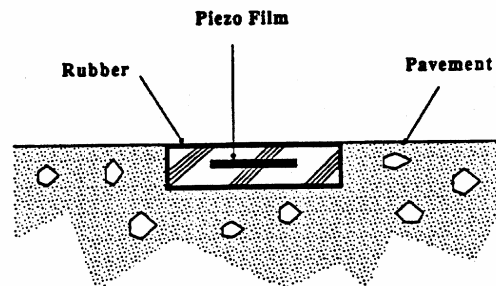
Appendix Figure 6   Piezoelectric film moulded in rubber and inlayed into the pavement

Single sensor systems are usually combined with at least one inductive loop mounted upstream to alert the system of approaching vehicles. If speed and axle spacing are to be registered, it might be necessary also to use a second inductive

---

[168] Source of this and the two following illustrations: Piezoelectric Film Technology, Texas Transportation Institute, 1990

loop downstream.
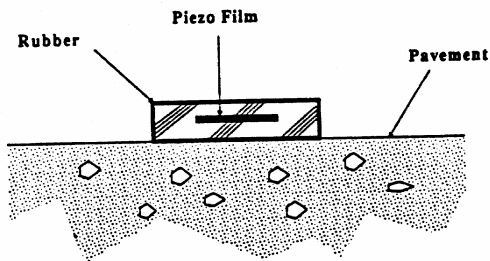
Appendix Figure 7   Piezoelectric film moulded in rubber and attached on top of the pavement
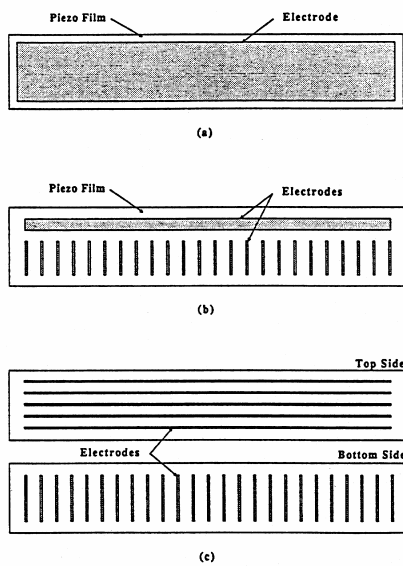
Appendix Figure 8   Sensors for measurement of contact area and tire pressure[169]

---

[169]   Source: Texas Highway Dept.: Forecasting Truck Volume

Capacitive mats

Capacitive sensors are based on the fact that their capacitance (ability to hold electric charges) changes when different the loads are applied to them. This fact has been utilised in different concepts related to WIM measurements. Capacitive mats are commercially available for WIM measurements today. The principle of such a mat is illustrated in Appendix Figure 9. When a wheel rolls over the mat, the three metal plates are pressed together. When the mat is mounted as a condenser in an electronic circuit, the frequency will be dependent on the capacitance of the condenser, which will change when the metal plates are pressed together. This change is measurable and may be converted into a measure of weight applied to the mat.
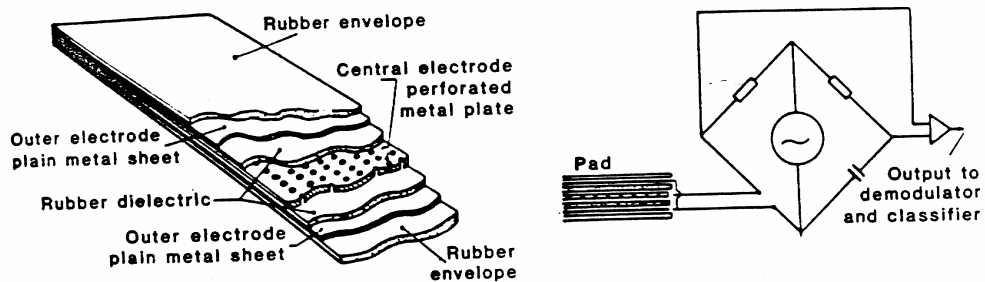
Appendix Figure 9   Illustration of a capacitive mat[170]

The accuracy of the measurements from capacitive mats depends strongly on a proper mounting to the road surface. This may be both time consuming and costly. The mat is movable, but a site-specific calibration is a must.

---

[170] Source: Freight Vehicle Overloading and Load Measurements, OECD
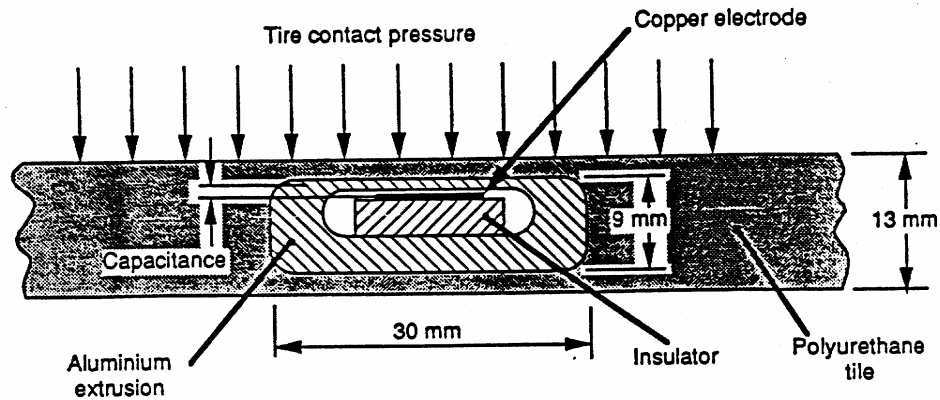
Appendix Figure 10 The principle of a capacitive strip

Appendix Figure 10 illustrates a further development of the capacitive mat, called the *capacitive strip*. This is to some extent flexible, enabling the device to be fixed to rougher surfaces.

Load cells

The typical load cell system consists of a single load cell placed across the traffic lane. This load cell has two in-line scales that operate independently. Depending on the site design these systems may be classified as ASTM Type I, II, III or IV. Once again the typical deployment would also comprise at least one inductive loop for notification of approaching vehicles (mounted upstream). In order to cater for speed/axle spacing measurements, another loop may be fitted downstream. A typical lay-out of a load-cell system is given in Appendix Figure 11.

Bridge Weigh-In-Motion systems

Bridge WIM-systems involve attaching strain transducers to the soffit of a bridge and placing axle detectors on the road surface. The axle detectors provide information on truck velocity, axle spacings and the position of the truck. This information along with the measured strain is fed into the bridge weigh-in-motion

algorithm to determine the axle and gross vehicle weight of the truck[171].   In principle the bridge is functioning as an ordinary scale in this case.

NOTES

1. Cabinet and base
2. 1-2" conduit scale leads in a 6" wide by 6" deep excavation
3. Junction box
4. Drill through shoulder for 1" conduit
5. 3/8" sawcut for loop wire home run Twist 3 turns per foot
6. 3/8" sawcut - axle sensor leads
7. 2" PVC - drain

LOOP 2

AXLE SENSOR

WIM SCALE

LOOP 1

POWER AND TELEPHONE (BY OTHERS)

DRAIN

Note: This layout is similar to the Bending Plate layout shown in Figure 3.1

Appendix Figure 11 Example of Load Cell System Layout[172]

As the measurements are performed through the period in which the whole

---

[171] Source: COST323 – The European Weigh-In-Motion pages: http://www.zag.si/wim/

[172] Source: US Department of Transportation, F. H. A. (1997b).

vehicle is passing over the structure, the system is less influenced by dynamic effects. Bridge WIM systems also provide information about impact factors, lateral distribution factors and strain records which are used for further bridge analysis.

Due to the large mass of the bridge, dynamic fluctuations will be reduced and the length of the construction provides better opportunities for isolating static load better than with single sensor systems on the roads. However, not all bridges are suitable for WIM instrumentation because of limited accessibility, stiffness, high traffic volumes, vibrations etc.
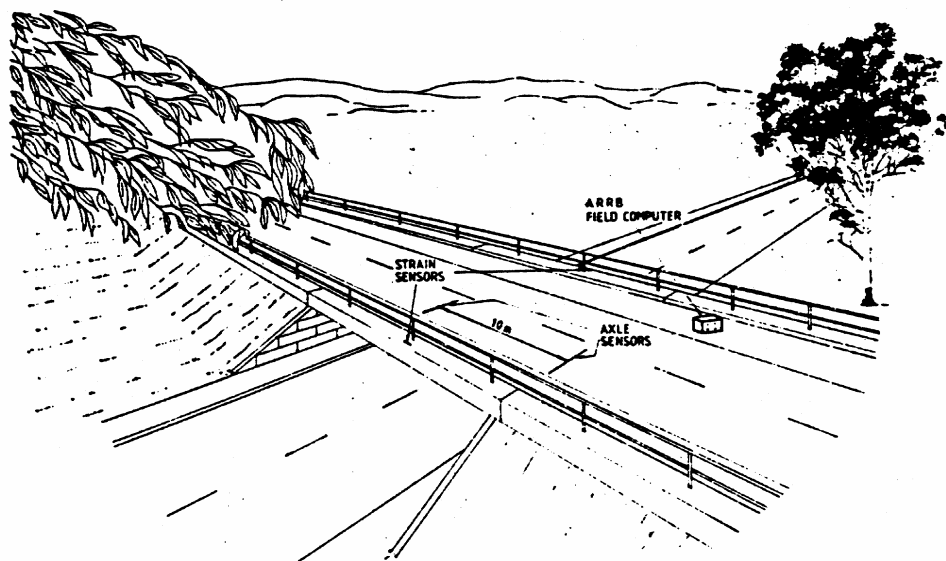
Appendix Figure 12 Illustration of a bridge-WIM deployment[173]

New approaches to isolating static load: Multiple sensor systems
In some uses of WIM-data (e.g. enforcement purposes) one is most interested in

---

[173] Source: AXWAY – A system to obtain vehicle axle weights, Australian Research Board 1984

the monitoring of static loads. To isolate the static part of the dynamic loads is often difficult. Especially when the WIM system contains only one, or a very limited number of sensors. To estimate the static load a *multiple sensor system* is required.

The dynamic interactions between road and vehicle may be represented by theoretical models based on the bouncing motion of the vehicle. The bouncing motion is illustrated in Appendix Figure 13. Here the factual oscillating dynamic loads are plotted (continuous line) along with an estimated single sine-wave representing the same dynamic loading as a function of road distance. The horizontal line (F_0) represents the static weight of the vehicle. The estimation of the sine-wave is based on measurements made by the N sensors deployed.

Recent research (Cebon et al. 1998) concludes that there are basically two different oscillation modes of a heavy vehicle: The so-*called body bounce* mode (1.5-4.5 Hz) and the higher frequency *axle-hop* (8-15 Hz). Algorithms founded on this research (e.g. *op.cit.* and Argoul et al. (1998) have been found to simulate this dynamic behaviour fairly well. The input to these algorithms stem from multiple sensor systems in the form of discrete observations of dynamic impacts.

Dolcemascolo and Jacob (1998) have experimented with *the number of piezoceramic strip sensors* (uniform spacing) connected to a WIM-system in order to find an optimal number of sensors when the purpose is to estimate static loads. Their conclusion is that very much is gained with a 5-7 sensor system. Increasing this to 13 sensors will improve accuracy even more and also provide a good robustness of the results. A further increase in the number of sensors does not improve the accuracy because the performance of the individual sensor then becomes the limiting factor. The tested 13 sensor system achieve almost the accuracy class B+(7)[174].

---

[174] See Appendix Table 2 below about the European Classification system.
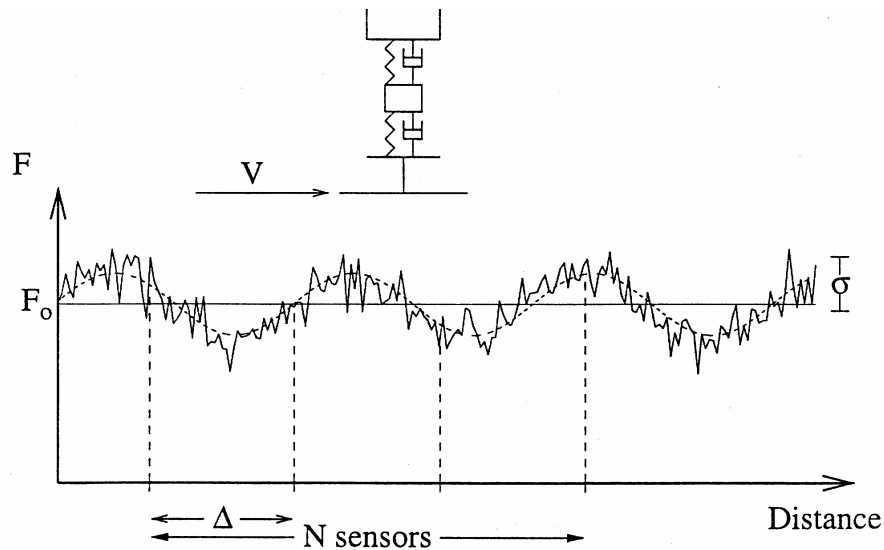
Appendix Figure 13 N-sensor WIM-system measurements of dynamic loads[175]

Elements of software needed for WIM measurements

Dedicated software is necessary to process the information generated by various WIM-technology deployments. Three separate (but communicating) software packages are generally necessary (US Department of Transportation 1997b):

1. On-site software

2. Communications software

3. In-house software

The *on-site software* interprets the signals from the WIM-measuring device and generates the on-site files which typically would include information such as:

1. Site identification

2. Time and Date of Passage

---

[175] Source: Cebon et al. 1998

3. Lane Number

4. Vehicle Sequence Number

5. Vehicle Speed and Classification

6. Weight of all Axles and Axle Groups

7. Code for Invalid Measurement

8. Optional Graphic Configuration

9. Equivalent Single Axle Loading (ESAL) value

10. Any other information collected on-site (e.g. from climatic sensors)

The typical *communications software* should allow for two-way communication between the WIM-deployment site and the site for gathering (and analysis) of the data. The in-house computer may have a need to send e.g. revised calibration data to the on-site computer. Equivalently, the on-site computer also needs to collect the information from the WIM-sites.

The main function of the *in-house software* system is to process the collected data and to generate reports based on the data. A typical system would enable the following tasks to be done (op.cit.):

1. View real time vehicle selectable by lane

2. Reset the system clock

3. Monitor system memory in terms of storage remaining

4. Set-up and initiate the generation of summary reports on data previously collected by the system

5. View generated reports

6. Generate and view error reports including time down, system access, auto-calibration, and improperly completed records.

7. Transfer selected raw data files or generated reports from the site system to the office host computer

8. Purge old data files from the system

Cost comparison of WIM systems

The factual costs of installing and operating a permanent WIM installations will of course depend on the selection of equipment, but also on the total layout of the system, the traffic volumes etc. Taylor and Bergan (1993) presents the figures in Appendix Table 1. The performance column is estimated under ideal ASTM site conditions. Maintenance includes all cost items such as power and communication, structural costs (roadway and scale frames) and core WIM-system costs.

WIM System	Performance (Percent error on GVW at highway speeds)	Estimated Initial Cost per Lane (Equipment and Installation)	Estimated Average Cost per Lane (12-year life span including maintenance)
Piezoelectric Sensor	+/- 10%	$ 9,500	$ 4,224
Bending Plate Scale	+/- 5%	$ 18,900	$ 4,990
Double Bending Plate Scale	+/- 3-5%	$ 35,700	$ 7,709
Deep Pit Load Cell	+/- 3%	$ 52,500	$ 7,296

Appendix Table 1   Cost comparison of WIM Systems (Source: Taylor and Bergan, 1993)

Major conclusions from this cost analysis are:

1. The Piezoelectric Sensors are by far the cheapest to install.
2. Over a 12-year life-span the picture is not so simple. Bending Plate Scales come down to almost the same cost per lane in the long run.
3. The most significant improvement in performance seems to be achieved going from piezoelectric sensors and the (single) bending plate.
4. The highest accuracy, and the highest costs relate to the double bending plate scale and the deep pit load cell.

It seems that there is a *trade-off between performance and price*. However, the "optimal" solution will be rather dependent on the needed level of accuracy, and on the relevant time-horizon.

Evaluation criteria and classification of WIM-systems

What are we looking for? The "perfect" WIM-system ?
I have briefly reviewed the principal WIM-solutions available and the major relationships between road use and road wear. As mentioned earlier the major focus of this review is to evaluate WIM-systems as a means of collecting information about factual loadings related to different vehicles with different weights and different axle configurations.

There are many different potential uses of the data generated from WIM deployments, mainly related to the detection of overloading, studying the relationship between factual loadings and pavement and bridge wear, and the enforcement of restrictions. There is also a potential for more active road pricing related to weight using WIM systems. The many varied applications of WIM systems makes it rather difficult to establish one "perfect" WIM system. Still, for most purposes a system capable of measuring factual loadings with a highest possible degree of accuracy would tend to be the ultimate goal. One problem is that for some purposes the factual dynamic loadings are what we are looking for (e.g. to establish the infrastructure responses to the loadings), in other applications we would prefer to isolate the static element of the loading (e.g. related to enforcement of loading restrictions).

There is a trade-off between costs and the desired high accuracy of the systems. We are therefore mainly looking for the most cost-efficient ways of achieving a *sufficient* degree of accuracy and quality, rather than the most accurate solution possible from a pure technological view. The degree of necessary accuracy will be different for different applications, hence we need to develop a classification system that relates to the different needs for different purposes. The development of such classification systems, have been some of the major concerns of the recent

306

research activities, both in Europe and in the USA. In the two subsequent sections I present a proposed European classification system and the American classification system. After that I try to relate the different technological concepts to the European classification system.

The European WIM-specification system

The European WIM-specification system has been established by the COST 323 management committee, and is described in Jacob and O'Brien (1998). This specification's main purpose is to give a description of what accuracy is needed related to different WIM applications, *and* what accuracy the different system specifications is likely to provide. The benefit of such a classification of WIM systems is rather obvious: This will be a great help when users are to choose the most suitable system for a specific need. Moreover, since the users to a great extent are public agencies, such a system forms a very good platform for tendering specifications. At a later stage this European WIM-specification will form the basis for the establishment of a more formal *standard* (like the American one – ASTM).

The European WIM specification also covers topics related to finding the best locations for WIM measurements, procedures for calibration and verification of WIM deployments and a classification scheme which is also applied to the requirements imposed by different uses of the systems. The specification forms a very thorough document which includes both "mandatory" items and more suggestive information. In the subsequent sections I will try to cover the main contents of the specification without becoming too concerned with the more detailed technicalities.

Accuracy Classification

The division of WIM systems into classes is made subject to the different needs for accuracy related to different purposes. For a specific WIM installation to comply with a given accuracy class, the probability that results are within the

interval $[(W^S(1-\delta), W^S(1+\delta)]$, where W^S is the accepted reference value, must exceed a specified minimum, π_0. The confidence interval width is determined by the factor δ, hence the accuracy classes can be identified as different requirements related to this factor. The required confidence interval width is different for different types of measurement as indicated in Appendix Table 2 which defines the accuracy classes of the European WIM specification.

Criteria (type of measurement)	Domain of use	Accuracy Classes: Confidence interval width δ (%)						
Accuracy class:		A(5)	B+(7)	B(10)	C(15)	D+(20)	D(25)	E
Gross weight and Axle load	Gross weight >35kN	5	7	10	15	20	25	>25
	Axle load >20kN							
Group of axles		7	10	13	18	23	28	>28
Single axle		8	11	15	20	25	30	>30
Axle of a group		10	14	20	25	30	35	>35

Appendix Table 2   European WIM classification (Jacob and O'Brien 1998)

The minimum probability, π_0, is a function of the test conditions (repeatability, reproducibility, duration etc), and the sample size. These are features of the test procedures recommended in the European specification:

Test procedures, calibration and verification

Because all WIM installations need to be adapted to the local conditions there is a need for calibration of the systems. The European specification contains guidance on several testing procedures that may be adopted for this purpose. As a means of quality assurance procedures for verification of the system (initial and in-service) should also be applied. Basically the methods for these purposes are the same. The most commonly used way of calibrating and testing WIM deployments is statistically to compare repeated loadings with a known static weight (e.g. a pre-

weighed calibration truck) with the output from the WIM system.

There are several ways of carrying out such a test:

- *Full repeatability conditions:* One vehicle passes several times at the same speed, load and lateral position.

- *Extended repeatability conditions:* One vehicle passes several times at different speeds, different loads and with small variations in lateral position (in accordance with typical traffic).

- *Limited reproducibility conditions:* A small set of vehicles (typically 2 to 10) representative in weight and silhouette of typical traffic, is used. Each vehicle passes several times, at different combinations of speed and load and with small variations in lateral position.

- *Full reproducibility conditions:* A large sample of vehicles (some tens to a few hundred), taken from the traffic flow and representative of it, pass over the system, each of them passing only once.

The further down this list one moves, the better would the calibration procedure be. However, undertaking a test under full reproducibility conditions is also rather complicated and costly compared to a full repeatability test procedure, so there is a trade-off between desired accuracy and cost here too. To really test the performance of a WIM site one would need to carry out a more long term test procedure that also would encompass *environmental* factors like weather conditions (temperature, rain/snow, freezing and thawing) and changing traffic patterns (vehicle types, tyre types, speed).

Jacob & O'Brian *(op. cit.)* present both standardised testing schemes and comprehensive procedures for analysing the test results proposed in the European specification. I leave the description of the testing, calibration and verification procedures here, and turn to the main application areas for WIM systems and the

corresponding requirements related to the above defined European classification system.

Main applications and their requirements

Three main types of applications of WIM-systems, with increasing demand for accuracy, are identified by COST 323 research (COST-323 1997):

- *Statistics:* Economic, geographical and technical studies of freight transport and traffic evaluation. Accuracy classes D(25) or D+(20) seem to be suitable for this application. In some cases lower classes such as E(30) and below, may be accepted for a rough evaluation, using simple and inexpensive devices or on poor pavements.

- *Infrastructure and preselection:* Detailed traffic analysis, road and bridge design and maintenance, pavement aggressiveness and fatigue studies, code calibration, preselection for enforcement. Classes C(15) or B(10) are required.

- *Enforcement of legal weight limits:* If the legislation allows the use of WIM (Low speed WIM or eventually High speed WIM), Class A(5) is likely to be required. It is possible that B+(7) may also be accepted if the vehicle exceeds the legal limit by a sufficient amount.

The requirements are only indicative, the necessary performance level must be considered in each case dependent on the local circumstances.

Choice of WIM site

The choice of location for the installation of a WIM site may significantly affect the expected accuracy of the results. The major concern is related to *road roughness* because a more coarse road surface will increase the magnitude of the dynamic loadings and thereby make the calculation of static loads more difficult. In the European specification the site conditions are divided into three classes (I=Excellent, II=Good and III=Acceptable) depending on measured dynamic and quasi-static deflections and the evenness (roughness) of the road surface (IRI

index). If the particular site in question doesn't measure up to the demands along these lines, it may be impossible to achieve the desired accuracy at certain locations. This is illustrated in Appendix Table 3.

Accuracy Class	WIM Site Class		
	I (Excellent)	II (Good)	III (Acceptable)
A (5)	Sufficient	Insufficient	Insufficient
B+(7)	Sufficient	May be sufficient	Insufficient
B(10)	Sufficient	Sufficient	Insufficient
C(15)	More than sufficient	Sufficient	Sufficient
D+(20)-D(25)	More than sufficient	More than sufficient	Sufficient

Appendix Table 3   WIM Accuracy Class Likely to be Achievable in a given WIM Site Class (Jacob and O'Brien 1998)

I have now briefly described the some of the major elements of the proposed European specification of WIM systems with special emphasis on the requirements related to the different fields of application, and on quality assurance procedures necessary to achieve reliable data from the WIM deployments.

The work with this European specification has mainly been done over the past three or four years, whereas the USA has had an adopted *standard* for WIM systems several years before that. Since the ultimate aim is to end up with a formal standard[176] on the European level too, I give a short presentation of the experiences with the American ASTM standard in the following section.

Some experiences with the American Classification system (ASTM)[177]
The first American standard was established in 1990[178]. Before that, in the 1970s and 1980s a lot of governmental agencies had started to use WIM equipment at a

---

[176] Parallel to the work carried out by COST 323 there has been a process aimed at establishing a standard for the requirements for in-road instruments which weigh road vehicles in motion for legal purposes (trade and law enforcement) – the so-called OIML standard. This is further described in Dunmill 1998.

[177] This chapter is based on Lee 1998.

[178] The so-called "ASTM Designation: E 1318 Standard Specification for Highway Weigh-In-Motion (WIM) System with

rather large scale. Being governmental agencies they were obliged by law to put the purchases of the WIM systems out to public tender (or so-called *competitive sealed bidding*). Among other things this meant that specifications had to be made as a part of the tendering process. These specifications were not only significantly different from one state to another, but also *within* states. This created a deficiency in the provision of WIM systems because so many different systems had to be developed and tested separately. The need for a common standard was acknowledged by both users and vendors of WIM technology.

The important features of the ASTM WIM system standard specification are related to:

- Terminology
- Functional Classes (Types) of WIM Systems
- Performance Requirements
- User Requirements
- Test Methods[179]

Very briefly the experiences with this standard is that the elements regarding the Types of WIM systems, The Performance Requirements for each type, and the User Requirements part are the most popular ones. The items related to test methodology and calibration have been less used. Little has been done to evaluate the potential capability of the Type IV systems designed to be used in enforcement situations.

User Requirements and Test Method". This was revised in 1994 and is designated E 1318 –94.

[179] I.e. Static Reference Weighing, On-site Calibration, Type Acceptance Test, On-site Acceptance Test

Type	Generation	Status (Company)	Estimated cost group	Accuracy Class (European specification)
Scales	1960-70	Disused		
Bending plates	1980-90	Operational (PAT)	€ 4000-8000	B/C
Load cells	1990	Operational (CAPTELS-IRD)	> € 8000	A (Low speed) B/B+ (High speed)
Capacitive mat	(1970-80) 1990	Operational (Mikros, Haenni)	€ 4000-8000	B/C
Capacitive strips	1980-90	Operational (Golden River)	€ 2000-4000	C/D
Piezoceramic strips	1980-85	Operational (ECM, Sterela, Lacroix)	€ 1000-2000	C/D Exceptionally B with an efficient self-calibration procedure
Piezocermaic nude cables	1990-95	Operational (Data Instrument, Sterela, Vectra)	€ 500-1000	D+/E
Piezoquartz strips	1995	Operational (Kistler / Golden River)	€ 4000-8000	B/C
Optical Fibre strips	(1990) 1997-1998	Under development (Alcatel)	Unkown	B/C ?
Piezopolymer strips	1995	Operational (Measur Sp., Peek Traffic, Atochem, Focas)	€ 500-1000	D+/E
Piezopolymer nude cables	1990	Operational (Measur Sp.)	< € 500	D/E
Piezoresistive	1980	Disused		
Multiple sensors (5 to 15 strips)	1993-98	Ongoing research Preliminary results	> € 4000	A?/B+

Appendix Table 4   Main WIM sensor types, accuracy class and estimated costs per installation (Coste 1998)

A more comprehensive standard is under work, also comprising additional items such as data storage formats, transmission and processing of data, simpler calibration procedures, site qualities (roughness), vehicle classification etc.

313

Classification of current (and past) systems

As I have pointed out on several occasions the performance of any WIM depends strongly on case-specific factors like the characteristics of the site (road roughness especially) and the calibration and configuration of the system. Still the different technical concepts have different potential with respect to accuracy and it is therefore possible to allocate different systems to the different accuracy classes established in the European specification. This is done in op.cit., and this information is reproduced in Appendix Table 4.

The Norwegian ATK-deployments as WIM devices

In the previous chapter a general classification of the major WIM-concepts related to the European specification was presented. As I have already mentioned, the factual applicability of the different systems will also depend on case- and site-specific conditions. One of the major factors among these environmental conditions is *temperature*. Special focus has, both in COST 323 and in WAVE, been put on the performance of WIM systems in a cold climate environment. Since the performance of WIM systems in Norway is the primary focus of this chapter, I start out with a brief review of the outcomes of this particular research activity before I turn to an evaluation of the Norwegian ATC deployments as WIM installations.

COST 323 and WAVE on cold climate environment for WIM

The research on WIM performance in cold climate environments is reported in Henau and Jacob (1998), Jehaes and Hallström (1998), Huhtala et al. (1998), Hallström (1999) and Jehaes (1999).

The research carried out in the Cold Environment Test (CET) focused on (Hallström 1999):

- Effects on sensor and electronics

- Effects on road surface characteristics

- Effects on vehicle mechanical characteristics

- Effects on driver behaviour

- Effects on the quality of the reference weighings

In a European setting the cold environment problems are most relevant in the Alpine and northern regions. The effects were therefore studied in two Swiss (St. Gotthard and San Bernadino) and one Swedish (Luleå) test sites. The Swedish test was managed by the Swedish National Road Administration and analysed by the Belgian Road Research Centre and the Finnish Technical Research Centre (VTT). The Swiss tests were managed and analysed by the Swiss Federal Institute of Technology (ETH).

The results from the Luleå test site
In Luleå four different pavement WIM systems and one bridge WIM system were tested, these were (*op.cit.*):

- A two bending plate system (DAW 100 by PAT)

- A two nude 3mm diameter piezoceramic cable (Datainstrument)

- A two piezoquartz sensor (Lineas and Marksman 660 combined by Kistler and Golden River)

- A prototype instrumented steel and concrete structure (Omni Weight Control)

- An instrumented bridge WIM system (Trinity College, Dublin)

Data were collected over 7 test periods over a full year (ending in June 1998). The road (pavement) condition was evaluated as a class II (good) according to the European specification. Between November and May the average temperature over a 24 hour period is normally below 0ºC, and in spring and autumn the daily variations in temperature may be above 20ºC at the surface of the pavement.

The performance of the systems tested at Luleå, according to the European specification, is summarised in Appendix Table 5 (based on tables 8 and 9 in (Jehaes 1999). One interesting point of this test is the loss of accuracy as we turn from summer to winter conditions. This is most visible for the PAT and the OWC systems. All systems recovered their initial accuracy during the following summer.

System[180] Season	KI/GR Summer	DI Summer	PAT Summer	OWC[181] Summer	KI/GR Winter	DI Winter	PAT Winter	OWC Winter
Measurement of:	Class	Class	Class	Class	Class	Class	Class	Class
- Single axles	C(15)	E(60)	B(10)	E(30)	C(15)	E(50)	D(25)	E(60)
- Axles of a group	B(10)	E(65)	B(10)	E(35)	C(15)	E(45)	D(25)	E(60)
- Group of axles	C(15)	E(65)	C(15)	E(35)	C(15)	E(50)	E(30)	E(65)
- Gross weight	D+(20)	E(65)	C(15)	E(30)	C(15)	E(50)	E(30)	E(75)
Accepted class	**D+(20) ~C(15)**	**E(65)**	**C(15)**	**E(35)**	**C(15)**	**E(50)**	**E(30)**	**E(75) ~E(70)**

Appendix Table 5   Classification of the different WIM installations in the Cold Environment Test in Luleå

The results from the St. Gotthard and San Bernardino test sites

The Swiss WIM tests (the so-called Alpine Tests) comprised WIM installations at both entrances of two motorway tunnels. On each site, one day of measurement with pre/post weighed vehicles was realised per year since 1996. These tunnels are both located on the A2 motorway which has a very high traffic load and also a high proportion of HGVs (21 %). The main focus of the Alpine Tests was on *initial and periodical calibrations*, rather than on the analysis of normal measurement periods (op.cit.).

The pavement conditions were classified as excellent (class I) for the Gotthard site and good (class II) for the San Bernardino site according to the European classification system. Average temperatures varies between –3ºC/-1ºC in winter

---

[180] GR=Golden River Traffic Ltd. (UK)   KI=Kistler Instrumente AG (CH)   PAT=PAT Equipement Corp. (US) DI=Datainstrument AS (N) OWC=Omni Weight Control

[181] Note: The OWC-system was a prototype and was modified several times during the test.

316

and 13ºC/16ºC in summer for Gotthard and Bernardino respectively. In winter roads are heavily salted and the water brought into the tunnels is not washed away by rain, causing a very high concentration of corrosive substances.

At St. Gotthard a Golden River system comprising four capacitive strips and one inductive loop was installed (on both directions) in 1995. The strips had to be adjusted in 1997, and eventually replaced by a combination of crystal piezo-quartz sensor (Kistler) and electronic (Golden River) in June 1998.

At San Bernardino a Pietzsch system was installed in January 1996 comprising two bending plates fixed in a steel frame combined with two inductive loops. Both systems work without automatic self-calibration or temperature compensation.

The performance, according to the European specification, of the alpine installations is summarised in Appendix Table 6 (based on tables 5 and 6 in op.cit.). The PAT installation in San Bernardino showed a remarkably stable performance over the 3 year test period.

Year and system[182]	1996 GR	1998 KI/GR	1996 PAT	1997 PAT	1998 PAT
Measurement of:	Class	Class	Class	Class	Class
- Single axles	C(15)	C(15)	B(10)	B(10)	B(10)
- Axles of a group	B(10)	C(15)	C(15)	B(10)	B(10)
- Group of axles	C(15)	C(15)	C(15)	C(15)	C(15)
- Gross weight	C(15)	C(15)	B(10)	C(15)	B(10)
Accepted class	**C(15)**	**C(15)**	**C(15)**	**C(15)**	**C(15)**

Appendix Table 6   Classification of the different WIM installations in the Alpine Test

---

[182] GR=Golden River Traffic Ltd. (UK)   KI=Kistler Instrumente AG (CH) PAT=PAT Equipement Corp. (US)

Conclusions and recommendations from the cold environment test

Some recommendations are given by Hallström (1999) as a conclusion drawn
from the CET:

- Select a *WIM site* with the following properties:
- Well-known road and surface characteristics and behaviour through the
  various seasons
- Free floating traffic
- Traffic characteristics[183]
- Lane discipline[184]
- Easy access to electricity and telephone lines (for data retrieval)
- High quality reference weighing facilities should be available
- Use sensors not affected by changes in pavement rigidity; nude piezoceramic
  cables seem to be less adapted for such climates, above all on low volume
  roads because of the use of a self-calibration procedure.
- The piezoquarz crystal and the bending plate systems gave rather good results.
  The first one was stable along the year, while the latter had some temperature
  sensitivity.
- Bridge WIM is a promising up-coming technology for cold climates.
- A new prototype large instrumented structure may also have a promising
  development, because of its insensitivity to the studded tyres, salt etc.

The use of very flexible pavements in the northern countries seem to disadvantage
most of the strip sensors (Jehaes 1999), and above all the nude cables (in
particular the thinner ones) do not perform well. Snow in the winter-time makes
the road more uneven and also causes traffic to have a different transverse
position in the lanes in wintertime. These effects, together with the presence of

---

[183] This item is included in Hallströms list, but it is not quite clear what is meant here.

[184] I.e. Choose a site where the vehicles is more or less forced to pass the measuring device properly (no option of going
around)

318

sand and salt which may affect the instruments, cause WIM measurements under cold winter conditions to be less accurate.

Evaluation of the Norwegian ATC[185] deployments as WIM installations

The Norwegian ATC-deployments are automatic speed surveillance systems comprising a double piezo-electric sensor (nude parallel cables with 3m longitudinal spacing), an automatic camera and a instrumentation unit. These systems are situated on major roads that historically have a combination of above limit average speed and a high frequency of accidents. Vehicles that exceed the speed limit (by a certain margin) are photographed and the driver is subsequently fined via an identification of the number plate and the vehicle register.

Related to the BUAB[186]-project, conducted by the Norwegian Public Road Administration (NPRA), tests were conducted of the possibility of using these ATC-deployments as WIM measurement devices in the early 1990s. Software and (auto)calibration procedures were developed to turn these deployments into a simple form of a WIM system. The benefit of the use of autocalibration procedures[187] is that one does not have to control environmental factors like temperature and moisture.

The test runs made are unfortunately not very well documented, but the achieved accuracy of the systems is reported in Senstad (1994) and reproduced in Figure 4-3. For the gross weight of the vehicle one found that 84% of the heavy vehicles were registered with an accuracy of +/- 20 percent relative to controlled static weight. The accuracy was even higher for the front axle.

---

[185] ATC (Automatisk Trafikk Kontroll)=Automatic Traffic Control

[186] BUAB (Bedre Utnyttelse Av vegers Bæreevne)=Better utilisation of the bearing capacity of roads.

[187] Autocalibration means that the system is calibrated through a number of passages of ordinary cars. Knowing the approximate wheel-load of these cars one can calibrate the recordings according to the signal given from these passages.

These trials indicate that the current ATC deployments in Norway do have a potential as crude WIM measurement devices. However, there is still much research to be done before the expected performance of these systems can be established. Bearing in mind that these tests were conducted under rather idealistic (summer) conditions the expected average accuracy of a typical ATC system would probably be lower than the figures reported above.

WIM systems: State of the art

A young technology, already useful in some applications

This literature study has briefly described the recent developments in a new technology for weighing vehicles in motion. The technology is a relatively young one and important improvements and experiences are still being made every year. However, a lot of progress has been made in developing reliable and accurate systems over the past decades. The US FHA has been a major contributor to the knowledge in this field, not least through the systematic collection of "States best practices". These give practical hints on how to install and run different WIM systems in order to achieve the best results. The US was also a forerunner in defining a standard (ASTM), mainly for making public tenders possible. This also created a better environment for private enterprises that wanted to take part in the development of new systems.

In the 1990s the European research carried out in the major international programs COST 323 and WAVE has been the major source of new knowledge about WIM systems. These programmes have supported research work in many countries ranging from pure technological development and new conceptual approaches to massive testing of different systems in different environments. Even legal issues have been addressed related to the potential use of WIM systems for direct enforcement.

320

WIM systems may have many different uses, and the need for accuracy will be different for different purposes. Many of the current deployments based on existing technology have already proven useful for statistical purposes (European accuracy class D). Some existing systems also meet the requirements for more detailed traffic analysis and may help evaluating, and possibly improving, prevailing road and bridge design algorithms. These systems, which comply with the demands of accuracy classes B or C, are mainly being used for preselection of vehicles exceeding the legal limits for vehicle or axle weights.

The full potential of the WIM systems will not be released until it is possible to have systems of class A or B+ in full and reliable operation. Only then it will be possible to use the systems for direct enforcement of weight limits. The potential for improving the design procedures will also be very much extended if such high quality systems were developed. The most crucial challenge is to design systems that can capture the dynamic behaviour of the vehicles and still be able to estimate the static weight of the vehicles with a high level of accuracy. The most promising technologies here are the different multi-sensor systems that have been introduced over the past years.

The applicability of WIM systems for cost allocation and pricing of HGVs
As stated earlier in this review, my main interest area has been the potential use of WIM systems for road cost allocation and for future road pricing systems for HGVs. I have described the major mechanisms of road wear (cracking and rutting) briefly. I have also focused on the problem related to measuring static and dynamic impacts with the WIM systems. My two interest areas are really asking for both of these figures. Road wear is dependent on the prevailing dynamic forces, hence the measurements made by the WIM systems should be directly relevant to studies of road cost allocation. A differentiated pricing regime applied to HGVs, on the other hand, would probably have to be based on static loads, possibly adjusting for factors that influence dynamic behaviour (suspension type,

tire type, tire pressure etc., i.e. the "road friendliness" of the vehicle).

This means that the current class B and C systems should be useful for cost allocation studies, if one relates road wear to the characteristics of the vehicle. The problem arises when one is to transform the outcome of such studies to a practicable charging system based on static loads. The systems needed for such purposes would probably need to be of classes A or B+ and here more research is clearly needed to develop the required technology.

Future perspectives on technological development and uses of WIM systems
The major applications of the future WIM systems will probably be in the areas of (Jacob and O'Brien 1998):

1. Improving road and bridge dimensioning procedures
2. Detection, control and repression of the overloads, and transport management
3. Road Pricing
4. Weighing of Aeroplanes

The focus in the development of WIM systems has so far been on the first two items, which will still be important in the future. However the potential use connected to road pricing is one of the coming areas. The huge increase in air traffic has also created an interest from the airport operators to use WIM equipment to measure factual loads imposed on the pavement under takeoff and landing.

Jacob (*op.cit.*) expresses a belief that the new techniques such as optic fibres, instrumented bridges and multiple sensors will prove satisfying to most of the expressed needs in the near future. There is still some development work to be done, and - not least - a lot of testing and evaluation, before we know the actual performance of these more refined systems.

322

Some of the major challenges related to further WIM developments are (according to Jacob *op.cit.* and to Giblin 1999):

1. To ensure that the new technology is tested, approved and implemented
2. To ensure that the findings of national and international research in this field is widely distributed
3. To establish a European standard for WIM systems and eventually; a worldwide standard.
4. To integrate WIM systems into more comprehensive "intelligent road" and "intelligent vehicle" concepts.

These challenges indicate that the potential for the pure technical performance of the WIM systems as such is close to being reached, and that the future research and development work will be focused on the problems of integrating this new technology in the general systems related to road building and road management (in the widest sense of the notion).

APPENDIX B: DOCUMENTATION   OF   THE   FAMAROW   DATA PROCESSING

Prior to the estimation of the FAMAROW-model presented in Section 4, a lot of data preparations have been done. The raw data have been extracted from a number of external databases, and processed by Visual Basic macros attached to MS Excel Workbooks. The macros are documented in Hjelle (2003a), and the overall structure of the data processing is illustrated in Appendix Figure 14. Here I will only give a brief overview over the data processing steps.

Data extracted from the Norwegian Road Data Bank

The Norwegian Public Roads Administration (NPRA) maintains a comprehensive database called the Road Data Bank (RDB). This database has many registers, each holding specific information about the network of national roads. Unfortunately the user interface of the RDB is not very well developed, and a particular lack is automated procedures for extracting information across road sections, counties and time. The interface with other data processing programs could also have been better, although many reports may be given in Excel and dBase formats. For the purposes of this thesis, this has meant a tremendous lot of work related to the extraction and processing of the data needed. From the RDB I have extracted data for:

- Roughness measurements
- Rut depth measurements
- AADT measurements
- Maintenance actions

All of these data have been extracted for each relevant road section (+/- 100 m around the WIM-sites), and for road networks representing the years 1990 to

2000. Each 200 m section may be divided into many smaller sections within the databank, in which case I have added the links together by calculating sums or averages representing the 200 m stretches.

For the extraction of IRI and ruth depth data I have used "Report 1" to generate *.sdv files (fixed format text files) for each month and each observation year. The IRI-files are first processed by the macro "LagIRIBok" (this, and all subsequently mentioned macros could be found in (Hjelle 2003a). Apart from merging the files, and arranging them into Excel Workbooks (one for each county) with spreadsheets for each observation year, this macro also corrects road identification numbers that have changed over the period. The next step is taken by the macro "FixIRIFil" which takes the ten county-files and merges them into one Workbook. Here the observations get assigned an id-number that matches the WIM-measurements, average (weighted) IRI values are calculated for each road section, and a time-series of IRI-loggings is generated based on the spreadsheets with observation dates and loggings for each year in the observation period.

A very similar procedure is applied to the rut-depth measurements by the macros "LagSporBok" and "FixSporFil".

Measurements of average annual daily traffic (AADT) figures are extracted by using "Report 2" in the Traffic Register in the RDB. This report also generates fixed format text-files for each month and year in the observation period. The files are merged, sorted and new road identification numbers are assigned to the sections in the macro "LagADTBok". As for the Rut-depth and IRI-files, the next step, performed by "FixADTFil" is to assign WIM-ID-numbers, calculate weighted average values for the 200 m sections, and organising the loggings into time-series for each observation site. The AADT-loggings have been done with various intervals within the observation period, and as I needed to estimate the traffic at specific dates (i.e. the dates of the IRI and Rut-depth loggings), a simple

regression analysis is performed for each observation site by the macro. The resulting estimators of this second order equation are merged into the spreadsheet, paving the road for using these for predictions of the traffic levels at the observation times.

The fourth, and last, data that was extracted from the RDB, is the recorded pavement actions. "Report 1" in The Pavement Register in the RDB was used to extract *.txt files for each county. These files have been processed by the macro "FixActions". The maintenance actions are not formatted in the same way as the other road network data, and therefore a procedure has been written to match this information with the other network data that have been extracted

The processing of the WIM-data from the ATC-units

The Weigh-In-Motion (WIM) data are collected from a selection of the Automatic Traffic Control (ATC) units deployed on the Norwegian road network. Most of this data is recorded in a computer-unit located at the observation site, and then this data is downloaded to a central computer at the NPRA in Oslo. This has partly been done manually and partly by a mobile phone connection to the observation unit. The data (which contains each axle passage on each observation point over the observation period) is then compressed into a special compact file format that could be read by a dedicated program called Traffic. This program has then been used to extract readable text-files (*.grb) allowing further processing for my purposes. Once again the software (i.e. Traffic) has not been designed for analytic purposes, and extracting data is extremely cumbersome. Adding to these problems is the problem of handling the vast amount of information contained in these files. Fully extracted this amounts to more than 40 GB of text-files. Needless to say, the need for automated and efficient ways of handling this has been importunate. The macro "FixATKFil" was designed for this purpose, still a full run of this macro lasted for almost two full days!

External databases

Data processing (MS Excel / Visual Basic)

Road Data Bank: Roughness register (#12) Report 1

iri0100.sdv
…
iri1299.sdv

SamleIRIFiler i Arbeidsbok.xls
Macro: LagIRIBok

iri01.xls
…
iri18.xls

IRI.xls
Macro: FixIRIFil

IRI-Sluttfil.xls

Road Data Bank: Rut depth register (#10) Report 1

spor0100.sdv
…
spor1299.sdv

SamleSporFiler i Arbeidsbok.xls
Macro: LagSporBok

spor01.xls
…
spor18.xls

Spor.xls
Macro: FixSporFil

Spor-Sluttfil.xls

Road Data Bank: Traffic register (#3) Report 2

adt0100.sdv
…
adt1299.sdv

SamleADTFiler i Arbeidsbok.xls
Macro: LagADTBok

aadt01.xls
…
aadt18.xls

AADT.xls
Macro: FixADTFil

ADT-Sluttfil.xls

ATK-Sluttfil.xls

LSE-DATA.xls
Macro: FixLSEDatafil

Traffic (Bilfunn): ATC/WIM-data

b006????.grb
…
zn65????.grb

ATK.xls
Macro: FixATKFil

Road Data Bank: Pavement register (#17) Report 1 with physical sections

dekke01.txt
…
dekke18.txt

Actions.xls
Macro: FixActions

Actions-Sluttfil.xls

FAMAROW SPSS/PC LS-Estimation

LSE-Data-Sluttfil.xls

Appendix Figure 14 FAMAROW Data preparation files and procedures

The macro also converts the micro-information into information about the number of axle passages in each 1 tonne category, also discriminating between single and multiple axle configurations (definitions based on axle spacing). The macro also deals with some practical problems related to some files containing duplicates of some observation periods.

The final merger of the data

Finally, the macro "FixLSEDatafil" merges the information processed, and also includes data from the Norwegian Meteorological Institute (DNMI) on precipitation, snow-depths and temperatures. This data was delivered readily in MS Excel-format. Based on a manually computed table linking the various DNMI-stations to the closest WIM-observation site, this data was merged with the traffic and road network data described above. Two spreadsheets were computed, one for the roughness models, and one for the rut-depth models. The interval between observations of IRI and rut-depth defines the observation units (records) of the resulting database. The number of axle passages in each axle category is computed by applying the estimated AADT-models for establishing the absolute traffic level in the observation period, and then these figures are merged with the distribution on axle-categories extracted from the WIM-data. Any observation periods containing the date of a registered maintenance action are left out of the database. A proxy-variable for the use of studded tyres is computed: Winter-days without snow. Based on the 4th-power law, the number of Equivalent Standard Axle Loads (ESALs) is computed for each observation period. Equivalently ESAL-figures with deviating exponents is also added to each record.

Finally, the resulting two spreadsheets containing the databases to be used for the regression of rut-depth and roughness against traffic loads are imported in SPSS for further analysis and computation of derived variables.

APPENDIX C: RECOMMENDATIONS   FROM   THE   CAPRI PROJECT

The main conclusions from the project Concerted Action on Transport Pricing Research Integration (CAPRI) is reported in Nash et al. (2001). Since the focal point in this thesis is on pricing of roads, I will here present the recommendations from the CAPRI project that has direct relevance to this sector only[188]. The following sections are **direct quotations[189]** from *op.cit.*:

Recommendations for Pricing Principles – for Infrastructure use by all Modes

A1   **An understanding of marginal social costs should form the basis for the development of pricing policy** since under marginal social cost pricing, users pay the costs that they cause through additional infrastructure use, and thus economic efficiency is maximised.

A2   **Revenue contributions in addition to marginal costs may be justified** – to meet governments' and private operators' revenue needs, and to take account of equity considerations.

A3   **Prices based on short run marginal costs should incorporate all significant cost categories** – including:
   - *operating costs* (except those costs borne by the individual user);
   - *infrastructure wear and tear;*
   - *congestion costs* (except those costs borne by the individual user);

---

[188]   Thus leaving out the recommendations made for rail and air transport.

[189]   Apart from some formatting.

- *opportunity cost for the use of capacity* (when other users are displaced);
- *accidents* (except those costs borne by the individual user)
- *noise;*
- *air pollution; and,*
- *global warming.*

A4.   **Prices should vary more according to location and travel time** – reflecting the characteristics of the above cost categories. However, changes to existing mechanisms should only occur when the benefits from more variable pricing exceed the costs of implementing complex systems.

A5   **These principles should be applied evenly across all passenger and freight modes of transport** – and, in addition, in related sectors such as energy.

Recommendations on Valuation of Externalities

B1   **Externalities *within* the transport sector are of equal relevance as externalitites that are caused *outside* the transport sector** – it is the impact of one individual/organisation's behaviour on the other individuals/organisations that is important for pricing, not whether the cost is within or outside the transport sector as a whole.

B2   **All of the key externalities can be valued and incorporated in the development of pricing structures** – although substantial uncertainty exist in relation to cost estimation, in principle there is no reason to exclude any of the cost categories listed under "Pricing Principles".

B3   **Evidence of external benefits from increased private use of transport infrastructure is weak** – in contrast to public transport, where external benefits arise due to increased demand resulting in improved service levels to the benefit of other public transport users.

B4   **External costs of congestion, scarcity and accidents should be valued using willingness to pay approaches** – and it is essential that the internal element that the user already "pays" is separated from the price-relevant external element.

B5   **Valuation of air pollution is best undertaken by the impact pathway approach** – that incorporates emission, dispersion and dose-response relationships, with valuation of the final impacts on health, buildings, crops etc.

B6   **Human health impacts should be valued using the years of life lost approach** – and not the "value of a statistical life" method. Ideally, quality life years lost should be taken into account.

B7   **Regulatory policy may often be more powerful than pricing policy in the control or reduction of some categories of environmental emission** – particularly for aspect such as noise, where in some circumstances the marginal costs are very low.

B8   **At present there is no consensus on the values that should be placed on emissions of global warming gases** – thus, valued used in pricing should be based on political decisions about target emission levels.

Road Pricing – Urban and Inter-Urban

C1   **More differentiation in road charges by time period and area is necessary** – externalities are severe in congested urban and inter-urban situations where travel patterns are heavily peaked by time of day, day of week or season of the year.

C2   **Comparative analysis of the performance of road pricing versus the performance of other transport strategies is essential** – existing mechanisms or policies often have under-exploited potential. It is also important that the level of technological sophistication is also justified, relative to more basic road pricing schemes.

C3   **Low levels of political and public acceptability imply the need for a *staged* implementation strategy** – beginning with simple systems with low charge levels, and introducing more complex charging systems over time.

C4   **To increase acceptability, revenue from road pricing should be earmarked for specific spending programmes** – for activities such as public transport, additional infrastructure expenditure, improving environmental conditions in towns, etc. However, both economic viability and public support will be undermined if revenue raised is not spent wisely.

C5   **Increased choice can result in increased acceptability** – enhancing alternative modes in urban areas, retaining free parallel inter-urban routes, an allowing variation in charges by time of day all increase choice and improve acceptability.

C6   **Modelling studies for urban and inter-urban road pricing indicate that proposed price changes can induce small, but significant changes in behaviour** – small changes in behaviour can make a major contribution to the reduction of congestion and other externalities.

C7   **In contrast, demonstrations of urban road pricing often suggest that unacceptably large price changes may be needed to influence behaviour** – however, these magnitudes of response should be treated with caution as they may underestimate price sensitivities,- although these demonstrations have provided valuable evidence on behaviour, their short-term focus and use of compensation to volunteer participants who chose to use their car (as opposed to charging those who did) affects the results obtained.

C8   **The main impact of more variable road charging is likely to be travel at different times or by different routes by the same mode** – user's first preference will often be to continue to user their vehicle, but in a different way (different departure time, route, etc.).

Conclusions on Likely Impacts of Implementing Efficient Pricing

F1   **Pricing based on marginal costs may result in price *reductions* for some modes as well as price rises for others** – internalisation of externalities does not necessarily imply lower travel demand, or a shift to modes that are viewed as more environmentally sustainable – because current levels of taxation and charging have to be taken into consideration.

F2   **For inter-urban passenger travel in uncongested conditions, it is likely that road-based modes are over-priced** – due to the combination of

existing charging and taxation systems.

F3   **Inter-urban passenger travel by rail is often over-priced** – despite generally low taxation to account for externalities, current fares are often in excess of the marginal cost of providing additional services.

F4   **For inter-urban freight transport, evidence suggests that there is often significant under-charging for both road and rail-based modes** – evening uncongested conditions. The outcome of efficient pricing applied across modes is likely to be fairly neutral in terms of mode shares, however.

F5   **Urban transport by means of road-based modes is typically dramatically under-charged** – implying that efficient pricing will have the greatest impact in reducing externalities in urban areas.

F6   **The need for radical pricing reform has to be made on a case-by-case basis** – in over half the situations examined in European pricing projects the case for new pricing systems was not established, relative to the potential performance of existing pricing measures.

F7   **More variable pricing is likely to result in more dramatic changes *within* modes, than in switching of trips *between* modes** – greater differentiation of prices by time of day / vehicle or engine type / location is likely to change how users make use of their existing (and currently preferred) mode. For example, the dominant impacts of more variable charges by time of day / location are likely to be departure time and route adaptation, as opposed to switching between modes.

End of quotation.

APPENDIX F[190]: THE UNITE PREFERRED METHODOLOGY FOR THE CALCULATION OF MARGINAL EXTERNAL ACCIDENT COSTS

In this thesis the calculations of external accident costs are based on a model established by Rune Elvik in the early 1990s. This model provides *average external cost estimates*, and not genuinely *marginal external costs* which would be the proper ones to found a charging regime on. To calculate such costs one needs knowledge about the various *risk elasticities* (defined below). Such information is currently not available for the Norwegian road network, and it has been beyond the limits of this doctoral work to establish a model for estimating these elasticities. However, to illustrate how this could be done in the future, I will in this appendix briefly present the preferred methodology for the calculation of marginal external accident costs, as presented by the UNITE project in Bossche et al. (2001):

The total cost of an accident could be defined as the sum of three components, a, b, and c. The first component (a) is the value of a statistical lifte (VOSL). The second (b) is the equivalent figure for relatives and friends of the victim, and cost component c comprises the costs for the rest of society (mainly material costs). If there is A accidents in one year, then the total accident costs could be written:

Equation F - 1 $\qquad TC_{accident} = A(a + b + c)$

The marginal cost with respect to traffic volume (Q) is then

---

[190] As noted in the introduction to the Appendices, Appendix D and Appendix E are published in separate binders.

Equation F - 2 $$MC_{accident} = \frac{dA}{dQ}(a+b+c)$$

Marginal external costs (MEC) equals marginal cost minus marginal private cost (MPC):

Equation F - 3 $$MEC_{accident} = MC_{accident} - MPC_{accident}$$

The MPC will be different for a victim (v) and an injurer (i). The former may internalise cost component a, and possibly also b, while the latter will have no marginal private costs (assuming he will not be charged for liabilities). The users are divided into victims (with traffic volume Q_v) and injurers (with traffic volume Q_i). The victim's risk is π (A/Q_v) and the injurer's risk is θ (A/Q_i). We are now able to define the corresponding *risk elasticities*, i.e. the relative change in risk resulting from a relative change in traffic:

Equation F - 4 $$E_v = \frac{d\pi}{dQ_v} \cdot \frac{Q_v}{\pi}, \text{ and } E_i = \frac{d\theta}{dQ_i} \cdot \frac{Q_i}{\theta}$$

Using these elastiticities, the expressions for marginal external costs of the victim and the injurer becomes:

Equation F - 5

$$MEC_{accident,v} = \pi(1+E_v)(a+b+c) - \pi(a+b) = \pi E_v(a+b+c) + \pi c$$

$$MEC_{accident,i} = \theta(1+E_i)(a+b+c) - 0 = \theta E_i(a+b+c)$$

A road user may be both a victim and an injurer. If we allocate a probability β for becoming a victim and (1- β) for becoming an injurer, the expected marginal cost becomes

338

Equation F - 6

$$MEC_{accident} = \beta\pi(1+E_i)(a+b+c)+(1-\beta)\theta(1+E_i)(a+b+c)+\beta\pi c$$

This equation forms a general model for calculating the marginal external accident cost of road users. However, one might also argue that the cost of risk avoiding behaviour should be added to this. Taking this into account, *op. cit.* concludes that *Equation F - 6 represents a lower bound* for the external marginal accident cost, and that a pure risk elasticity approach would represent the *upper bound* of the estimate:

Equation F - 7 $\qquad MEC_{accident,i} = \dfrac{dA}{dQ_i}(a+b+c)$

This means that the preferred methodology for estimating marginal accident costs comprises the following steps:

- Estimation of the risk of the injurer and the victim
- Estimation of the relationship between traffic volume and accident frequency A(Q); and calculate the marginal increase of the expected number of accidents according to this relationship dA/dQ, i.e. the *risk elasticity*.
- Evaluate the monetary value of these changes dA/dQ (a+b+c)
- Estimate the parts of this added cost that are internal and external. The difference between the marginal accident cost and the internal/private cost gives the external marginal accident cost.

REFERENCES

Archilla, A. R., and Madanat, S. (2000). "Development of a Pavement Rutting Model from Experimental Data." *Journal of Transportation Engineering*, 2000(July/August), 291-299.

Argoul, P., Sainte-Marie, J., Jacob, B., and Dolcemascolo, V. "Multiple Sensors WIM Using Reconstruction Algorithms of the Dynamic Axle Loads." *2nd European Conference on Weigh-In-Motion of Road Vehicles*, Lisbon, Portugal.

Bang, J., Flugsrud, K., Holtskog, S., Haakonsen, G., Larssen, S., Maldum, K. O., Rypdal, K., and Skedsmo, A. (1999). "Utslipp fra veitrafikk i Norge - Dokumentasjon av beregningsmetode, data og resultater." *99:04*, Statens Forurensingstilsyn (SFT), Oslo.

Bastiaans, R. (1998a). "PARIS Deliverable D4: Pavement Deterioration Models." VTT, Helsinki.

Bastiaans, R. (1998b). "Pavement Deterioration Models." *Paris/WP V/01/Final*, VTT, Helsinki.

Baumol, W. J., and Bradford, D. F. (1970). "Optimal departures from marginal cost pricing." *American Economic Review*, 60, 265-83.

Begg, D. "Road user charging - the way forward." *Imprint-Europe, 3rd seminar*, Brussels.

Bickel, P., and Schmid, S. (2002a). "Marginal Cost Case Study 9D: Urban Road and Rail Case Studies Germany. Appendix 9D to Deliverable 11 from the UNITE project." University of Leeds, Leeds.

Bickel, P., and Schmid, S. (2002b). "Marginal Cost Case Study 9E: Inter-Urban Road and Rail. Case studies Germany. Appendix 9E to Deliverable 11 of the UNITE project." University of Leeds, Leeds, UK.

Bohm, P. (1987). *Social Efficiency. A concise introduction to welfare economics*, MacMillan, London.

Bossche, M. v. d., Certan, C., Goyal, P., Gommers, M., and Sansom, T. (2001). "Marginal Cost Methodology. Deliverable 3 of the UNITE project." University of Leeds, Leeds. UK.

Calthrop, E., and Proost, S. (1998). "CAPRI Deliverable 2: General Economic Principles of Pricing Transport Services." University of Leeds, Leeds.

Capros, P., and Mantzos, L. (2000). "Kyoto and technology at the European Union; Costs of emission reduction under flexibility mechanisms and technology progress." *International Journal of Global Energy Issues*, 14, 169-183.

Carlsson, A., Centrell, P., and Öberg, G. (1995). "Dubbdäck, Samhällsekonomiska konsekvenser." *756-1995*, VTI, Linköping.

Cebon, D., Stergioulas, L. K., and MacLeod, M. D. "Enhancing Multiple -Sensor WIM Systems." *2nd European Conference on Weigh-In-Motion of Road Vehicles*, Lisbon, Portugal.

Coase, R. H. (1960). "The Problem of Social Cost." *Journal of Law and Economics*, 3.

COST-323. (1997). "Collection and Analysis of Requirements as Regards Weighing Vehicles in Motion. First version (revised)." COST-323.

Coste, J.-F. "Conclusions on the Scientific Achievements of the Conference." *2nd European Conference on Weigh-In-Motion of Road Vehicles*, Lisbon, Portugal.

De Borger, B., Peirson, J., and Vickerman, R. (2001). "An overview of policy instruments." Reforming Transport Pricing in the European Union, B. De Borger and S. Proost, eds., Edward Elgar, Cheltenham, UK.

De Borger, B., and Proost, S. (2001). *Reforming Transport Pricing in the European Union. A modelling approach.*, Edward Elgar, Cheltenham, UK / Northamton, MA, USA.

De Kluizenaar, Y., Passchier-Vermeer, W., and Miedema, H. (2001). "Adverse effects of noise exposure to health. Report prepared for the EC Project UNITE." TNO, Leiden, NL.

Dolcemascolo, V., and Jacob, B. "Multiple Sensor Weigh-In-Motion: Optimal Design and Experimental Study." *2nd European Conference on Weigh-In-Motion of Road Vehicles*, Lisbon, Portugal.

EC, H. (1999a). "Final Report on Estimating Transport Costs." EC, The High Level Group on Transport Infrastructure Charging, Brussels (?).

EC, H. (1999b). "Final Report on Options for Charging Users Directly for Transport Infrastructure Operating Costs." EC, The High Level Group on Transport Infrastructure Charging.

EC, H. L. G. (1998). "Final Report." EC, The High Level Group on Transport Infrastructure Charging, Brussels (?).

EC-DG7. (1997). "COST 324 Long Term Performance of Road Pavements. Final Report of the Action." EC, Brussels.

EC-DG7. (1998). "Fair Payment for Infrastructure Use: A phased approach to a common transport infrastructure charging framework in the EU." *COM(1998) 466 final*, EC, Brussels.

EC-DG7. (1999). "PARIS, Performance analysis of road infrastructure. Final report." Office of Official Publications of the European Communities, Luxembourg.

ECMT. (1998). "Efficient Transport for Europe. Policies for internalisation of external costs." European Conference for Ministers of Transport, Paris.

Elvik, R. (1993). "Økonomisk verdsetting av velferdstap ved trafikkulykker. Dokumentasjonsrapport." *203/1993*, TØI, Oslo.

Elvik, R. (1994). "The external costs of traffic injury: Definition, estimation, and possibilities for internalization." *Journal of Accident Analysis & Prevention*, 26(6), 719-732.

Enei, R., Leone, G., Schmid, S., and Bickel, P. (2002). "Marginal Cost Case Study 9h: Inter-Urban road and rail. Appendix 9h to Deliverable 11 of the UNITE project." University of Leeds, Leeds, UK.

Eriksen, K. S. (2000). "Calculating External Costs of Transportain in Norway." *European Journal of Transport and Infrastructure Research*, 0(0).

Eriksen, K. S., Markussen, T. E., and Pütz, K. (1999). "Marginale kostnader ved transportvirksomhet." *464/1999*, Transportøkonomisk Institutt, Oslo.

Fosser, S., and Sætermo, I. A. (1995). "Vinterdekk med eller uten pigger - betydning for trafikksikkerheten." *310/1995*, Transportøkonomisk Institutt (TØI), Oslo.

Fridstrøm, L. (1999a). "Econometric models of road use, accidents, and road investment decisions, Volume 1." *456/1999*, Institute of Transport Economics (TØI), Oslo.

Fridstrøm, L. (1999b). "Econometric models of road use, accidents, and road investment decisions, Volume 2." *457/1999*, Institute of Transport Economics (TØI), Oslo.

Fridstrøm, L. (2000). "On the Theory of Marginal Cost Road Pricing, With Particular Attention to Accident Externalities." University of Oslo, Oslo.

Friedrich, R., and Bickel, P. (2001). *Environmental External Costs of Transport*, Springer-Verlag, Berlin.

Giblin, J.-P. "Conclusions and Future Perspectives." *Final Symposium of the project WAVE (1996-99)*, Paris, France.

Glazer, A., Link, H., May, T., Milne, D., and Niskanen, E. "Barriers to transport pricing - review of research." *Dissemination seminar of the UNITE project*, Brussels.

Grue, B., Larsen, O. I., Rekdal, J., and Tretvik, T. (1997). "Congestion Cost and Congestion Pricing in Urban Traffic." *363/1997*, Institute of Transport Economics (TØI), Oslo.

Hagen, K.-E. (1993). "Samfunnsøkonomisk regnskapssystem for trafikkulykker og trafikksikkerhetstiltak." *182/1993*, TØI, Oslo.

Hajek, J. J., Tighe, S. L., and Hutchinson, B. G. (1998). "Allocation of Pavement Damage Due to Trucks Using a Marginal Cost Method." *Transportation Research Record*, No. 1613.

Hallström, B. "Durability Implications of Cold Climates on WIM systems." *Final Symposium of the project WAVE (1996-1999)*, Paris, France.

Haugødegård, T. (1993). "ALFRED test. Utprøving av måleutstyr for spor og jevnhet." Statens Vegvesen, Vegdirektoratet, Vedlikeholdskontoret, Oslo.

Haugødegård, T. (1995). "Jevnhet på langs på nylagte vegdekker. ALFRED-systemet: Målenøyaktighet på nylagte vegdekker." Statens Vegvesen, Vegdirektoratet, Vedlikeholdskontoret, Oslo.

Haukeland, J. V. (1991a). "Trafikkskadd - hva så?" *0092/1991*, TØI, Oslo.

Haukeland, J. V. (1991b). "Velferdstap ved trafikkulykker." *0091/1991*, TØI, Oslo.

Henau, A., and Jacob, B. "European Test Programme of WIM systems: Cold Environment Test & Contintental Motorway Test." *2nd European Conference on Weigh-In-Motion of Road Vehicles*, Lisbon, Portugal.

Hensher, D. A. (1994). "Stated Preference Analysis of Travel Choices: The State of Practice." *Transportation*, 21.

343

Herry, M. "Barriers to Interurban Transport Pricing." *The First Meeting of the IMPRINT-Europe Thematic Network*, Brussels.

Hesselborn, P.-O. (2000). "Översyn av förutsättningarna för marginalkostnadsbaserade avgifter i transportsystemet." *2000:10*, SIKA, Stockholm.

Hesselborn, P.-O. (2001). "Trafikens Externa Effekter." *2001:7*, SIKA, Stockholm.

Highway Research Board, N. A. o. S. (1962a). "The AASHO Road Test, Report 5 Pavement Research." The Highway Research Board, Washington D. C.

Highway Research Board, N. A. o. S. (1962b). "The AASHO Road Test, Report 7 Summary Report." The Highway Research Board, Washington D. C.

Hjelle, H. M. (1992). *En modell for bilisters veivalg*, Universitetet, Oslo.

Hjelle, H. M. (2000). "New technology for Weigh In Motion (WIM) of road vehicles : a comparative literature study with focus on applications related to road wear cost allocation and road pricing." *2000:1*, Molde University College, Molde.

Hjelle, H. M. (2003a). "A Foundation of Road User Charges. Appendix D: Visual Basic macros for the FAMAROW model." The Norwegian University of Science and Technology (NTNU). Department of Civil and Transport Engineering., Molde/Trondheim.

Hjelle, H. M. (2003b). "A Foundation of Road User Charges. Appendix E: Output from the CATERU-model. Scenario A-D." The Norwegian University of Science and Technology (NTNU). Department of Civil and Transport Engineering., Molde/Trondheim.

Holtskog, S. (2001). "Direkte energibruk og utslipp til luft fra transport i Norge, 1994 og 1998." *2001/16*, Statistics Norway, Oslo - Kongsvinger.

Holtsmark, and Hagem. (1998). "Emission Trading under the Kyoto Protocol." *1998-01*, CICERO, Oslo.

Huhtala, M., Halonen, P., and Miettinen, V. "Cold Environemental Test at Luleå; Calibration of WIM systems Using an Instrumented Vehicle." *2nd European Conference on Weigh-In-Motion of Road Vehicles*, Lisbon, Portugal.

Jacob, B., and O'Brien, J. "European Specification on Weigh-In-Motion of Road Vehicles." *2nd European Conference on Weigh-In-Motion of Road Vehicles*, Lisbon, Portugal.

Jacobsen, T., and Hornwall, F. (1999). "Wear from studded tyres on asphalt pavements. Summary of results from test roads and control sections 1990-98." *862*, Swedish National Road and Transport Research Institute (VTI), Linköping.

Jansson, J. O. (1994). "Accident Externality Charges." *Journal of Transport Economics and Policy*, 1994(January).

Jehaes, S. "Tests of WIM systems in Cold Climates." *Final Symposium of the project WAVE (1996-99)*, Paris, France.

Jehaes, S., and Hallström, B. "Accuracy Analysis of WIM systems for the Cold Environment Test." *2nd European Conference on Weigh-In-Motion of Road Vehicles*, Lisbon, Portugal.

344

Jensen, T. (1998). "Alternativ transportteknologi. Reduserte utslipp fra transportsektoren." *413/1998*, Institute of Transport Economics (TØI), Oslo.

Johansson, M. V. (1999). "Economics Without Markets. Four papers on the Contingent Valuation and Stated Preference Methods.," Umeå Economic Studies No. 517, Umeå University, Umeå, Sweden.

Jones, J., and Nix, F. P. (1995). "Survey of the Use of Highway Cost Allocation in Road Pricing Decisions." *3*, Transportation Association of Canada, Ottawa, Canada.

Jones-Lee, M. W. (1994). "Safety and the saving of life: The economics of safety and physical risk." Cost-benefit analysis, R. Layard and S. Glaister, eds., Cambridge University Press, Cambridge, UK.

Lindberg, G. "External Accident Cost of Heavy Goods Vehicles." *UNITE Seminar*, Paris.

Lindberg, G. (2001b). "Traffic Insurance and Accident Externality Charges." *Journal of Transport Economics and Policy*, 35(3), 399-416.

Lindberg, G. "Implementing interurban road pricing reforms; Concensus and constraints (Draft)." *Imprint-Europe, 3rd seminar*, Brussels.

Lindberg, G. (2002b). "UNITE Deliverable 9: Infrastructure Cost Case Studies, Case Study 5b: Marginal cost of road maintenance for heavy goods vehicles on Swedish roads. Preliminary draft." University of Leeds, ITS, UK / VTI, Sweden, Leeds.

Link, H., Dodgson, J. S., Maibach, M., and Herry, M. (1999). *The Costs of Road Infrastructure and Congestion in Europe*, Physica-Verlag, Heidelberg, New York.

Link, H., and Lindberg, G. (2000). "UNITE Interim Report 5.3. Marginal Cost Methodology for Infrastructure Costs." University of Leeds, ITS, UK / DIW, Germany / VTI, Sweden, Leeds.

Mayeres, I., Proost, S., Quinet, E., Schwartz, D., and Sessa, C. (2001). "Alternative Frameworks for the Integration of Marginal Costs and Transport Accounts." *4*, University of Leeds, Leeds.

Metronomica. (2001). "Monetary valuatio of noise effects. Draft Final Report prepared for the EC UNITE project." Bath, UK.

Nash, C. A., and Johnson, D. "The measurement of marginal social cost - lessons from the UNITE project." *Imprint-Europe, 3rd seminar*, Brussels.

Nash, C. A., and Matthews, B. "Why reform transport prices? A review of European research." *IMPRINT-EUROPE Key requirements for Implementing Pricing Reform in Transport*, Bruxelles.

Nash, C. A., Sansom, T., and Matthews, B. (2001). "Final Report for Publication. Concerted Action on Transport Pricing Research Integration (CAPRI)." *ST-97-CA-2064*, Institute for Transport Studies, University of Leeds, UK, Leeds.

Nellthorp, J., Sansom, T., Bickel, P., Doll, C., and Lindberg, G. (2001). "Valuation convensions for UNITE (Unification of accounts and marginal costs for Transport Efficiency)." University of Leeds, Leeds.

Newbery, D. M. (1985). "The Fundamental Theorem of Road User Charges." *No. 59*, Centre for Economic Policy Research, London.

Newbery, D. M. (1988a). "Charging for Roads." *The World Bank Research Observer*, 3(2).

Newbery, D. M. (1988b). "Road Damage Externalities and Road User Charges." *Econometrica*, 56(2), 295-316.

Newbery, D. M. (1988c). "Road User Charges in Britain." *The Economic Journal*, 98.

Nielsen, O. A., and Herslund, M.-B. "The AKTA Road Pricing Experiment in Copenhagen." *European Transport Conference*, Cambridge, UK.

NPRA. (1998). *Standard for drift og vedlikehold 1999. Retningslinjer.*, The Norwegian Public Roads Administration, Oslo.

OECD, I. D. p. "Dynamic Interaction Vehicle-Infrastructure (DIVINE)." *DIVINE European Concluding Conference*, Paris.

OECD, R. T. R. (1988). "Heavy Trucks, climate and pavement damage." OECD, Paris.

OFV. (2002). "Bil og Veistatistikk, Nettutgave." Opplysningsrådet for Vegtrafikken.

Paterson, W. D. O. (1984). "Summary of Distress Prediction Models from Major Studies: Brazil, Kenya, Arizona and Texas." World Bank Transportation Department, Baltimore and London.

Paterson, W. D. O. (1987). "Road Deterioration and Maintenance Effects. Models for Planning and Management." *10083*, The World Bank, Baltimore and London.

Pearce, D. W., and Turner, R. K. (1990). *Economics of Natural Resources and the Environment*, Harvester Wheatsheaf, New York.

Rennings, K., Ricci, A., Sessa, C., and Weinreich, S. (1999). "CAPRI Deliverable D3: Valuation of Transport Externalities." University of Leeds, Leeds.

Rideng, A. (2001). "Transportytelser i Norge 1946-2000." *515/2001*, Transportøkonomisk Institutt (TØI), Oslo.

Rosendahl, K. (1999). "Vurdering av skadekostnadsmetoden til bruk på vegprosjekt - en case studie." *1999/5*, Statistics Norway, Oslo.

Roy, R. (2000). "Revenues from Efficient Pricing: Evidence from the Member States." EC, DG TREN, London.

Sansom, T., Nellthorp, J., Proost, S., Mayeres, I., Maibach, M., Niskanen, E., Quinet, E., and Schwartz, D. (2000). "The Overall UNITE Methodology. Deliverable 1 from the UNITE project." University of Leeds, Leeds, UK.

Sayers, M. W. (1995). "On the Calculation of IRI from Longitudinal Road Profile." *1501*, Transportation Research Board, Washington D.C.

Sayers, M. W., and Karamihas, S. M. (1998). *The Little Book of Profiling. Basic Information about Measuring and Interpreting Road Profiles.*, The University of Michigan, Michigan.

346

Senstad, P. (1994). "Sluttrapport for etatsatsingsområdet: Bedre utnyttelse av vegens bæreevne." *75*, Statens Vegvesen. Veglaboratoriet, Oslo.

Small, K. A., and Winston, C. (1988). "Optimal Highway Durability." *American Economic Review*, 78(3).

Small, K. A., Winston, C., and Evans, C. A. (1989). *Road Work. A new Highway Pricing & Investment Policy*, The Brookings Institution, Washington D. C.

(SSB), S. N. (2000). "Road Goods Transport - Sample Survey 1996-1998." Statistics Norway (SSB), Oslo.

Sulkjær, P. "AKTA - project about Road Pricing in Copenhagen." *Imprint-Europe. 2nd seminar*, Brussels.

Sund, E. K. (2001). "DVE-prosjektet - Vegdekker (HIPS). Sluttnotat." *Project nr. 22G113*, SINTEF, Trondheim.

Suter, S., Sommer, H., Marti, M., Wickart, M., Schreyer, C., Peter, M., Gehrig, S., Maibach, M., Wühtrich, P., Bickel, P., and Schmid, S. (2002). "UNITE Deliverable 5, Appendix 2: The Pilot Accounts for Switzerland." University of Leeds, Leeds.

Sælensminde, K., and Hammer, F. (1994). "Verdsetting av miljøgoder ved bruk av samvalgsanalyse." *251/1994*, Transportøknomisk Institutt (TØI), Oslo.

Taylor, B., and Bergan, A. (1993). "The Use of Dual Weighing Elements (Double Threshold) to Improve the Accuracy of Weigh-In-Motion Systems, and the Effect and Accuracy on Weigh Station Sorting." Oregon Department of Transport and International Road Dynamics Inc., Oregon.

Tervonen, J., Hämekoski, K., Otterström, T., Anton, P., Bickel, P., and Schmid, S. (2002a). "Marginal Cost Case Study 9a: Urban Passenger Car for Finland. Appendix 9a to Deliverable 11 from the UNITE project." University of Leeds, Leeds, UK.

Tervonen, J., Hämekoski, K., Otterström, T., Anton, P., Bickel, P., and Schmid, S. (2002b). "Marginal Cost Case Study 9b: Heavy Goods Vehicle for Finland. Appendix 9b to Deliverable 11 of the UNITE project." University of Leeds, Leeds, UK.

Teule, O. "Pay per kilometer. Progress report." *Imprint-Europe, 2nd Seminar*, Brussels.

US Department of Transportation, F. H. A. (1997a). "1997 Federal Highway Cost Allocation Study." Washington D. C.

US Department of Transportation, F. H. A. (1997b). "States' Successful Practices Weigh-In-Motion Handbook."

US Department of Transportation, F. H. A. (2000). "Addendum to the 1997 Federal Highway Cost Allocation Study. Final report." Federal Highway Administration, Washington D.C.

Verhoef, E. T. (1994). "External Effects and Social Costs of Road Transport." *Transportation Research A*, 28A(4).

Verhoef, E. T. (2000). "The implementation of marginal external cost pricing in road transport. Long run vs. short run and first-best vs. second-best." *Regional Science*, ?(?).

Verhoef, E. T. "Marginal Cost Based Pricing in Transport. Key Implementation Issues from the Economic Perspective." *IMPRINT 1st seminar: Key Requirements for Implementing Pricing Reform in Transport.*, Bruxelles.

Vougioukas, M., Clement, L., and Monsigny, M. (1999). "CAPRI Deliverable 4: Road Transport Pricing Issues with particular reference to Urban Road Pricing." University of Leeds, Leeds.

Vougioukas, M., Clement, L., and Monsigny, M. (2000). "CAPRI Final Report. Annex C: Road Transport Pricing Issues." University of Leeds, Leeds.

Walters, A. A. (1968). "The Economics of Road User Charges." *World Bank Staff Occational Papers no. 5*, The World Bank, Baltimore and London.

Winston, C. (1985). "Conceptual Developments in the Economics of Transportation: An Interprative Survey." *Journal of Economic Literature*, XXIII.