ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Schmidtchen, Dieter; Koboldt, Christian; Monheim, Jenny; Will, Birgit E.; Haas, Georg

Working Paper

The internalisation of external costs in transport: from the polluter pays to the cheapest cost avoider principle

CSLE Discussion Paper, No. 2007-03

Provided in Cooperation with:
Saarland University, CSLE - Center for the Study of Law and Economics

Suggested Citation: Schmidtchen, Dieter; Koboldt, Christian; Monheim, Jenny; Will, Birgit E.; Haas, Georg (2007) : The internalisation of external costs in transport: from the polluter pays to the cheapest cost avoider principle, CSLE Discussion Paper, No. 2007-03, Universität des Saarlandes, Center for the Study of Law and Economics (CSLE), Saarbrücken

This Version is available at:
https://hdl.handle.net/10419/28012

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

CSLE Discussion Paper 2007-03

The Internalisation of External Costs in Transport: From the Polluter Pays to the Cheapest Cost Avoider Principle

December 2007

Prof. Dr. Dieter Schmidtchen
Dr. Christian Koboldt
Dr. Jenny Monheim
Diplom-Kauffrau Birgit E. Will
Diplom-Kaufmann Georg Haas

Center for the Study of
Law and Economics

Center for the Study of Law and Economics
http://www.uni-saarland.de/fak1/fr12/csle/
Saarland University,
Campus C 3.1
66123 Saarbrücken, Germany
Tel: +49 (0)681-302-4387
Fax: +49(0)681-302-3591,
email: csle@rz.uni-saarland.de

Center for the Study of Law and Economics

Contents

Executive Summary **v**

Introduction **1**

1 A critical assessment of the CE study **7**

1.1 Introduction and background 7

1.2 The CE study is based on a flawed notion of external costs 15

1.3 The CE study lacks a coherent framework assessing external costs 21

1.4 The CE study lacks a coherent framework for internalisation 29

2 The polluter pays vs. the cheapest cost avoider principle **43**

2.1 Introduction 43

2.2 Coase versus Pigou 44

2.3 Paradigm shift 69

2.4 The logic of the "cheapest cost avoider" principle 76

2.5 Comparative Institutions: the polluter pays principle 86

2.6 Comparative Institutions: concerns about the application of the cheapest cost avoider principle 97

3 Case Studies **111**

3.1 Introduction 111

3.2 Methodological considerations 112

3.3 The case of the missing A44 114

3.4 The sectoral ban on the Inn valley motorway 121

Conclusions **127**

Recommendations **130**

References **131**

Appendix **137**

Contents

Executive Summary **v**

Introduction **1**

1 A critical assessment of the CE study **7**

1.1 Introduction and background 7

1.1.1 A brief summary of the CE study 8

1.1.2 Estimation of external costs 8

1.1.3 Scenarios for internalisation 11

1.2 The CE study is based on a flawed notion of external costs 15

1.2.1 The economic notion of external effects and external costs 16

1.2.2 Externalities as a reciprocal problem and the cheapest cost avoider: a simple stylised example 17

1.3 The CE study lacks a coherent framework assessing external costs 21

1.4 The CE study lacks a coherent framework for internalisation 29

1.4.1 What is internalisation? 30

1.4.2 The objectives of internalisation in the CE study are unclear 33

1.4.3 The CE study does not provide a comparative assessment of internalisation tools 35

1.4.4 Second best issues, implementation costs and regulatory failure 37

1.4.5 Interrelation with other policy instruments 40

2 The polluter pays vs. the cheapest cost avoider principle **43**

2.1 Introduction 43

2.2 Coase versus Pigou 44

2.2.1 The Pigovian tradition and the polluter pays principle 44

2.2.2 The Coase theorem: optimality without government intervention 50

2.2.3 A simple illustration 52

2.2.4 The zero transaction costs scenario 52

2.2.5 The high transaction costs scenario 56

2.2.6 The Coase theorem, Pigovian taxes and welfare with zero transaction costs 57

2.2.7 Pigovian taxes and welfare with high transaction costs 63

2.2.8 The joint cause of damage 68

2.3 Paradigm shift 69

2.3.1 There is no negative externality without rivalry 70

2.3.2 The reciprocal nature of the externality problem 71

2.3.3 Some externalities are the result of exclusive rights 71

Contents

2.3.4 It is socially desirable to eliminate only the Pareto relevant negative externalities 72

2.3.5 All harm is jointly caused 73

2.3.6 Efficiency requires double responsibility at the margin 74

2.4 The logic of the cheapest cost avoider principle 76

2.4.1 Cheapest cost avoidance: care as a strictly binary variable 77

2.4.2 Cheapest cost avoidance: care as a continuous variable 79

2.4.3 Putting the cheapest cost avoider principle into practice 81

2.5 Comparative Institutions: the polluter pays principle 86

2.5.1 The polluter pays principle and corrective justice 87

2.5.2 The polluter pays principle and distributive justice (fairness) 93

2.5.3 The polluter pays principle and fairness between transport modes 94

2.5.4 The polluter pays principle and administrative costs 95

2.6 Comparative Institutions: concerns about the application of the cheapest cost avoider principle 97

2.6.1 Valuing non pecuniary factors such as life, pain and suffering 97

2.6.2 "Soft" variables 99

2.6.3 Empirical data concerns 99

2.6.4 Fairness issues 99

2.6.5 Rational individual welfare maximisation 101

2.6.6 The value of the protection of nature 102

2.6.7 The interests of future generations 103

2.6.8 Costs of the cheapest cost avoider principle 105

2.6.9 Social constraints 107

3 Case Studies 111

3.1 Introduction 111

3.2 Methodological considerations 112

3.3 The case of the missing A44 114

3.3.1 Cheapest cost avoider analysis 114

3.3.2 Legal reasoning on the B7 ban 120

3.4 The sectoral ban on the Inn valley motorway 121

3.4.1 Cheapest cost avoider analysis 121

3.4.2 The reasoning of the European Court of Justice 124

Conclusions 127

Recommendations 130

Contents

References **131**

Appendix **137**

Executive Summary

In March 2000, in what has come to be known as the "Lisbon strategy" European leaders committed the EU to become, by 2010, the most dynamic and competitive knowledge-based economy in the world capable of sustainable economic growth with more and better jobs and greater social cohesion, and respect for the environment. Achieving these aspirations requires a careful and comprehensive analysis and assessment of likely economic, social and environmental impacts, both direct and indirect, of all regulatory measures. Regulatory impact assessments, as part of the Better Regulation Agenda, are a key tool for the European institutions in designing better policies and laws. By providing a detailed study of the economic, social and environmental impacts, this procedure helps the European Union to reach the goals of Better Regulation, and its underlying Lisbon and Sustainable Development strategies.

As a follow-up to the Lisbon statement, the European Council concluded during its meeting in Göteborg (2001) that a sustainable transport policy should tackle rising volumes of traffic and levels of congestion, noise and pollution, and encourage the use of environment-friendly modes of transport as well as the internalisation of social and environmental costs. At the heart of this sustainable transport policy should be a fair, transparent and efficient system of charging for all modes of transport.

Given this mandate, and required by the upcoming revision of the Eurovignette Directive (2006/38/EC) and more generally by the EU "Better Regulation" initiative, the European Commission recently launched the debate on how to internalise external costs caused by all modes of transport.

As a first step, the European Commission mandated the consulting firm CE Delft to develop a *framework for the internalisation* of external costs and to devise a number of potential *internalisation scenarios* for further analysis. The results of their preliminary research are published in a CE discussion paper, and the Commission is now formulating draft legislative proposals which will be subjected to a consultation of stakeholders.

The present report was commissioned by IRU from the Center for the Study of Law and Economics, University of the Saarland, in order to contribute to the ongoing debate about the appropriate treatment of external costs in the transport sector. More specifically, we have been asked to:

- offer a critical assessment of the above mentioned CE study;
- provide a theoretical examination of the relative strengths and weaknesses of the polluter pays principle and more modern concepts such as the cheapest cost avoider principle in relation to the problems arising in the road transport sector;
- demonstrate how the cheapest cost avoider principle can be *applied in practice* by providing examples of external costs in road transport; and to
- put forward *recommendations* for the appropriate principles that should guide the Commission's further activities in this area.

The results of our study are grouped into four corresponding parts which are now discussed in turn.

Part I: A critical assessment of the CE study

1. The CE study addresses two distinct issues, namely:

- how **external costs** are to be defined and measured for various modes of transport with regard to a number of potential sources of external costs; and
- why and how external costs should be internalised, what problems arise with such **internalisation** in practice, and what internalisation scenarios should be further considered.

Reliable and robust answers to both of these questions are essential if the Commission wishes to adopt policies that promote economic growth and address problems of growing congestion and pollution in an equitable and efficient manner – or, in short, policies that are required by the pursuit of the Lisbon agenda.

Having carefully studied the CE report from a scientific point of view, the following assessment can be made:

2. Although crucial for any attempt to estimate external costs or assess the most appropriate way of internalisation, **a proper definition of external costs and a consistent and coherent discussion of external effects are completely missing from the CE study. It is long on assertion and short on analysis. Subsequently the CE lacks a sound and reliable basis for the evaluation of potential internalisation scenarios and policy options.**

3. The CE study does not take into account that both the measurement of external costs and the choice of an appropriate internalisation strategy crucially depend on the insight that externality problems are reciprocal, and that external costs are caused jointly by all parties involved.
The CE study endorses a naive view of external cost as being caused by one particular party – the polluter – and suppresses the more sophisticated understanding of external costs as costs arising from competing demands for scarce resources, which is now a generally accepted view in economics. It neglects the basic insight that external costs are caused jointly by all the parties involved and that the problem is of a reciprocal nature; avoiding pollution and thus improving the pollutee's situation would inflict harm on the polluter.
In the CE study, the question is thought of as one in which a polluter inflicts harm on a pollutee; what has to be decided is how to restrain the polluter. This is the logic underpinning the polluter pays approach. It obscures the nature of the choice that has to be made. The real question that has to be decided is: Should the polluter be allowed to harm the pollutee, or should the pollutee have the right to restrain the polluter? The aim should be to avoid the most serious harm; its solution is the cheapest cost avoider principle.

4. Efficiency can require charging a polluter, a pollutee or even that both parties pay. In brief, no single party can enjoy an automatic immunity from this joint responsibility as implied by the polluter pays principle.
The question how one should most appropriately deal with external effects is more complex than simply asking – as the CE study does – how much the party "causing" the external effect should be paying (and on what particular variables the payable amount should depend).

5. It seems inappropriate, and potentially dangerous, to rely on the cost estimates presented in the CE study as a basis for internalisation.
The CE study lacks a coherent framework for measuring external costs, or

assessing external cost estimates prepared by others. The CE study does not explore the potential reasons for the differences in cost estimates. Neither does it sufficiently take into account the **enormous variation in the presented numbers**, some examples of which are:

- The CE study quotes results form a UIC study, which estimates the total **external costs of road transport** as € 650 billion whereas a UNITE study estimates costs at € 129 billion.

- Regarding accident costs, the CE study admits the under-reporting of fatalities and injuries in official statistics, as well as substantial differences in the **valuation of lives** lost. Estimates for the Value of a Statistical Life (VoSL), which are used in the estimation of external costs range from less than US$200,000 to US$30m.

- Maximal air pollution and climate change externality costs present respectively 500% and 383% of the minimal values.

- Establishing the costs of **climate change** is complex owing to the long term effect of greenhouse gas emissions and to the difficulty in anticipating risk patterns. The share of climate change in generating natural disasters is difficult to establish.

Even though one of the declared objectives of the CE study is the collection of such third-party cost estimates, it is unclear how the various estimates can be compared and used in an internalisation framework without an appropriate methodology for the measurement of external costs.

6. The CE study fails to provide any reliable guidance with regard to the choice of potential internalisation scenarios, let alone help in relation to the assessment of their relative costs and benefits. This is because the CE study does not put forward a consistent set of criteria against which one could judge the relative merits of different internalisation options. It lacks a coherent framework for internalisation.

7. The objectives of internalisation in the CE study are unclear. Clarity about policy objectives cannot be obtained by listing the different motives for internalisation, and then – as the CE study does – suggesting that in practice the objectives underlying internalisation may be a bit of everything. The potential objectives may not only be poorly aligned, but may actually be conflicting.

8. The CE study does not provide a comparative assessment of internalisation tools.
What is missing is a comprehensive list of instruments that can be used, and a comparison of these different instruments with regard to their effectiveness and error tolerance, which is of particular importance given the substantial uncertainty about external costs implied by the vastly differing estimates collated in the CE study.

9. The CE study fails to identify the implications for the assessment of different policy options. There is also neither a detailed analysis of implementation costs and of the consequence on the economy of regulatory failure, nor any regulatory impact assessment. It correctly notes that marginal cost pricing methodologies might not be appropriate because they are only optimal under certain theoretical assumptions that are not satisfied in practice. However, the CE study does not sufficiently insist on the fact that interventions to

address externalities can and must take into account the restrictions characterising a second-best world.

Part II: Polluter pays vs. cheapest cost avoider principle

From an economics point of view, the polluter pays principle is an outdated and limited approach. It is widely agreed in economic circles that its shortcomings have been exposed and its suitability as sound basis for internalisation policies soundly superseded by the Cheapest Cost Avoider Approach developed by Ronald Coase. In the cheapest cost avoider framework, the "polluter pays" is one possible outcome of the analysis, but not a generally applicable principle.

To better understand the two economic approaches, it is necessary to first define them:

1. According to the polluter pays principle, the polluter should on all occasions pay the bill of the external costs he produces, usually via a tax. The elevation of the "polluter pays" idea to a principle is not justifiable.

2. The cheapest cost avoider principle requires that the party which can prevent (or abate) the damage at the lowest cost overall should take action. For example, (see case study 1 in Part III) it would be better for economic and social welfare overall to build a direct-route motorway rather than to make HGVs take a detour, simply because the extra economic and environmental costs of the detour are higher than the costs of building the motorway. Another simplified example would be noise emission: When a truck drives through the open fields the question of noise emissions plays only a minor role. However, when the truck uses a road nearby a house there is a conflict of interests. The house owner wants quiet and the truck needs to emit noise. The question that needs to be answered is: Is it more effective to build a sound barrier to solve the noise-problem or does it make more sense to just charge the truck user?
The cheapest cost avoider principle is applied in all areas of public decision-making under the heading of "regulatory impact assessment". It is not currently employed in the context of the regulation of transport related externalities.

Furthermore it is essential to understand the fundamentals of external costs:

3. The **fundamentals of external costs** are:

- External costs are always the **result of conflicting interests in the use of a scarce resource**: the environment. They arise only when there are competing uses of scarce resources.
- Without **rivalry** for the use of a scarce resource there are no external costs.
- Consequently, damage (external cost) is **jointly caused**.
- Damage is to be considered as a **loss of value to somebody** from a change in the quality (state) of the environment.
- If there is rivalry in the use of the scarce resource, policy makers are faced with a "tragic" **choice**: Furthering the interests of one group necessarily damages the interests of others. This is known as the **reciprocal nature of the problem**.

Bearing this in mind, the polluter pays principles has the following weaknesses:

4. The underlying economic logic of the polluter pays principle is fallacious.
The mere existence of externalities does not, of itself, provide any reason for governments to induce polluters to take action, because the polluters might well be the highest cost avoiders. Coming back to the example of truck noise, it might be that it is more costly for the overall economy to impose general charges on trucks than to build a sound barrier where there are conflicts of interests.

5. The polluter pays principle has the same flaws as the CE study which endorses it:

- it does not take into account the fact that externalities are **jointly caused** by all involved parties;
- it does not take into account that the externality problem is of a **reciprocal nature**. Improving the pollutee's situation necessarily inflicts harm on the polluter.

Most regulatory policy decisions now require some form of what has become known as regulatory impact analysis, regulatory impact appraisal or regulatory impact assessment. The polluter pays principle discounts such an assessment because the decision over who should pay has automatically been taken. By contrast the cheapest cost avoider principle is free from such preconceptions and includes a built-in cost-benefit analysis which can be applied fairly and efficiently to each situation requiring an internalisation of external costs. For the example of truck noise it could mean that a cost-benefit analysis establishes whether it is the truck user or the house owner or a third party like the state who can avoid the damage at the lowest cost for the overall economy by taking appropriate action.

6. The cheapest cost avoider principle presents a number of clear advantages over the polluter pays principle:

- It **guarantees efficiency**, i.e. no waste of resources, which is in turn fundamental in the pursuit of the European Community's **Lisbon goals of growth, jobs and competitiveness**.
- It is a better means to achieve **fairness** than the polluter pays principle.
- It studies a broader set of options. In contrast to the polluter pays principle, it can **lead to the choice of innovative projects**
- Its use of some form of cost-benefit analysis in a welfare economics framework makes it take a much **broader range of relevant variables** into account, such as administration costs or moral values.
- The logic of the cheapest cost avoider principle helps to **avoid regulatory failure** and contributes to the success of the Commission's Better Regulation Agenda at the heart of which is regulatory impact assessment.
- **The cheapest cost avoider analysis incorporates "polluter pays" as one possible outcome.** In contrast to the polluter pays approach, it does not make this outcome a principle.

7. The cheapest cost avoider principle is clearly more sophisticated than the polluter pays principle in terms of efficiency, and it is also superior when taking into account values such as corrective justice, distributive

justice, undamaged environment or the interests of future generations. Thus it meets challenges which the polluter pays principle cannot

Part III: Case studies

To show that the cheapest cost avoider principle can, is and must be applied, the study commissioned by the Centre for the Study of Law and Economics illustrates the methodology in **two exemplary cases**:

- The question of the construction of the missing part of the A 44 motorway near Kassel in Germany; and
- The problem of the sectoral ban on the use of the Inn valley motorway in Austria.

Both making the trucks drive a detour (A44 case) and banning certain trucks from using the Inn valley motorway can be interpreted as regulatory measures following the idea of the polluter pays principle.

1. In the first case study, called the A 44 case, there exist plans for the motorway A 44 connecting Dortmund to Kassel to be extended in the direction Erfurt, Chemnitz and Dresden, via Eisenach. However, except for a small section, these plans have never been realised. Instead, there is a direct road B7 between Kassel and Eisenach along the route of the planned motorway. This route is closed to trucks, who are obliged to drive a detour of 42 km via the motorways A7 and A4. This detour generates both extra private and external (pollution) costs, amounting to between 34 and 590 million € per year, depending on the source.

2. The A 44 case: The cheapest cost avoider analysis of this case involves the cost-benefit analysis of two scenarios:
Scenario 1 serves to address the question whether the interests of society are better served by building the A 44 compared to the detour. Scenario 2 serves to address the question whether the interests of society are better served by lifting the ban to use the B7 compared to the detour.

3. The second case study concerns the prohibition of a 46 km long section of the Inn valley motorway for trucks transporting goods belonging to a number of sectors. The sectoral ban on the Inn valley motorway is estimated to cost Germany approx. 250 million €, plus the losses of 2500 jobs due to insolvencies, increases in costs, and a reduction in demand.

4. Most cheapest cost avoider principle scenarios suggest that it would be beneficial to build the A44 and to lift the ban on the Inn Valley motorway.
The detour and the ban are inefficient.

5. Legal reasoning by the local court in the A 44 case and by the European Court of Justice implicitly applies the cheapest cost avoider approach.

Conclusions and recommendations

1. Due to its many shortcomings, the **CE Study cannot be regarded as a reliable or robust basis for policy prescriptions.**

2. It is not necessarily only the transport industry (i.e. the polluter) which should be made liable for externalities. Other actors, such as the state or the pollutee may well be in a better position to take measures to reduce externalities, and they should do so in the interest of economic efficiency and fairness.

3. Not all harm caused should automatically be internalised. Internalising too much of the damage would cost society more than it would benefit it. An efficient level of damage should be accepted.

4. When introducing new measures to reduce pollution by the transport industry, all existing levies should be taken into account in order to create optimal incentives. Among other taxes, the transport industry pays VAT and fuel tax. Environmental requirements for vehicles also present a cost to the transport industry.

5. The type of measure taken to make the transport industry pay must be based on a broad impact analysis. For example, a km toll would not sufficiently reduce the harm caused by pollution if trucks could employ an alternative route with the same (or even more) emissions. **Furthermore, making the polluter pay might not solve the pollution problem and it will give no incentive for the pollutee to solve the problem either.**

6. The polluter pays principle should not be used because its underlying economic logic is fallacious. It neglects the basic insight that external costs are caused jointly by all parties involved and that the externality is a problem of reciprocal nature.

7. The PPP can lead to the levying of unjustified, i.e. socially inefficient, taxes.

8. The cheapest cost avoider principle, for which Ronald Coase received the Nobel Prize for Economics, **should clearly be used**, also for transport, because it guarantees efficiency and fair competition. It is based on some form of cost benefit analysis (which is also part of regulatory impact assessment), it is a better means to achieve fairness, and it finally leads to better incentives for all parties involved. **Not using the cheapest cost avoider principle and thus not conducting some form of a cost-benefit analysis might lead to yet another regulatory failure.**

9. The cheapest cost avoider principle is partly already in use, especially in courts, which handle conflicts of interests. Regulation agencies active in road transport should follow the same maxims.

10. The cheapest cost avoider principle presents the fundamental tool in the pursuit of the European Union "Lisbon Goals" of growth, jobs and competitiveness.

Introduction

In March 2000, in what has come to be known as the "Lisbon strategy" European leaders committed the EU to become, by 2010, "the most dynamic and competitive knowledge-based economy in the world capable of sustainable economic growth with more and better jobs and greater social cohesion, and respect for the environment".

Achieving these aspirations requires a careful and comprehensive analysis and assessment of likely economic, social and environmental impacts, both direct and indirect, of all regulatory measures. Regulatory impact assessments, as part of the Better Regulation Agenda, are a key tool for the European institutions in designing better policies and laws. By providing a detailed study of the economic, social and environmental impacts, this procedure helps the European Union to reach the goals of Better Regulation, and its underlying Lisbon and Sustainable Development strategies.

As a follow-up to the Lisbon statement, the European Council concluded during its meeting in Göteborg (2001) that a sustainable transport policy should tackle rising volumes of traffic and levels of congestion, noise and pollution, and encourage the use of environment-friendly modes of transport as well as the internalisation of social and environmental costs. At the heart of this sustainable transport policy should be a fair, transparent and efficient system of charging for all modes of transport.[1]

According to Directive 2006/38/EC of the European Parliament and the Council amending Directive 1999/62/EC on the charging of heavy goods vehicles for the use of certain infrastructures, a "fair" system of charges is based on the "user pays" principle and the application of the "polluter pays" principle. The directive specifically mentions "the polluters [sic] pays principle for all modes of transport, by means of the internalisation of external costs".[2]

The 2006 directive requires the Commission to provide a model for the assessment of all external costs of road transport, which are to be internalised in the form of infrastructure charges. A study was commissioned from CE Delft and INFRAS (hereafter "CE study") to offer proposals on how external costs are to be defined and measured, and on why and how external costs should be internalised.[3]

The present report was commissioned by IRU from the Center for the Study of Law and Economics, University of the Saarland, in order to contribute to

[1] See European Commission 2001.

[2] See Directive 2006/38/EC, recital 18. See also recitals 18 and 19.

[3] See Directive 2006/38/EC, especially recital 19.

the ongoing debate about the appropriate treatment of external costs in the transport sector. More specifically, we have been asked to:

- offer a critical assessment of the above mentioned CE study;
- provide a theoretical examination of the relative strengths and weaknesses of the polluter pays principle and more modern concepts such as the cheapest cost avoider principle in relation to the problems arising in the road transport sector;
- demonstrate how the cheapest cost avoider principle can be *applied in practice* by providing examples of external costs in road transport; and to
- put forward *recommendations* for the appropriate principles that should guide the Commission's further activities in this area.

Our study offers a critical assessment of the CE study (part I), in which its narrow focus on the "polluter pays approach" is identified as a fundamental methodological flaw. Moreover, given that the need for a careful assessment of costs and benefits associated with any policy intervention is a well-established principle of community policy[4] one would expect such assessment being undertaken. But the CE study fails to do so.

Although crucial for any attempt to estimate external costs or assess the most appropriate way of internalisation, a proper definition of external costs and a consistent and coherent discussion of external effects are completely missing from the CE study. It is long on assertion and short on analysis. Subsequently the CE lacks a sound and reliable basis for the evaluation of potential internalisation scenarios and policy options.

The CE study does not take into account that both the measurement of external costs and the choice of an appropriate internalisation strategy crucially depend on the insight that externality problems are reciprocal, and that external costs are caused jointly by all parties involved.
The CE study endorses a naive view of external cost as being caused by one particular party – the polluter – and suppresses the more sophisticated understanding of external costs as costs arising from competing demands for scarce resources, which is now a generally accepted view in economics. It neglects the basic insight that external costs are caused jointly by all the parties involved and that the problem is of a reciprocal nature; avoiding pollution and thus improving the pollutee's situation would inflict harm on the polluter.
In the CE study, the question is thought of as one in which a polluter inflicts harm on a pollutee; what has to be decided is how to restrain the polluter. This is the logic underpinning the polluter pays approach. It obscures the nature of the choice that has to be made. The real question that has to be decided is: Should the polluter be allowed to harm the pollutee, or should the pollutee have the right to restrain the polluter? The aim should be to avoid the most serious harm; its solution is the cheapest cost avoider principle.[5]

---

[4] See European Commission 2001, 2005.

[5] See Schmidtchen 2003.

Efficiency can require charging a polluter, a pollutee or even that both parties pay. In brief, no single party can enjoy an automatic immunity from this joint responsibility as implied by the polluter pays principle.
The question how one should most appropriately deal with external effects is more complex than simply asking – as the CE study does – how much the party "causing" the external effect should be paying (and on what particular variables the payable amount should depend).

The CE study not only lacks a coherent framework for measuring external costs, where the appropriate measure of external cost is closely linked to the underlying objective of internalisation. The discussion of cost functions and external cost measures in the CE study is also conducted largely in terms of inappropriate variables. There are a number of inconsistencies that cast serious doubt on the scientific robustness of the analysis. It seems entirely inappropriate to rely on the cost estimates presented in the CE study as a basis for internalisation.

Additional flaws of the CE study are:

- The CE study does not put forward a consistent set of criteria against which one can judge the relative merits of different internalisation options;
- It is unclear about the effects of various instruments,;
- It fails to take into account the problems that arise from the fact that any practical internalisation policy will face substantial limitations; and
- It does not properly consider the interplay between internalisation policies and existing charging schemes.
- It can lead to the levying of unjustified, i.e. socially inefficient, taxes and charges.

In summary, the CE study does not provide a sound and reliable basis for the evaluation of potential internalisation scenarios and policy options.

Part II compares the polluter pays principle with the cheapest cost avoider principle. The underlying economic logic of the polluter pays principle turns out to be fallacious because the mere existence of externalities does not, of itself, provide any reason for governments to induce polluters to take action.[6] Indeed, the polluters might well be the highest cost avoiders. Thus the full internalisation of external costs is not always socially useful in that it does not necessarily maximise welfare. Moreover, the polluter pays principle does not take into account the fact that externalities are caused jointly, i.e. both the polluter and the pollutee cause of the damage. The fundamentals of external costs are these:[7]

[6] See Schmidtchen 2003.

[7] On external costs, see Schmidtchen 2003.

- External costs are always the result of conflicting interests in the use of a scarce resource: nature (environment). They arise only when there are competing uses of scarce resources.
- Without rivalry for the use of a scarce resource there are no external costs.
- Consequently, damage (external cost) is jointly caused.
- Damage is to be considered as a loss of value to somebody from a change in the quality (state) of the environment.
- If there is rivalry in the use of the scarce resource nature (environment), policy makers are faced with a tragic choice: Furthering the interests of one group we necessarily damage the interests of others. This is known as the "reciprocal nature of the problem".

The cheapest cost avoider principle requires policy makers to make a cost benefit analysis of alternative uses of the scarce resource and of corresponding abatement costs by all actors. Consequently, the party which can prevent (or abate) a damage at the lowest cost overall should take action.

The cheapest cost avoider principle, for which Ronald Coase received the Nobel Prize for Economics, presents a number of clear advantages over the polluter pays principle:

- It guarantees efficiency, i.e. no waste of resources, which is in turn fundamental in the pursuit of the Lisbon goals of growth, jobs and competitiveness of the European Community.
- It is a better means to achieve fairness than the polluter pays principle.
- It studies a broader set of options. In contrast to the polluter pays principle, it can lead to the choice of innovative projects.
- Its use of some form of cost-benefit analysis in a welfare economics framework makes it take a much broader range of relevant variables into account, such as administration costs or values.
- The logic of the cheapest cost avoider principle helps to avoid regulatory failure and contributes to the success of the Commission's Better Regulation Agenda at the heart of which is regulatory impact assessment.[8]
- The cheapest cost avoider analysis incorporates "polluter pays" as one possible outcome. In contrast to the polluter pays approach, it does not make this outcome a principle.

There is no criticism that can be made of the cheapest cost avoider principle that is not also valid for the polluter pays principle.

[8] See European Commission 2005.

To show that the cheapest cost avoider principle can, is and must be applied, part III illustrates the methodology in two exemplary cases:

- The question of the construction of the missing part of the A 44 motorway near Kassel in Germany; and
- The problem of the sectoral ban on the use of the Inn valley motorway in Austria.

Both making the trucks drive a detour (A44 case) and banning certain trucks from using the Inn valley motorway can be interpreted as regulatory measures following the idea of the polluter pays principle.

In the first case study, called the A 44 case, there exist plans for the motorway A 44 connecting Dortmund to Kassel to be extended in the direction Erfurt, Chemnitz and Dresden, via Eisenach. However, except for a small section, these plans have never been realised. Instead, there is a direct road B 7 between Kassel and Eisenach along the route of the planned motorway. This route is closed to trucks, who are obliged to drive a detour of 42 km via the motorways A 7 and A 4. This detour generates both extra private and external (pollution) costs, amounting to between 34 and 590 million € per year, depending on the source.

The A 44 case: The cheapest cost avoider analysis of this case involves the cost-benefit analysis of two scenarios:
Scenario 1 serves to address the question whether the interests of society are better served by building the A 44 compared to the detour. Scenario 2 serves to address the question whether the interests of society are better served by lifting the ban to use the B7 compared to the detour.

The second case study concerns the prohibition of a 46 km long section of the Inn valley motorway for trucks transporting goods belonging to a number of sectors. The sectoral ban on the Inn valley motorway is estimated to cost Germany approx. 250 million €, plus the losses of 2500 jobs due to insolvencies, increases in costs, and a reduction in demand.

Most cheapest cost avoider principle scenarios suggest that it would be beneficial to build the A 44 and to lift the ban on the Inn Valley motorway. The detour and the ban are inefficient.

Legal reasoning by the local court in the A 44 case and by the European Court of Justice implicitly applies the cheapest cost avoider approach.

Insights of the economic analysis of law, which is largely efficiency orientated, are increasingly used in policy making and legal reforms. In part inspired by the American model that requires a fully fledged cost-benefit analysis to be undertaken in preparing major regulatory proposals,[9]

[9] See Ogus 2006, p 282; European Commission 2001, 2005.

government departments in Europe, when making proposals, have to prepare what has become known as a "regulatory impact analysis", a "regulatory impact appraisal", or a "regulatory impact assessment", which includes some form of cost-benefit analysis.[10]

Cost benefit analysis in its strict form is a procedure for comparing the aggregate (social) gains to be attained by the application of a regulatory proposal against aggregate losses. Maximising the welfare of society, i.e. the sum of the welfare of the members of society, requires a comparison of the costs and benefits of different regulatory options, and the choice of the option which promises the highest net benefit to society. Such a cost-benefit analysis is concerned both with the determination of the optimal goal of a regulatory proposal and with the means of realising it at the lowest cost.

Some methods of regulatory impact assessment apply what we would like to call a "weaker" form of cost-benefit analysis which is known as cost-effectiveness analysis. Cost-effectiveness analysis has two principal functions: to determine how to maximise benefits for a given level of costs, or to determine what regulatory intervention will generate specified benefits at lowest cost. Compared to cost-benefit analysis in the strict sense, cost-effectiveness analysis is a less ambitious mode of economic appraisal, since it does not address the problem of specifying the optimal level of an economic policy goal, such as the optimal reduction of pollution. However, both the cost-benefit analysis in the strict sense and the cost-effectiveness analysis can be used to identify the cheapest cost avoider.

Note that at the heart of the cheapest cost avoider principle is the insight that the environmental performance of all modes of transport is crucial in order to encourage sustainable transport in the Community. Insofar, there is no difference to the polluter pays principle. However, the cheapest cost avoider principle shows much better than the polluter pays principle how to reconcile the EU goals of the Lisbon strategy, i.e. the respect for the environment with becoming the most dynamic competitive and knowledge-based economy in the world capable of sustainable economic growth with more and better jobs and greater social cohesion.

[10] See Ogus 2006, p 279-292.

1 A critical assessment of the CE study

1.1 Introduction and background

Directive 2006/38/EC of the European Parliament and of the Council of 17 May 2006, amending Directive 1999/62/EC, requires the Commission to present, "after examining all options including environment, noise, congestion and health related costs, a generally applicable, transparent and comprehensible model for the assessment of all external costs to serve as the basis for future calculations of infrastructure charges. This model shall be accompanied by an impact analysis of the internalisation of external costs for all modes of transport and a strategy for a stepwise implementation of the model for all modes of transport"[11] no later than 10 June 2008.

In order to assist the Commission's work in this area, CE Delft and INFRAS prepared a discussion paper ("the CE study") covering potential methodologies for the estimation of external costs and possible internalisation scenarios subject to further consideration. The paper was discussed in a workshop in Brussels on 15 March 2007, and the consultants are expected to provide a final report in November 2007. The findings of this final report are expected to feed into the expected Commission Communication on proposed policy options (and an accompanying impact assessment) which will be published in June 2008 and which will form the basis for a full stakeholder consultation.

The CE study addresses two distinct issues, namely:

- How external costs are to be defined and measured for various modes of transport and with regard to a number of potential sources of external costs (Chapter 2); and
- Why and how external costs should be internalised, what problems arise with such internalisation in practice, and what internalisation scenarios should be further considered (Chapter 3).

Tackling either of these issues on its own would be a formidable task, and dealing with both of them is extremely ambitious. Unfortunately, the CE study does not live up to this ambition, and falls short of what is required in order to provide a sound underpinning of important policy choices. It is long on assertion and short on analysis. It does not seem to provide a robust and coherent framework that allows one to assess the extent to which external effects in various modes of transport can lead to economic inefficiency, i.e. waste of resources, and how any such inefficiencies can best be avoided. It is confused and unclear about the objectives of internalisation, and seems to pay little attention to the potential risks of intervention and the associated welfare losses that could flow from ill designed or ill informed policies. It notes the significant amount of uncertainty that exists with regard to external cost estimates, but fails to highlight the implications of such

[11] Directive 2006/38/EC of 17 May 2006, Article 11.

uncertainty. In short, it should not be regarded as a reliable or robust basis for the design of policies.

The remainder of this section provides a brief overview of the approach taken in the CE study and its findings. This is followed by a discussion of the main methodological flaws of the CE study, namely:

- That it is based on a flawed notion of external costs; and
- That, even based on its inappropriately narrow notion of external costs, it does not provide a consistent and coherent framework for the measurement of such costs;
- That it lacks a consistent and coherent framework within which one could assess the different options for internalisation.

In addition, the CE study is strewn with inconsistencies, ambiguities and mistakes that cast serious doubts on its reliability and robustness. We will provide some examples of these in order to highlight concerns in this respect.

1.1.1 A brief summary of the CE study

The CE study is made up of two distinct parts.

- The first part (Chapter 2) is dedicated to a discussion of how to establish external costs estimates, and the collection of such estimates from third party studies.
- The second part (Chapter 3) presents a number of possible internalisation scenarios, following a discussion of potential objectives for internalisation, theoretical considerations relevant to internalisation, and practical issues that need to be addressed in order to internalise externals costs.

1.1.2 Estimation of external costs

The first part begins with a brief overview of the general approach, setting out the practical problems confronting any attempt to establish external costs associated with various modes of transport. The authors give their definition of external costs, discuss different methods for measuring such costs, and then review the best practice estimates for different cost categories.

According to the CE study, external costs are simply costs "not paid by the transport users" who are "thus faced with incorrect incentives for transport supply and demand, leading to welfare losses" (p 6).[12] However, one can distinguish between:

- The "scientific discussion, focussing on welfare optimisation and efficient pricing", in which "the term "external" is not of major importance", but where one started from the "marginal social cost of

[12] References are to the CE study unless otherwise specified.

transport" which formed the basis of optimal infrastructure pricing, and which implicitly defined external costs as the difference between these costs and marginal revenues (e.g. from infrastructure charges and fuel duties); and

- The "transport accounts discussion, where external costs are the difference between the total social costs of transport and the costs already paid by the user".

The CE study then considers different categories of external costs, defined with reference to the various effects that might arise from the decisions of transport users, namely congestion and scarcity, accidents, air pollution, noise, climate change, and other external costs.

With regard to ***scarcity and congestion*** costs, the CE study argues that it is important to distinguish between congestion (which "denotes the social loss due to the fact that users do not care for the additional costs and inconvenience they cause to others", p 10), delays (which are the effects of congestion), reliability (which is equally a consequence of congestion), and scarcity (which "denotes the economic costs to users and operators occurring when infrastructure can not [sic] be used at the desired time due to overcrowding", p 11).[13]

Best practice approaches for measuring the cost associated with scarcity and congestion are based on deriving time estimates from speed flow models and applying the opportunity cost of time associated with these (for road transport) and opportunity cost approaches for scarce tracks (for scheduled transport modes). This in turn suggests that it is appropriate to differentiate between various traffic networks (e.g. urban/inter urban) with their own specific speed-flow characteristics. The assumed value for the opportunity cost of time is a crucial determinant of the resultant estimates, and it is in this regard that considerable variations can be observed. While there are concerns about the quality of the underlying traffic data, methodological uncertainties are described as being comparably low.

Regarding ***accident costs***, the CE study states that external costs are "those costs which are not covered by risk-oriented insurance premiums" (p 12), so that the level of external costs depends on the insurance system. Discussing best practice approaches, the CE study finds that there are "two different approaches leading to rather different results", namely:

- A top-down approach which "estimates total and average accident costs considering national accident statistics and insurance systems", focussing on "material damages and administrative costs (usually covered in the insurance premiums), medical costs (including other

[13] These quotes are typical for the somewhat loose use of terminology throughout the CE study, which often makes it difficult to establish the precise meaning of a statement or to understand the logic underlying a particular argument. Though this is a pervasive problem of the CE study, we will not address it in the remainder of this section but rather attempt to reconstruct the intended meaning wherever possible.

insurance systems), production losses and societal valuation of risks (usually external)" (p 12); and

- A bottom-up approach which "aims at estimating marginal costs" which "depend on risk elasticity (correlation between traffic levels and accidents) and on the assumption of risk values" (p 13).

However, according to the CE study, there is no consensus on which of these two approaches should be considered to be the best practice approach, although the authors declare their preference for the top-down approach as it "is more transparent and considers the insurance system properly" (p 13). The CE study also notes concerns about under-reporting of fatalities and injuries in official statistics, and the substantial differences in the valuation of lives lost. Estimates for the Value of a Statistical Life (VoSL), which is used in the estimation of external costs, range from less than US$200,000 to US$30m.

In relation to ***air pollution costs***, the CE study refers to health costs, damage to buildings, crop losses and further costs of damage to the ecosystem arising from the emission of air pollutants. The most appropriate method for calculating external costs of air pollution is the so-called impact pathway approach, which is based on a detailed assessment of the link between activities and emissions, their transport and conversion, which result in specific concentrations and the deposition of pollutants. These cause a response from receptors and a physical impact, which in turn is valued in terms of associated changes in utility and welfare losses expressed in monetary terms. While this very detailed bottom-up approach is capable of considering detailed input variables and thus produces tailored estimates of air pollution costs for very specific traffic situations, deriving representative average figures for a whole country, say, is relatively costly. The CE study also notes a number of critical aspects and uncertainties, mainly related to the underlying data and the causal links between pollution and health, for example, as well as uncertainties with regard to VoSL and the appropriate social discount rates for costs imposed on future generations.

Noise costs are linked to the fact that noise is a nuisance which reduces the quality of life of those exposed to it, and can also have an impact on health. Noise costs can be estimated using a bottom-up impact pathway approach, tracing the effect of accommodating an additional vehicle on a particular route (for a given traffic volume, speed distribution, technology etc.), or a top-down approach based on national data on noise exposure and some measure of the willingness to pay for silence, derived for example from hedonic pricing studies providing the impact on property prices of an increase in noise levels. According to the CE study, both approaches are valid, but subject to similar uncertainties, as for example approaches for the estimation of air pollution costs (e.g. uncertainties about the appropriate figure for VoSL) as well as uncertainty with regard to the threshold value above which noise should be considered a nuisance.

Establishing the costs of ***climate change*** is complex owing to the long term effect of greenhouse gas emissions and the difficulty in anticipating risk patterns. According to the CE study, a "differentiated approach (looking both

at the damages and the avoidance strategy) is necessary" (p 19). In particular, when assessing the potential damage associated with climate change, "there is a general lack of knowledge about the physical impacts caused by global warming" with some "possible impacts, such as extended flooding or hurricanes with higher energy density" often not being taken into account because of "the lack of information on the relationship between global warming and these effects. Secondary impacts such as socially contingent damages (e.g. regional conflicts) are even more difficult to assess" (p 19). According to the CE study, an "alternative approach which avoids the uncertainties associated with assessing damage costs of climate control is to assess the costs of avoiding CO_2 emissions. (…) The method is based on a cost-effectiveness analysis that determines the least cost option to achieve a required level of greenhouse gas emissions reduction. (…) In practice the avoidance costs approach is more feasible, since the approach is more transparent and refers to climate change policy" (p 19 f). Obviously, the choice of target level, both in terms of the total value and the sectors to which it applies, and the estimation of the reduction potential of various technologies, play a critical role in this regard.

Other external costs, such as costs for nature and landscape, soil and water pollution, costs in urban areas or costs of up- and downstream processes are often considered, but not widely estimated owing to complex impact patterns.

The first part concludes with a presentation of figures for total external costs and various unit cost measures for the different cost categories across the various transport modes compiled from a number of third party studies (which produce a substantial range of values), and a brief discussion of the issues that need to be considered in order to transform these figures into values that can be used in internalisation scenarios. More specifically, the CE study proposes to use the unit values as "the basis for calculating the values for the various traffic situations, modes, types of vehicles and countries", noting that such "value transfer is useful and appropriate to save expenditures for detailed estimation of external costs in specific traffic situations, vehicle types, modes of transport and countries" (p 32). Such a "value transfer" would have to consider differences in dose-response functions, be based on local data or appropriate adjustments of unit values (such as VoSL, which would be adjusted on the basis of differences in GDP per capita). Further adjustments will have to be made for changes in the base year, and in order to obtain values for future costs. Last but not least, the CE study discusses how such figures might be used in internalisation scenarios, concluding that "the internalisation of congestion costs requests a road pricing scheme which differentiates at least between urban areas and interurban bottlenecks", while the "internalisation of environmental costs can be linked to km or fuel charging" (p 33).

1.1.3 Scenarios for internalisation

The second part starts from the observation that transport "gives rise to various types of external effects" for which there is no economic market so that "the market clearing process does not lead to the most optimal outcome, from a societal point of view" (p 35). The CE study distinguishes

between inter-sectoral externalities – external costs inflicted by transport users on others outside the transport sector (noise, pollution and climate change) - and intra-sectoral externalities where the external effects are experienced by transport users themselves (e.g. congestion and accidents). The solution to the problem of external costs, in the tradition of Pigou, is to "introduce a regulatory charge equal to the marginal external costs" so that transport users will "take account of the external effects on one another and others, and may or may not adapt their decision, depending on whether their marginal benefit is lower or higher than the marginal external costs to others. This is the basic idea behind internalisation of external costs" (p 35).

Having set out this "basic idea" – ensuring that decision makers face the cost that their actions may impose upon others – the CE study then progresses to discussing various potential objectives, or aims, of internalisation, namely

- Influencing behaviour (to reduce environmental impacts or allow a freer flow of traffic);
- Generating revenues (to cover infrastructure costs, fund new investments or contribute to the general budget); and
- Increasing fairness (in terms of making polluters pay; identical taxes and charges for everyone; changing, or prevent changes in income distribution; and levelling the playing field between modes).

According to the CE study, all of these objectives are to some degree present in Directive 2006/38/EC.

The discussion of potential motives for internalisation is followed by a review of the theoretical framework for internalisation, starting from the claim that the "internalisation of external costs can be done by a wide variety of methods and instruments". Even if one focuses on market based instruments "there are (…) many different ways of implementing pricing policies, for example with regard to price structures and price levels" (p 37). According to the CE study, the "optimal internalisation strategy depends on the underlying aims and motifs. If internalisation takes place out of equity considerations, intersectoral externalities are especially relevant, because these make up the "unpaid bill" that transport imposes upon society" (p 38). The CE study notes that marginal social cost pricing (which involves appropriate Pigovian taxes) would be "optimal for optimising economic efficiency" and "would, under some conditions, lead to allocative efficiency in a static perspective", but that in "the dynamic real world, deviations from marginal social cost pricing may be more appropriate or practical" (p 38) because:

- Marginal social cost pricing might not guarantee the full recovery of infrastructure costs;
- Pigovian taxes are only optimal under certain theoretical assumptions that may not be satisfied in practice (namely that all other prices in the economy correspond to marginal social costs, i.e. that there are no distortive taxes or economic profits, for example, that would lead to deviations from allocative efficiency); or

- The cost of implementing full marginal social cost pricing would be too high.

Following a brief discussion of the scope for differentiating existing charges in order to achieve internalisation, and the potential benefits from combining different policy instruments into acceptable packages, the CE study then considers how revenues raised from the imposition of Pigou taxes should be used. While efficiency considerations suggest that it would be inappropriate to earmark the revenues from charges aimed at internalisation for transport projects, various institutional conditions can offset the inefficiency that would otherwise be associated with earmarking. Equity considerations would not strongly support earmarking, but earmarking revenues may increase acceptability.

The CE study further notes that there are a number of legal constraints that may need to be taken into account, such as the exemption of jet fuel from minimum tax levels set out in Directive 2003/96/EC.

Before defining potential internalisation scenarios, the CE study then considers a number of "cross-cutting issues" (p 45 ff), namely:

- The most important cost categories for each mode of transport;
- The "incentive base" for charging (i.e. the dimensions along which charges should be varied);
- The question of how existing taxes and charges should be dealt with (in particular whether they can be considered to be aimed at internalising costs);
- How revenues raised from internalisation charges should be used;
- How to deal with different potential sources of external costs (air pollution, congestion, noise, accidents, and climate change); and
- The special issue of whether a toll system for passenger cars is needed.

The findings can briefly be summarised as follows:

- Generally, for most modes of transport, climate change and air pollution are the most important categories. For rail and aviation, scarcity and noise can be added to the most important cost categories. With regard to road transport, all cost categories with the exception of scarcity have to be considered as being most important.
- With regard to most cost categories, charges would have to be differentiated according to a combination of criteria. With the exception of climate change costs, where fuel use (and the CO_2 content of fuel) are closely correlated with the marginal (social) cost level, no single variable provides a good proxy for marginal cost drivers.
- Internalisation should in some form take account of existing charges in spite of some of the theoretical problems associated with relating those charges to marginal social costs. The CE study claims to take account of marginal taxes and charges that are not related to marginal infrastructure costs, the most prominent being fuel excise duty. The CE study assumes that fuel excise duties are used partly for the internalisation of climate change costs to the extent that revenues from

fuel excise duties for a particular transport mode exceed the variable infrastructure costs of this mode. Fixed charges are considered on a case-by-case basis.

- As a consequence of this approach, revenues from CO_2 taxes should be assumed to be used in order to "lower fuel excise duties to the extent that they exceed marginal infrastructure costs". Revenues from congestion charges are assumed to be used for investment within each mode, or for inter-modal funds, and taxes and charges for other external effects are assumed to be used for investments to reduce the external effects and rewarding the best in class (p 54).
- The CE study proposes further to tighten the Euro standards (or other standards) for road vehicles and locomotives, vessels and aircraft. In relation to congestion, tolls should be levied on congested links, with a recommendation to earmark revenues for road infrastructure investments. Noise can be best addressed through regulation and standards for road transport, and track prices and landing charges are linked to noise emission for rail and aviation. Regarding accident costs, the CE study proposes to charge insurance companies in line with estimated external costs, and rely on the insurance companies to pass on these charges in the most effective form to its customers. Climate change costs should be borne by each mode, and where legal constraints prevent passing through these costs in the form of fuel duties, it is proposed that the transport modes should be included in the existing emissions trading system.
- Given the previous considerations, there would seem to be little benefit from a sophisticated toll system based on per kilometre charging for passenger cars.

The CE study concludes with the definition of six different scenarios (from which four to six are to be picked for further detailed evaluation).

- The current situation defines Scenario 1, and serves as reference scenario.
- Scenario 2 comprises a differentiation of existing taxes and charges to bring them more in line with cost drivers, combined with regulatory intervention. Scenario 2 does not seek to introduce new charges, but to modify existing charges based on external cost estimates, with arbitrary limits on deviation from the average charge level. For example, CO_2 taxes or fuel charges would be set on the basis of external climate costs, while at the same time additional standards can be introduced to reduce emissions of pollutants. Existing road pricing or toll schemes would be modified to achieve a better spread of traffic between peak and off-peak, and so on.
- Scenario 3 aims to achieve full internalisation of external costs. This would require the introduction of new charges which should adhere as much as possible to marginal cost pricing, including a variable kilometre charging system for road traffic, and charging of insurance companies for accident costs. Two subscenarios are defined with regard to revenue use, where in one case (3A) revenues are used in different ways,

whereas in the other case (3B) all revenues are used in order to reduce charges that are not linked to marginal costs, and any remainder is used to reduce taxes on labour (which are assumed to have the largest distortive effect).

- Scenario 4 aims to achieve the full internalisation of inter-sectoral external costs only, i.e. it does not take into account costs the users of a particular transport mode impose on each other (namely congestion and accident costs). Revenues would be used in the same way as in scenario 3B.
- Scenario 5 consists of a mix of scenarios 2 and 3 in that only some new instruments are to be introduced. For example, unlike scenario 3, the scenario would not envisage a variable kilometre-based charging system.
- Scenario 6 assumes that certain optional measures laid out in Directive 2006/38 are made compulsory and thus applied to the maximum. For example, the scenario assumes that all Member States will introduce tolls for freight road infrastructure costs, differentiated by Euro standard and based on infrastructure costs as indicated in the Directive.

These scenarios were presented at the workshop, and a subset is expected to be assessed in more detail in the final report.

1.2 The CE study is based on a flawed notion of external costs

Although crucial for any attempt to establish external costs or assess the most appropriate way of internalisation, a proper definition of external cost and a consistent and coherent discussion of external effects are completely missing from the CE study. The CE study notes that "the production of unit values ready for internalisation needs a modelling approach with a clear definition of external costs" (p 5), but no such clear definition is forthcoming. In its place one finds the somewhat imprecise statement that "external costs are not paid by the transport users" (p 6), followed by a distinction between the "scientific discussion" and the "transport accounts discussion", which appear to differ with regard to whether they consider the gap between *marginal* social and private costs, and the gap between *total* social and private costs respectively. The section entitled "Definition of external cost and level of externality" then proceeds with a discussion of what costs transport users consider, a number of – unsubstantiated and unexplained - assertions with regard to the level of externalities for different cost components and transport modes, and some claims relating to the extent to which some of these costs may already be internalised, which – as the study claims – is "crucial for the definition of external costs".

Overall, this section does not provide any clear definition of external costs that could form the basis for measuring such costs (or establish the robustness of cost estimates from third party studies), and would be the precondition for assessing internalisation strategies.

1.2.1 The economic notion of external effects and external costs

In very general terms, externalities are effects of a consumption or production decision made by one agent on the consumption set, utility function or production function of other economic agents which do not work through the price system.[14] Externalities may be positive or negative, i.e. they may generate a benefit for, or impose a cost on other agents, which is by definition not taken into account by the decision maker. External costs can then be defined as the utility loss (expressed in monetary terms) or the increase in production costs suffered by those agents affected by the decision. By contrast, private costs are the costs faced by the decision maker. Social costs comprise both private and external costs.

A number of implications follow from this:

- External costs do not exist in a void, but are derived from utility losses or higher production costs of economic agents. For example, effluents pumped into a river are not an external cost; the increase in production costs of a downstream plant having to clean up the water before it can be used in their manufacturing process, by contrast, are external costs, as are the utility losses suffered by those who are (or would be) using the river banks for their recreation.
- External effects, and thus external costs, arise only when there are competing uses of scarce resources. Without a downstream factory using the water in its manufacturing process, or swimmers and sunbathers using the river as a leisure facility, there would not be an external effect, and there would not be any external costs.
- The fact that there are competing uses of scarce resources also means that reducing the harm to one party inevitably implies harming the other party. Reducing the cost faced by the downstream firm for cleaning river water by reducing the amount of wastewater pumped into the river upstream means imposing costs on the upstream firm, which will have to find other ways of disposing of wastewater, treating the water, or moving elsewhere. Externality problems are of a reciprocal nature: reducing the cost imposed on one party imposes costs on the other. Total costs depend on the actions of all of the parties involved – the decision-maker as well as those affected by the decision. For example, the extent to which using the river as a deposit of waste water causes external costs is likely to depend on the

[14] Strictly speaking, we are dealing here with so-called 'technological' externalities as opposed to 'pecuniary' externalities, which do work through the price system. For example, an increase in demand for a particular product by an economic agent may cause the market price for that product to increase, which obviously affects other buyers who now face higher prices. However, such pecuniary externalities are not Pareto-relevant, i.e. they do not affect the Pareto-optimality of competitive equilibria provided that all assumptions underpinning the first fundamental theorem of welfare economics hold, and thus prices only have the effect of equating supply and demand. For a definition and discussion of externalities see Laffont 1987.

production technology chosen by the downstream firm, which in turn affects how much cleaning is required. The utility loss suffered by swimmers and sunbathers might be avoided if they chose a spot upstream, although in this case they would experience higher travel costs.

The last two points are of particular importance, as they highlight that the question how one should most appropriately deal with external effects is more complex than simply asking how much the party "causing" the external effect should be paying (and on what particular variables the amount payable should depend). It may be efficient to charge the upstream "polluter" for pumping wastewater into the river, but not necessarily so. Welfare may be higher if the downstream firm were left to face the cost of cleaning the water that it requires for its production. Or efficiency may require that both parties pay - the upstream polluter some internalisation charge, and the downstream firm some cost of cleaning up the water. Which of these options should be chosen depends on the effectiveness with which the various parties can reduce the externality.

The CE study appears to acknowledge this fundamental insight arising from the work of Coase – albeit in a very cursory manner – by stating that "the quantification of environmental costs has to consider the interrelation between the cause and the effect of the externality. Noise is a good example: According to the Coase theorem, it must not only be the causer (transport) who has to pay for the externality" (p 7). However, the implication of this insight – namely that there is no justification for assuming that only "the causer" has to pay on the basis of the reciprocal nature of the externality problem – appears not to play any role in the remainder of the study, which proceeds without taking into consideration the fact that external costs can be avoided by multiple parties. Rather, with the exception of the brief reference to the Coasian perspective, the study is based on the presumption that internalisation is all about making the "causer" pay, based on the "cost" its actions are causing to other parties.[15] As the following example shows, this can lead to inefficiencies and welfare losses.

1.2.2 Externalities as a reciprocal problem and the cheapest cost avoider: a simple stylised example

A very simple stylised example may be helpful in illustrating the difference between the naive view of external cost as being caused by one particular party, and the more sophisticated understanding of external cost as cost arising from competing demands for scarce resources.

[15] One possible exception is the discussion of external costs of climate change, where the CE study acknowledges that the crucial issue is how to reduce greenhouse gas emissions with the least cost. However, as we discuss, this discussion is again superficial and incomplete, and only refers to the cheapest cost avoider principle without taking it seriously.

Consider the case of a driver using a particular stretch of road, and assume that the noise caused by driving along this road harms a local resident. Assume further that driving at greater speed benefit the driver (e.g. because he spends less time on the road), but also increases the harm for the resident. More specifically, consider that the value to the driver and the harm to the resident associated with various speeds – both expressed in monetary terms (i.e. as the amount the driver would be willing to pay for driving at the indicated speed[16], and as the amount the resident would be prepared to pay to avoid a car passing at the indicated speed) - are as set out in Table 1-1:[17]

Table 1-1 External cost from noise – a simple stylised example

Speed	Benefit to driver	Harm to resident	Total value (equals social welfare)
80 km/h	3	2	1
100 km/h	5	3	2
120 km/h	6	7	-1
140 km/h	5	11	-6

The harm to the resident of driving at higher speeds (or, indeed, of driving at all) is not taken into consideration by the motorist and can therefore be considered as the proper measure of "external costs" that require internalising. Maximising his own benefit, the motorist would choose to drive at 120 km/h, which is clearly inefficient, as the resident would gain more from a reduction in speed than the motorist would lose. Indeed, it would be socially optimal if the motorist chose to drive at 100 km/h, at which speed social welfare would be maximised. A lawmaker knowing the respective figures for benefit and harm with sufficient certainty could ensure a welfare maximising outcome by simply imposing a speed limit of 100 km/h (ignoring for the moment that it can of course be costly to enforce such a speed limit).

Alternatively, efficiency can be achieved by making sure that the motorist takes account of the harm he causes the resident, i.e. by transforming "external cost" into costs faced by the motorist when deciding how fast to drive. Assume that it is possible to charge the motorist for the use of that particular stretch of road, and to differentiate charges according to the

[16] This willingness to pay in turn captures the difference between the value of arriving faster and the direct cost borne by the motorist associated with travelling at higher speeds, e.g. in terms of higher fuel consumption or increased accident risk.

[17] From these values, it is straightforward to calculate incremental benefits and costs. For example, the incremental value to the driver of driving at 120 km/h rather than 100 km/h would be given by the difference in her valuation, i.e. 1. Similarly, the incremental harm caused to the resident by increasing the speed from 100 km/h to 120 km/h is 4.

speed chosen by the motorist (ignoring the cost of implementing such a charging scheme). Then, by setting charges equal to the harm suffered by the resident, the motorist will take account of the full cost of his decision, and choose to drive at 100km/h. This is the internalisation model that appears to underpin the CE study.

What this model ignores, however, is that the underlying problem is not caused by the motorist alone, but by the motorist and the resident jointly. It is only because the resident is there that the noise caused by the motorist causes harm; harm is not caused by the noise emission of the travelling car per se, but by the fact that this emission affects someone who values peace and quiet and suffers a loss in utility as a result of the noise. The fact that the harm is caused jointly by the resident and the motorist matters to the extent that the resident can take actions to affect the magnitude of such harm. For example, assume that the resident could reduce the amount of harm by installing sound insulation in his property at a cost of 2, and that the reduced harm resulting from this is as shown in Table 1-2.

Table 1-2 External cost from noise with mitigation efforts

Speed	Benefit to driver	Harm to resident if sound insulation installed*	Total value, including cost of installing sound insulation (equals social welfare)
80 km/h	3	0 (2)	1
100 km/h	5	1 (3)	2
120 km/h	6	1 (3)	3
140 km/h	5	5 (7)	-2

* Figures in parentheses show the sum of harm and installation costs

In this case, the welfare maximising outcome would be for the motorist to drive at 120 km/h. Moreover, total welfare in this case is higher than the maximum welfare that can be achieved by naively focusing on the motorist as the party responsible for causing harm, and pursuing the corresponding internalisation strategy (as described above). This is because, by installing sound insulation, the resident can reduce the harm he suffers by more than the motorist ever could as a result of driving more slowly, and at any given speed up to 140 km/h, can do so more cheaply.

Perhaps even more importantly, attempting to internalise external costs on the basis of the naive view of cost causation outlined above will in all likelihood lead to welfare losses relative to what could be achieved if one took into account the fact that some harm is unavoidable (i.e. that any attempt to reduce the amount of harm due to noise suffered by the resident will go hand in hand with other costs that have to be incurred, or with lost

benefits to the motorist from having to drive more slowly), and that in order to minimise the magnitude of the (inevitable) welfare loss one has to ensure that the party with the lowest cost of reducing the amount of harm suffered has an incentive to do so. In order to see this, we need to focus on the incentives of the resident to install sound insulation on his property.

Without charging the motorist for any "external cost", or imposing a speed limit, the motorist would choose to drive at 120 km/h, and the resident would obtain a net benefit of 4 from installing sound insulation.[18] By contrast, if the motorist drove at 100 km/h because of a speed limit or because of being faced with external costs determined on the basis of the naive view of who is responsible for causing harm, the resident would not have any incentive to install insulation.[19] By contrast, doing nothing to confront the motorist with the harm he causes would lead to the right outcome.

This last conclusion rests on the assumption that the resident is unambiguously the cheapest cost avoider[20], which may of course not hold in practice. For example, assume that the benefits to the motorist from driving at 140 km/h were 7, say. In this case, absent any charging mechanism or restriction, the motorist would choose to drive at this speed, even though as a result total welfare would decrease. It may therefore be appropriate to confront the motorist with the external cost he causes – but these costs need to be measured appropriately and taking into account the reciprocal nature of the problem and the fact that the harm is caused jointly by both parties rather than one party alone. More specifically, in the above example the appropriate external cost charge (or Pigou tax) would need to be set at the level of harm suffered by the resident under the condition that the resident has taken appropriate measures to reduce the quantum of harm.[21] The following table presents the difference between the external cost charge that results from the naive view of cost causation, and the understanding that externality problems are reciprocal in nature.[22]

[18] The resident would reduce noise costs from 7 to 1, i.e. gain benefits of 6. Given that the installation of sound insulation costs 2, this leaves a net benefit of 4.

[19] The harm suffered at this speed is 3, the cost of installing insulation is 2, and the harm suffered at this speed with insulation is 1, as shown in the above Tables.

[20] Also called "cheapest cost avoider" in the literature.

[21] See Rose-Ackerman 1989.

[22] See Baumol 1972.

Table 1-3 Pigovian tax based on correct notion of external cost

Speed	Pigou tax based on reciprocal nature of externality*	Naive Pigou tax
80 km/h	2	2
100 km/h	3	3
120 km/h	3	7
140 km/h	7	11

*Calculated as the minimum of the harm suffered without insulation, and the sum of insulation cost and the harm suffered with insulation
Note that the figures in the table give the total amount of tax payable rather than a tax schedule giving the tax associated with changes in activity levels. However, converting the total tax amounts into marginal tax rates for driving at 80 km/h rather than not driving at all, driving at 100 km/h rather than at 80 km/h and so forth is straightforward.

This stylised example is of course very simplistic, and a more detailed discussion of the underlying principles will follow in part II. It demonstrates, however, that both the measurement of external cost and the choice of an appropriate internalisation strategy crucially depend on the insight that externality problems are reciprocal, and that external costs are caused jointly by all the parties involved. This insight is not limited to externalities from noise, but it is applicable to externalities in general.

1.3 The CE study lacks a coherent framework assessing external costs

In addition to ignoring the reciprocal nature of externalities (and the implications this has on the appropriate measure of external costs) the CE study does not set out a coherent "methodology" for the measurement of external costs.

In the first place, the CE study appears not to contain a methodological framework for the measurement of external costs, but rather presents a collection of assertions and examples of cost measurements that have been conducted by third parties. Even though one of the declared objectives of the CE study is the collection of such third-party cost estimates, it is unclear how the various estimates can be compared and used in an internalisation framework without an explicit discussion of an appropriate methodology for the measurement of external costs (and a subsequent assessment of any differences that might exist between the methodology used in producing the various estimates and the extent to which these differences could potentially explain the variation in the estimated values).

The following list of problems (without any claim to being complete) provides some indication of the underlying deficiencies of the CE study.

- The CE study does not appear to consider that the appropriate measure of external costs is closely linked to the underlying objective of internalisation. While it may be interesting to establish "total" external costs with reference to the difference between the overall cost caused

by a particular mode of transport, and the cost incurred by the users of this transport mode, such a measure is entirely inappropriate and irrelevant if one wants to correct incentives so that the level of an externality is reduced to its efficient level. A measure of total external costs may be relevant if one were to intervene purely under the heading of corrective justice (one of the many concepts of "fairness" mentioned as a potential policy objective)

- For a similar reason, it is entirely meaningless to calculate, for example, external costs per vehicle kilometre (which appears to be the prevailing method of calculation throughout most of chapter 2 of the CE study – see, for example, Tables 7 to 10) if vehicle kilometres are not the relevant driver of external costs (as they appear not to be, according to Table 13, which suggests that the level of external costs depends on a mix of factors in a potentially highly complex way). Making the polluter face some arbitrary average cost measure that is unrelated to the impact her behaviour has on the level of external costs not only fails to achieve internalisation in the sense of improving efficiency, but may even be counterproductive.
- Even if unit costs were expressed in terms of the correct driver of external costs, the relevant concept for internalisation is marginal rather than average cost (or unit cost).[23] Except in the case of constant marginal costs and no fixed costs, these two cost measures will differ. Of course, as the CE study points out, there are a number of reasons why the implementation of internalisation schemes based on marginal external costs is impractical or impossible, but this does not suggest that one can afford simply to ignore the impact that the recourse to some other, more easily quantifiable cost measure will have on the effectiveness and the welfare implications of internalisation. At the very least, in order to examine to what extent unit costs (or average costs) may be used as a proxy for marginal costs, one has to take a view on the shape of the external cost function. A detailed discussion of these aspects is missing. Although a comparison of average and marginal cost per vehicle kilometre is undertaken in Table 3 of the CE study, it is unclear to what extent these findings are driven by modelling assumptions which may or may not be realistic. In any case, the comparison is somewhat meaningless given that vehicle kilometres seem to be a poor proxy for the underlying drivers of external costs.

[23] The CE study indeed notes that "[a]ccording to theory and ongoing research, there is no doubt that marginal social cost figures are needed for optimal internalisation scenarios", but then continues by stating that "[I]n practice and related to the new Directive for HGV charging 2006/38, a separation between infrastructure costs (and recovery) and external costs is however decisive" (p 9). It is entirely unclear why (and if so, how) marginal external cost estimates would be incompatible with a separation of infrastructure costs, nor how such a separation relates to average costs.

- The CE study briefly mentions that it is "crucial for the definition of external costs" to understand which costs are already taken into account by the decision maker (p 7), but does not provide any systematic discussion of this issue. Any attempt to do so would presumably have demonstrated that, in order to examine this question, one not only has to consider the incentive properties of existing charging systems, but also a wider set of legal and institutional factors. Accident costs are a good example for this.

The CE study identifies as the external costs of accidents those that are "not considered in own risk anticipation and not covered by insurance" (p 7). Although this is later refined by stating that external accident costs are those that are "not covered by risk oriented insurance premiums" (p 12), the preferred approach for the measurement of external accident costs according to the CE study is a top-down approach, which can, but does not necessarily have to restrict its attention to insurance premiums that are based on risk (as set out on page 12).[24] The CE study completely ignores that (a) those "causing" accidents[25] are often liable for damages that are not covered by their insurance (under general tort law), so that a greater proportion of costs would be internalised than is suggested by looking at insurance premiums and (b) that insurance systems in which the premium paid is unrelated to the insured's risk characteristics and behaviour can be the cause of externalities rather than contribute to their internalisation. Put differently, if liability for the damage caused by one's behaviour were not covered by insurance, then the incentives to take account of the likelihood of an accident might well be taken into account to a greater extent than in the case where liability is covered by insurance and the insurance premium is unaffected by one's behaviour. Insurance premiums are of course not generally set in close relation to the risk characteristics of individual insurance customers owing to the information asymmetry between insured and insurer, limiting the insurer to designing a set of policies and relying on customer self-selection; monitoring costs limit the ability of the insurer to make insurance premiums contingent on the insured's behaviour. Sometimes, insurers also face restrictions on their ability to differentiate their insurance policies or their premiums that have been imposed in pursuit of some other policy objective (such as concerns about fairness or affordability). Such constraints can contribute to externalities. Similarly, the fact that under general medical insurance systems some health costs are shared by the community rather than being borne by those involved in an accident can give rise to externalities – but these are the consequence of a political choice to have in place a general system of health insurance rather than being caused by transport.

---

[24] Consequently, top-down approaches may lead to very different cost figures, as noted in the CE study.

[25] Please note again that accidents are generally 'caused' by two parties, and that either party may be able to undertake efforts to avoid the accident, or limit the level of expected harm that results from the accident.

The general lesson of this discussion is that one needs to consider a wide set of legal and institutional factors in order to assess what proportion of external costs is already internalised.

The complete lack of a systematic and robust methodology for the assessment of external costs is reflected in the fact that the discussion of cost functions and external cost measures in the CE study is conducted largely in terms of inappropriate variables, and that there are in addition a number of inconsistencies that cast serious doubt on the scientific robustness of any analysis that underpins the assertions made in the CE study. Table 1-4 brings together the entries in Table 3, Table 5 and Table 13 of the CE study in relation to external cost functions, showing these problems.

Table 1-4 The notion of external costs in the CE study

Cost component	Table 3: Difference between marginal and average costs (costs per veh. km)	Table 5: Cost function (costs per veh. km)	Table 13: Correlation between vehicle kilometre and marginal cost level	Comment
Costs of scarce infrastructure/ Congestions costs	Marginal costs above average costs	Increasing marginal costs	Low	Consistent – increasing marginal costs imply marginal costs above average costs
Accident costs	Marginal costs for infrastructure use unclear. Average costs as a proxy possible	Weak dependency; difference between marginal and average costs not proven	Low	Marginal cost meaningless if vehicle kilometres are not a good cost driver; average costs cannot be a meaningful proxy in this case
Air pollution costs	Linear doses response function: marginal costs similar to average costs	Complex: Increasing marginal cost curve	Low	Inconsistent – increasing marginal costs imply marginal costs above average costs and suggest that linear dose response functions are inappropriate
Noise costs	Decreasing dose response function; marginal costs below average costs	Decreasing marginal cost curve	Low	Broadly consistent, but CE study fails to draw out the implications for internalisation[26]
Climate change	Marginal damage costs similar to average costs (if no major risks included). For avoidance costs, marginal costs are higher than average costs	Proportional (marginal cost close to average costs)	Medium	Inconsistent, given that preferred approach is avoidance costs. Unclear why damage costs excluding major risks are of any relevance; if major risks included, highly unlikely that marginal and average costs correspond given that relationship between cause and effect is complex and potentially highly non linear
Nature and landscape	Marginal costs are significantly lower than average costs	Very low marginal costs, but high fixed costs	Low	Broadly consistent, but unclear why high fixed costs (potentially explained by other than choice of wrong cost driver)

[26] As Varian 1992, p 434, notes, with a convex external cost function, optimal behaviour can be achieved be imposing a tax that is equal to the marginal external cost at the optimal activity level. By contrast, with a non-convex external cost function, a non-linear tax schedule that sets the tax payable for each level of activity equal to the external cost caused by that activity will be required.

The lack of a coherent and consistent framework is particularly regrettable given that the cost figures estimated by various third parties and quoted in the CE study differ widely. This may simply indicate genuine uncertainty about the magnitude of external costs, but it may also reflect differences in the methodology used in these studies. In the latter case, it would be important to establish whether particular cost estimates should be given more or less weight (depending on whether the methodology used for estimating the value of costs is well or badly aligned with the underlying internalisation objectives), or whether some cost estimates should even be disregarded because they include or exclude specific effects.[27] It might even have been possible to narrow down the range of estimates by examining whether the different methodologies used in these studies were likely to result in over- or underestimations relative to an appropriate reference methodology.

However, no attempt is made in the CE study to explore the potential reasons for the differences in cost estimates. Without such an analysis, and in particular taking into account the enormous variation in the numbers presented, it seems entirely inappropriate, and potentially dangerous, to rely on the cost estimates presented in the CE study as a basis for internalisation.

A few examples suffice to show the extreme variations and inconsistencies in the cost figures presented:

- On page 27, the CE study refers to two studies looking at total external cost estimates, namely the UIC study, and the UNITE project. The former obtains a figure for total external costs of transport (excluding congestion costs and with a not further specified "high" scenario for climate change costs) as € 650 billion for 2000. 30% of total costs are accounted for by climate change costs, 27% by air pollution, 24% by accidents, 7% for noise and up- and downstream process respectively, and 5% by costs for nature and landscape and additional urban effects respectively. The latter estimates total external accident and environmental costs of € 120 billion. The corresponding figure from the UIC study covering the costs of climate change, air pollution and accidents amount to € 526.5 billion – more than four times the UNITE estimate. According to the CE study, the difference is due to "different methodologies for accident and air pollution costs and cautions valuation of external costs" (p 27). Assuming that the reference to cautiously valued external costs actually means the costs of climate change, and excluding these from the UIC figure, one still ends up with a figure of € 331.5 billion – more than two and a half times the UNITE estimate, with differences presumably being entirely explained by "different methodologies" for the estimation of accident and air

[27] For example, the different (and inconsistent) estimates in the third party studies quoted could be the result of differences in underlying assumptions, or they could reflect differences in the definition of the cost measures that are being established.

pollution costs. Relying in any way on such different estimates without fully understanding what drives the differences would seem to be wholly unacceptable.

- Matters become worse when one considers the various unit cost estimates presented in the CE study for various transport modes and cost categories. Given that road transport allegedly accounts for the large majority of external costs (83.7% of those external costs measured in the UIC study, and 94% of the external costs measured in the UNITE study, according to the CE study, p 27), we focus on the unit cost estimates for the external costs of road transport (Table 7 in the CE study). Table 1-5 reproduces Table 7 of the CE study, but gives in addition the difference between the minimum and maximum values as a proportion of the lower value.[28] This shows the considerable variation – with the exception of the accident costs associated with passenger cars in urban traffic, the maximum unit cost figure is at least twice the minimum figure; in most cases, it is four or five times the minimum figure, and it can be up to 24 times the minimum figure. Given these significant variations, there seems to be little point in even considering adjustments in order to arrive at applicable figures for particular countries, or in order to rebase cost estimates to a particular year (e.g. adjustments that would be made to cost estimates based on historic data in order to obtain cost figures in current terms). In the light of the massive uncertainty regarding the underlying cost estimates, any such adjustment would do no more than result in spurious precision, and does not add any more plausibility or reliability to the estimates.

[28] Note that Table 7 of the CE study distinguishes between petrol and diesel engines in relation to air pollution costs, but gives figures for the HGV segment in relation to petrol engines. Given that HGVs (heavy goods vehicles) tend to be powered by diesel engines, this is presumably a mistake and the figure given refers to diesel engines.

Table 1-5 : Unit costs (€ct/vehicle km) of road transport and difference between minimum and maximum value estimates, CE data

		Passenger car			HGV		
		Min	Max	Proportion of lower value	Min	Max	Proportion of lower value
Noise	Urban	0.8	3.4	425%	7	31	443%
	Interurban	0	0	N/A	0.1	0.2	200%
Congestion	Urban	2	28	1400%	6	84	1400%
	Interurban	0	15	N/A	0	7	N/A
Accidents	Urban	4.2	4.8	114%	3.2	11	344%
	Interurban	0.3	7.2	2400%	0.3	2.8	933%
Air pollution	Urban Petrol	0.1	0.3	300%	4.7	18	383%
	Urban Diesel	0.3	1.5	500%			
	Interurban Petrol	0.1	0.4	400%	2.1	7.5	357%
	Interurban Diesel	0.3	0.6	200%			
Climate change	Urban	0.6	2.3	383%	2	7	350%
	Interurban	0.3	1	333%	1.2	4.3	358%
Nature & landscape	Urban	0	0	N/A	0	0	N/A
	Interurban	0	0.4	N/A	0	1.2	N/A
Total	Peak, urban	7.7	39	506%	23	150	652%
	Peak, interurban	3.1	38	1226%	10	103	1030%
	Off-Peak, urban	5.7	26	456%	17	73	429%
	Off-Peak, interurban	1.1	25	2273%	4.5	26	578%

As an aside, it is difficult to see how, based on this Table, the CE study can claim that each and every of these cost categories is a "most important" cost category (which in itself is nonsensical) with regard to road transport (see Table 12 on page 46). The above table would suggest that the most important cost categories are accidents and congestion.

In any case, the numbers presented appear to be inconsistent with the total cost figures presented. Based on the unit cost figures, it is possible to calculate the proportion of external costs accounted for by different categories such as climate change, accidents etc., and to compare these proportions with those given for the total cost estimates. Given that road transport accounts for the large majority of external costs, it should be possible to find a similar distribution of cost across the different categories. However, this is not the case, as Table 1-6 shows. It gives the total unit cost for the two vehicle types and the two transport settings, using minimum and maximum cost figures from Table 7 of the CE study, excluding congestion costs, and the corresponding shares of these totals accounted for by the different cost categories. There does not seem to be any traffic pattern (split

between passenger cars and HGVs and urban and interurban traffic) capable of producing the distribution of costs in the UIC study, which, excluding costs for up- and downstream processes, are as follows:

- climate change: 32%
- air pollution: 29%
- accidents: 26%
- noise: 8%
- nature and landscape: 5%

Table 1-6 Proportion of external costs in different categories as implied by the unit cost figures for road transport

	Passenger car				HGV			
	Urban		Interurban		Urban		Interurban	
	Min	Max	Min	Max	Min	Max	Min	Max
Total costs, excluding congestion	5.8	11.4	0.8	9.1	16.9	67	2.65	12.25
Share of these accounted for by:								
Climate change	10%	20%	38%	11%	12%	10%	45%	35%
Air pollution	3%	8%	25%	5%	28%	27%	79%	61%
Accidents	72%	42%	38%	79%	19%	16%	11%	23%
Noise	14%	30%	0%	0%	41%	46%	4%	2%
Nature and landscape	0%	0%	0%	4%	0%	0%	0%	10%

Notes: calculations based on the numbers given in Table 7 of the CE study; air pollution costs for urban/interurban traffic settings for passenger cars are assumed to be the unweighted average of the corresponding cost figures for petrol and diesel cars

Taken together, these findings suggest that little reliance can be placed on the cost figures presented in the CE study as a basis for internalisation. No attempt is made to investigate potential reasons for the substantial differences and to potentially narrow down the range by examining differences in the underlying methodologies - and, as we discuss below, no implications are drawn from the fact that the large variations might indicate substantial uncertainty about the absolute and relative magnitude of external costs.

1.4 The CE study lacks a coherent framework for internalisation

Even if one were to ignore the concerns about the notion of external costs on which the CE study is based, and to accept the cost estimates presented in the study at face value, the CE study would fail to provide any reliable guidance with regard to the choice of potential internalisation scenarios, let alone help in relation to the assessment of their relative costs and benefits (which we understand to be a crucial part of the final report). This is because a proper analysis of potential policies has to be based on an understanding of what ought to be achieved, what tools are available to

achieve the objective, how these tools can be expected to perform, and how they interact with other policies. However, the CE study does not put forward a consistent set of criteria against which one could judge the relative merits of different internalisation options. It seems to be unclear about the effects of various instruments, does not seem to take into account the problems that arise from the fact that any practical internalisation policy will face substantial limitations, and – as briefly discussed above – does not properly consider the interplay between internalisation policies and existing charging schemes.

Before proceeding to set out the critique with regard to these aspects in more detail, it is however helpful to provide a brief definition of what we mean by "internalisation", not least because no attempt to provide such a definition is made in the CE study.

1.4.1 What is internalisation?

In very general terms, internalisation can be regarded as an intervention that leads to the decision maker facing the full social costs of his actions. This means that costs that would otherwise be "external" are now taken into account by the decision maker and affect his behaviour. In an ideal setting, internalisation restores the Pareto optimality of competitive equilibria. Pareto optimality requires that no actor can be made better off without making someone else worse off. This implies that there is no waste of resources.

There are many different ways in which internalisation can be achieved, but normally internalisation refers to interventions that rely on the price mechanism. Pigovian taxes are one particular form of internalisation, and the following simple diagram is helpful in terms of illustrating a number of important points.

Figure 1-1 Internalisation through a tax - a simple example

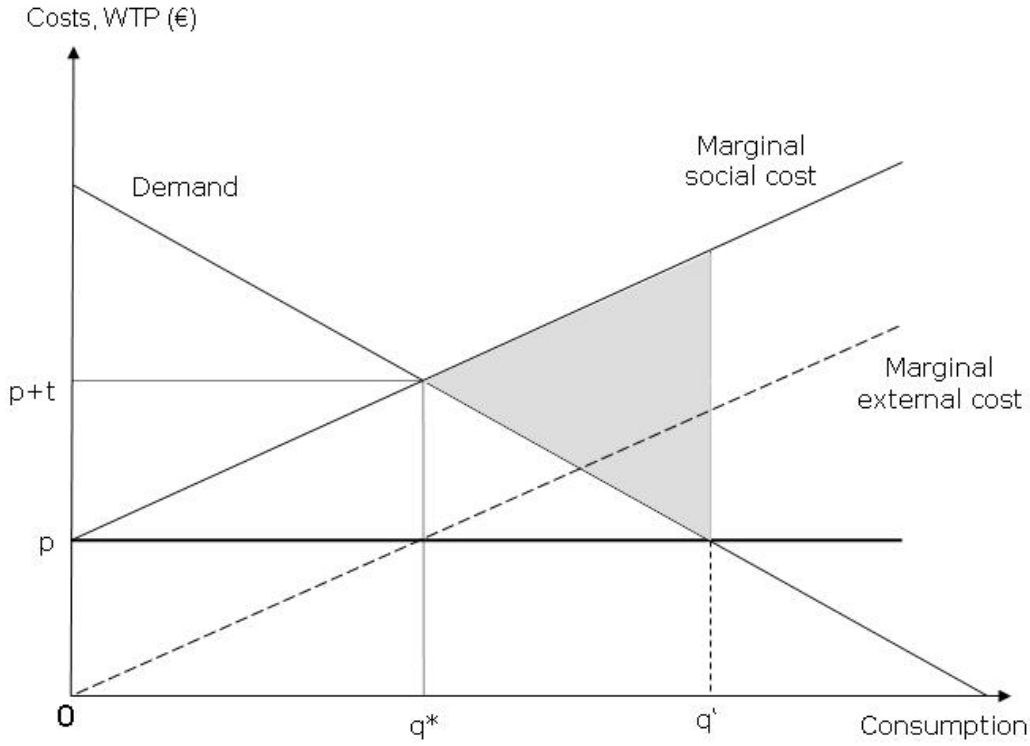


Figure 1-1 depicts a situation in which the consumption of a particular good has an impact on the utility or production options of other economic agents. The downward sloping line depicts demand for the good, and the horizontal line at p reflects the (perfectly elastic) supply, i.e. the price at which any particular amount of that good can be made available. As consumers have to pay this price for any unit they consume, p also reflects the marginal private cost of increasing consumption. The dashed upward sloping line shows the marginal external cost associated with particular levels of consumption, and the continued upward sloping line (which is obtained by adding up the external cost and the price) reflects marginal social cost. Ignoring the external costs associated with their decisions, consumers would demand an amount q', where willingness to pay equals price. By contrast, the socially optimal consumption level would be given by q*, where marginal social cost equals willingness to pay. This is because for any increase above q*, the additional social cost, given by the sum of marginal private and marginal external costs, exceeds the marginal social benefit, given by the willingness to pay for the additional unit as represented by the demand curve. The fact that consumption causes an external effect results in over-consumption by an amount equal to q'-q*. The welfare loss associated with this over-consumption is represented by the shaded triangle, indicating the difference between marginal social cost and marginal social value for all units between q* and q'. By levying a tax of t – which corresponds to the level of marginal external costs *at the optimum* q* - consumers can be made to choose the socially optimal consumption level q*.

This simple example allows us to highlight a number of insights that are relevant for the assessment of internalisation scenarios, but that are not at all discussed in the CE study:

- First, optimality does not imply complete avoidance of external costs, but is defined with reference to the marginal social cost being equal to marginal valuation. The tax on consumption at the optimum consumption level q* is equal to marginal external costs at that level of consumption, and drives a wedge between the price paid by the consumer (p + t) and the price received by the supplier (p).
- Second, the welfare loss associated with over-consumption is smaller than the external costs resulting from this over-consumption, which in the above example are represented by the area between the solid line and the price from q* to q'.
- Third, marginal external costs are equal to marginal avoidance costs only at the optimum, but not elsewhere. In the above example, avoidance costs are reflected by the loss in consumer surplus that would be associated with a reduction in the activity causing the external effect, i.e. the level of consumption. This is measured by the difference between the demand curve and price at any given consumption level (and thus increases as consumption is reduced further and further). In any case, it is inappropriate to use avoidance costs as a proxy for the external cost (as the CE study does for the cost of climate change[29]) without considering whether it is likely to be optimal to incur the cost of avoiding emissions, how much avoidance should be undertaken in the optimum, and by whom.
- Fourth, optimality requires that consumers face the marginal external costs of their behaviour at the margin. Any non decreasing tax schedule that yields a tax rate of t at q* will induce optimal behaviour. This also implies that optimality does not necessarily require that those "causing" the externality exactly bear the total social costs associated with their behaviour. In the above example, the total tax burden at the optimum (which is given by t · q*) exceeds the total external costs associated

[29] See pages 8, 9 (table 2) and 19 of the CE study. It is worth noting that the CE study further appears to confuse avoidance costs with shadow prices, stating that "[a]n alternative approach which avoids the uncertainties associated with assessing damage costs of climate control is to assess the costs of avoiding CO2 emissions. These are often referred to as avoidance costs or mitigation costs, and are expressed as so-called shadow values" (p 19). However, shadow prices generally refer to the amount by which an objective function (e.g. social welfare) changes if a constraint (e.g. the amount of CO_2 that firms are allowed to emit) is changed. Thus, the shadow price of CO_2 emissions would correctly be measured as the welfare impact of allowing an increase in emission volumes from a particular level. Requiring the shadow price of CO_2 emissions to equal the cost of reducing emissions is a condition that needs to hold at the optimal emission level rather than in general, and thus considering avoidance costs and shadow prices to be the same would seem to be a clear methodological flaw.

with this level of consumption (which is given by the area between the thick line and the price, from 0 to q*). Indeed, optimality does not require that those who cause the externality make any payment at all. For example, rather than taxing consumption, internalisation might be achieved by subsidising non consumption, i.e. by paying consumers for reducing their consumption levels.

- Fifth, although one might refer to any intervention that is aimed at achieving a consumption level of q* as internalisation, it is worthwhile to distinguish the effects of price and non price mechanisms. Consider, for example, an intervention that limits the amount each individual consumer is allowed to consume so that the aggregate level is q*. This is not guaranteed to lead to q*, however, because some consumers may not wish to exhaust their allowance at the prevailing price. Even if all consumers wanted to consume up to their limit, the overall outcome may not be welfare maximising, because there is no guarantee that total consumption is shared out amongst those that value it most. Put differently, consumers may differ with regard to their marginal valuation at their individual limits, which would be reflected in gains from trade. Unless consumers are allowed to trade allowances, however, these gains from trade would remain unexploited.[30]

1.4.2 The objectives of internalisation in the CE study are unclear

As the CE study notes, in order to discuss the relative benefits of different policy proposals for internalisation of external effects, it is necessary to be clear about the underlying policy objectives. However, clarity about policy objectives cannot be obtained by listing the different motives for internalisation, and then – as the CE study does – suggesting that in practice the objective underlying internalisation may be a bit of everything, and that in practice "implementation of pricing policies will generally contribute to more than one potential aim"(p 36). This is because the potential objectives may not only be poorly aligned, but may actually be conflicting.

This is obvious when one considers the first and second of the three potential motives for internalisation put forward in the CE study, namely the aim to influence behaviour in order to reduce environmental impacts and to allow a freer flow of traffic, and the aim of generating revenues. Obviously, the more successful an internalisation scenario is in terms of affecting the

[30] Handing out tradable consumption permits allowing consumption of q* in total would lead to a market price equal to t. It is worth noting that this would also be the price for 'consumption rights' if those who are affected by the consumption decision were given the alienable right to be free from external effects and those wishing to consume the good had to obtain the right to do so, or the compensation for avoiding consumption if consumers were given the right to 'cause' external effects, and agree not to exercise this right in exchange for compensation, provided that there are no transaction costs (and income effects). For a comparison of different internalisation mechanisms see Varian 1992, Chapter 24, or Laffont 1987.

behaviour of those who are faced with a particular charge or tax, the less revenue one can expect to be generated from such charges. Or, conversely, the more likely a particular charge is to generate revenues because those who will have to pay it have little or no alternative, the smaller the impact of such a charge will be on behaviour. Such charges can either raise revenue or the price of undesirable behaviour, but not both.

Some of the conceivable aims have potentially conflicting definitions. For example, increasing fairness in the sense of "making the polluter pay" may conflict with ensuring a fairer income distribution. This is because who pays the internalisation charge (or Pigou tax) in the first instance tells us little or nothing about who will ultimately foot the bill: tax incidence is quite different from tax collection. For example, an internalisation charge increasing the cost of road transport may ultimately be reflected in higher grocery prices, which would disproportionately be paid by consumers on lower incomes who spend a greater proportion of their income on groceries. This suggests that the "fairness" objective put forward in the CE study is at the very least ambiguous, and there are indeed a number of notions of fairness which can lead to different conclusions with regard to the costs and benefits of particular internalisation scenarios.

More worryingly than ignoring the potential conflicts amongst the objectives of internalisation put forward in the CE study, however, is that there does not seem to be an explicit reference to the one overarching objective commonly associated with the internalisation of externalities, namely to remove inefficiencies and increase welfare. Although efficiency shines through in some places (and the CE study claims to take the objective of a "more efficient economy" as the "primary aim of internalisation", p 37), it is not included in the list of potential aims of internalisation. The "influencing behaviour" motive can be read as a proxy[31] – but if it were meant to refer to economic efficiency, it would have to acknowledge that it cannot be an aim in itself to reduce environmental impacts or allow a freer flow of traffic, because the welfare impact of doing so also depends on the associated cost. As the discussion of externalities as a reciprocal problem above has highlighted, any reduction in environmental impacts comes at a cost, and a freer flow of traffic implies that some users are being priced off and thus face the cost of having to look for alternatives. Ignoring these costs will inevitably lead to recommendations that fail to maximise welfare (and may even lead to welfare losses compared with the status quo). There is little in the CE study to suggest that such costs would be properly taken into account.

Starting from the notion of economic efficiency would clearly have helped in the identification of potential conflicts in objectives, and would certainly have highlighted the cost of pursuing certain objectives which are

[31] As stated in the CE study, the primary objective of achieving efficiency is "particularly related to influencing behaviour by providing optimal incentives (…) [although] other motives may be politically relevant and will be taken into account, but less central" (p 37).

incompatible with economic efficiency. It would have shown, for example, that raising revenues is often incompatible with achieving efficiency, because it would be best achieved by charging those who have little flexibility rather than those who can easily avoid the activity on which the tax has to be paid. This is likely to lead to a situation in which less external harm is avoided at a much higher cost than would be the case if internalisation had been aimed at efficiency without any consideration of tax revenues – and in which therefore welfare is lower than it could be, and potentially lower than in the absence of intervention.

It would also have become clear that there is a notion of fairness that is perfectly aligned with the notion of economic efficiency, namely the objective of avoiding competitive distortions between different modes of transport.[32] Any such distortion can be expected to result in inefficient outcomes, and thus reduce overall welfare. This notion of fairness – which appears to be the one underpinning Directive 2006/38 - would a priori rule out internalisation schemes that disadvantage some transport modes and favour others, or internalisation schemes that limit flexibility in terms of how particular reductions in environmental impacts are achieved. By contrast, an alternative notion of fairness, namely the notion of corrective justice, is potentially conflicting with efficiency. If internalisation were aimed at making those who "cause" harm pay for the damage suffered by those suffering from pollution, it might seem natural to use revenues from internalisation charges to compensate the pollution victims.[33] This might seriously distort the incentives of those receiving compensation, who would not obtain any benefit from trying to mitigate their losses.

In summary, the CE study fails to put forward a clear objective for internalisation which would allow one to assess different internalisation proposals, to recognise potentially conflicting objectives and to establish the welfare losses that would be associated with the pursuit of such objectives.

1.4.3 The CE study does not provide a comparative assessment of internalisation tools

The CE study also appears to be unclear about the potential instruments that are available for the internalisation of external effects. It states that "[i]n general, internalisation relates to market based instruments, and pricing instruments in specific [sic]" (p 37), but the internalisation scenarios specified include a mix of instruments (including, for example, a tightening and extension of Euro standards), and there seem to be clear preferences for alternative tools, e.g. in relation to noise (which "may be better addressed by regulation and standards", p 56).

[32] For a discussion of the concepts of fairness and efficiency, see section 2.6.4.

[33] Note that if revenues from internalisation charges were used to fund additional infrastructure investment, this might have the perverse effect of increasing the level of activity that is regarded as responsible for the external effect in the first place, and which would thus increase the harm suffered.

What is missing is a comprehensive list of instruments that can be used, and a comparison of these different instruments with regard to their effectiveness and their error tolerance (which is of particular importance given the substantial uncertainty about external costs implied by the vastly differing estimates collated in the CE study). For example, there seems to be a basic choice between imposing a carbon tax and putting in place a carbon emission trading scheme. Both of these systems are likely to differ with regard to their setup and operating costs and the ease with which they can be extended. The informational requirements are likely to be different: in order to set the correct level of a carbon tax, it is necessary to have relatively good information about abatement costs and the damage caused by greenhouse gas emissions. By contrast, a carbon trading scheme requires one to take a view over the total amount of greenhouse gas emissions that is sustainable. Both schemes differ with regard to the uncertainty faced by polluters – where tax rates may not change frequently, the price of emission permits can fluctuate considerably in the short run.

All of these factors need to be considered and analysed in order to make recommendations about appropriate internalisation scenarios. The CE study appears to be largely devoid of any such consideration, and merely notes that different instruments "differ to the extent that they allow consumers to make their own decisions, and to the extent that the results can be predicted beforehand. They also differ to the extent that they may be related to particular government motives and aims. For example, communication & information may influence consumer behaviour, but cannot be related to generating revenue" (p 37).

In addition, the CE study appears to be confused about the nature of particular instruments. For example, emission trading is listed as one of the most relevant *other* instruments in Table 13 of the CE study, even though it is clearly a market-based instrument relying on the pricing mechanism. Similarly, auctioning of airport slots is listed in the category of other instruments, together with speed limits, emission standards, zoning and other typical regulatory instruments, even though the former is clearly an intervention based on price signals. It is also worth noting that even though these other instruments figure prominently in the various scenarios developed in the CE study, they are not the subject of any detailed investigation.

Last but not least, there seems to be some general confusion in the CE study about who would pay internalisation charges – sometimes it is transport users, sometimes operators. Given that the effects of internalisation charges may be very different, the apparent lack of detailed analysis of these effects should be of concern. For example, in order to affect end-user demand, internalisation charges will have to be reflected in end-user prices. This might be problematic if end user charges are differentiated by vehicle kilometres, for example, and internalisation charges are levied on the basis of some other, more appropriate variable and payable by transport operators. Even if the additional costs are fully passed through to transport users in the form of higher per-kilometre charges, the impact of demand may be different from the impact that has been assumed when setting the internalisation charges. These imperfections and frictions

can have a substantial impact on the effectiveness of various instruments, and would need to be considered.

1.4.4 Second best issues, implementation costs and regulatory failure

The CE study notes that marginal cost pricing methodologies might not be "appropriate because they are only optimal under certain theoretical assumptions that are not satisfied in practice", namely that "marginal social cost pricing is applied through the whole network considered, the whole transport sector and even throughout the economy" and that "governments use lump sum taxes to pursue any redistribution targets they may wish to meet" (p 39). In economic terms, this means that marginal cost pricing will only maximise welfare if the presence of the externality which it is meant to internalise is the only deviation from the textbook model of a competitive equilibrium. Where there are other market imperfections (such as distortive taxes, information asymmetries, or market power), there is no guarantee that marginal cost pricing will increase social welfare.

While this insight is correct, the CE study fails to identify the implications for the assessment of different policy options. What form (and what level) of intervention to address externalities would be appropriate in a so-called "second-best" world (where some of the optimality conditions that define the "first-best" competitive equilibrium in which social welfare is maximised are violated) is far from clear, and would require detailed analysis, which appears to be completely lacking from the CE study.

For example, one of the concerns raised in the CE study with regard to marginal cost pricing schemes – namely that revenues might not be sufficient to cover infrastructure costs – is a typical problem analysed in a second best framework: if it is not possible to use two part tariffs to cover total costs, the appropriate solution would be to charge mark-ups over marginal cost that are inversely related to price elasticities in order to minimise the welfare losses associated with a deviation from the first best solution.[34] The CE study, by contrast, argues that "one might want to use the revenues of external cost pricing of one mode to cover the fixed infrastructure costs of other modes" (p 40), without providing any justification as to why this might be better than two part tariffs or some average cost pricing, and indeed without even mentioning these alternatives.

More generally, the CE study seems to mix reasons for why it might not be *desirable* to use marginal cost pricing and reasons for why it might not be *possible* to do so, and to conclude – without much, if any analysis – that the two potential problems with marginal cost pricing are of a similar nature and may even cancel each other out. At the risk of oversimplification, the view

[34] This principle is known as Ramsey-Boiteaux pricing; where the pricing principle is applied to more than one product or activity, the relevant elasticities have to take into account substitutabiltiy or complementarity, i.e. one has to use the appropriate super-elasticities.

underpinning the CE study appears to be that it does not matter that much that we cannot practically apply marginal cost pricing because it might not be optimal to do so anyway.

The CE study claims that, owing to the limited scope of a pricing scheme, pricing measures "could lead to much less positive welfare effects" (p 39), but does not seem to consider the possibility that intervention could actually lead to welfare losses relative to not intervening at all. Given concerns about second-best issues and the practical limitations to internalisation mechanisms, however, such an outcome is not inconceivable. Unlike in the world of textbook models, in the real world it is not sufficient to diagnose a problem such as the presence of external effects in order to justify intervention. The case for intervention also needs to consider that available instruments may not work perfectly, that there is substantial uncertainty and that there are implementation costs of intervention, which taken together may mean that intervention makes matters worse.

The discussion of "Options for incentive base" (p 46 ff) certainly does not include any systematic assessment of the impact that the different options can be expected to have on the likely benefits that would be created if they were implemented.

The CE study is equally naive about the implications of system requirements. It does note that marginal cost pricing would require charges to be differentiated according to cost drivers, which is unlikely to be feasible in practice, where "a limited number of easy measurable parameters as a proxy of cost drivers" (p 40) have to be used. It does, however, fail to assess to what extent the use of such proxies may lead to inefficiencies, in particular given the considerable uncertainty about external cost functions and levels (as reflected in the wide range of estimates presented in the CE study), and simply assumes that welfare gains can be achieved by using internalisation measures that are "built on good proxies for cost drivers" and are "[n]ot too complex so as to limit implementation and transaction costs."

There is no detailed analysis of implementation costs. The discussion of the potential need for a toll system for passenger cars, for example, simply notes that "it may not be worthwhile to introduce sophisticated (and relatively expensive) systems for kilometre charging that apply to the whole network for the purpose of reducing external effects" (p 61). Given that road transport is assumed to account for the large majority of external costs, and that passenger cars are responsible for a substantial proportion of the external costs of road transport[35], it is entirely unclear how such a conclusion can have been reached without a detailed assessment of the relative performance of such a comprehensive toll system compared with other internalisation proposals, including the cost of the different solutions.

[35] According to the CE study, two-thirds of the external cost of road transport are caused by passenger transport (p 27), the majority of which is presumably accounted for by individual car use.

Overall, the CE study appears to be infused with the belief that intervention can only lead to welfare gains, which may admittedly be larger or smaller depending on how internalisation takes place, but that intervention will never lead to welfare losses. This is an inappropriate starting point for the analysis of policy options based on an assessment of costs and benefits, which has to consider the possibility that intervention could actually lead to welfare losses. It should by now be well understood that, in order to justify policy intervention, it is not sufficient to diagnose market failure, but that one also has to demonstrate that the risk of regulatory failure (or non market failure) is sufficiently small to make sure that the cure is not worse than the disease.[36]

As the brief description of internalisation provided above has shown, if internalisation is to be based on anything other than a tax schedule which sets taxes equal to marginal external cost at every possible activity or consumption level (i.e. a tax schedule that exactly tracks the marginal external cost function[37]) it is necessary to know the value of external costs at the optimal activity level (i.e. the marginal external cost at q*). If internalisation were based on setting a constant tax rate, for example, choosing the wrong rate may not only lead to welfare gains that are lower than those that could be achieved through the correct tax rate (namely if the tax rate is set too low), but can actually lead to welfare losses (namely if the tax rate is set too high, and thus too much of the activity causing external harm is priced off). There is nothing in the CE study to suggest that this consideration has been taken into account, even though the CE study acknowledges that external cost measures are imperfect and subject to substantial uncertainty. Attempts to internalise external effects on the basis of some arbitrarily defined unit cost figure may do more harm than good – and proposals to internalise on the basis of cost figures that are not only unrelated to any consideration of optimality, but also vary by a wide margin, seem positively dangerous.

The failure of the CE study to take into account the potential welfare losses that could arise from intervention as a result of uncertainty about the underlying external costs, the fact that it is costly to avoid external effects (which is at the heart of the cheapest cost avoider principle), and the implementation cost of any intervention is particularly surprising as the need for regulatory impact assessments and a careful weighing of costs and benefits is an explicit principle of Commission policy. The Commission's White Paper on European Governance of 2001 stresses that policy proposals "must be prepared on the basis of an effective analysis of whether it is appropriate to intervene at EU level and whether regulatory intervention is needed. If so, the analysis must also assess the potential

[36] For an early, but comprehensive discussion of non-market failure, see Breyer 1979 and Wolf 1979.

[37] Even a tax that exactly tracks the marginal external cost function may not be sufficient to ensure optimal behaviour if the external cost function is not convex (see Varian, 1992, chapter 24).

economic, social and environmental impact."[38] The Commission's Impact Assessment Guidelines of June 2005 stress that a full assessment of the impact of proposed policies demonstrates the "Commission's openness to input from a wide range of external stakeholders, and shows its commitment to transparency. Further, by providing a careful and comprehensive analysis of likely social, economic and environmental impacts, both direct and indirect, it also contributes to meeting the specific commitments of the Lisbon and Sustainable Development Strategies"[39], and acknowledge the risk of regulatory failure, including examples of environmental regulation or intervention that is capable of distorting competition. The Guidelines also explicitly refer to "compliance costs" as one important category in the assessment of non-expenditure measures. Consideration of such compliance costs would directly lead to the recognition of the costs incurred by the polluter, and would suggest that it is important to balance avoidance costs and benefits - or, in different terms, an acknowledgement that intervention should be based on the cheapest cost avoider principle. Given that the need for a careful assessment of costs and benefits associated with any policy intervention is a well-established principle of community policy, it is difficult to understand why the CE study fails to undertake such an assessment.

1.4.5 Interrelation with other policy instruments

The CE study includes a discussion of how existing charges and taxes should be treated in relation to proposed internalisation methods (albeit in an unsystematic manner), and how revenues raised by internalisation charges could be used, but it does not provide a systematic and comprehensive analysis of the interrelation between potential internalisation measures and other policy instruments. In particular, the CE study considers whether existing charges already imply some internalisation[40], how they might be differentiated, and whether some of them should be reduced in response to the introduction of new charges, but it does not consider that there may be other policy instruments in place that are aimed at internalisation, which should therefore be taken into consideration, or that existing charges may be the cause of distortions.

---

[38] COM 2001, 428.

[39] European Commission 2005.

[40] It is worth noting that the way in which existing charges are treated is haphazard and not supported by analysis. For example, the CE study simply assumes that "if the revenues of fuel excise duties of a mode exceed the variable infrastructure costs of that mode (…) the external climate costs are (partly) internalised by the fuel excise duties. To make this more transparent, we label this part of fuel excises duties in the scenarios as CO_2 taxes" (p 52). There is no justification for this assumption. There is no explanation as to why any difference between total revenues and total costs should be considered to give rise to partial internalisation, nor why it should be the internalisation of external climate costs rather than other external costs. There is also no explanation why infrastructure costs should be covered by revenues from fuel duties in the first place, given that the amount of fuel consumed is not a good proxy for the drivers of variable infrastructure costs, and there is no consideration of the fact that infrastructure costs may be covered by other charges.

- An example for the first category are policies to subsidise public transport, which could potentially be aimed at internalisation by increasing the opportunity cost of using cars[41] - although such subsidies may also have the effect of making transport too cheap overall.
- An example for the second category are specific tax differentiations, such as for example the exemption of jet fuel from excise duties. This makes air travel too cheap relative to other forms of transport, distorting not only competition between the modes of transport, but potentially also increasing the total amount of emissions.

Recognising this interrelationship with existing policies is important for two reasons:

- First, the option of changing existing policies should be considered alongside potential internalisation scenarios, in particular where existing policies are responsible for distortions of incentives and the potential source of externalities. Removing the cause of distortions is likely to be a better option than treating the symptoms, but should in any case be considered alongside potential additional intervention.
- Second, internalisation measures may conflict with existing policies. Where such conflicts arise, internalisation will have costs that have to be considered when assessing different policy options. For example, where certain forms of transport are subsidised for particular public policy reasons (e.g. public transport in rural areas may be subsidised as part of regional policy), internalisation that increases the cost of transport will counteract or undo the effects of such subsidies and jeopardise the achievement of these policy objectives.[42]

These effects would have to be taken into account in order to properly identify the most appropriate internalisation scenario. They do not seem to have been considered in the CE study.

Last but not least, the CE study does not discuss the potential impact that the use of revenues from internalisation charges will have on behaviour. Although the CE study purports to consider the efficiency effects associated with using such revenues, this consideration is limited to a few remarks on whether, and in what way, transport projects might be funded. It does not consider, however, what effect the use of revenues might have on the incentives of transport users, nor whether the use of revenues for transport investments might not actually increase the level of external effects. There would seem to be no reason to expect that spending income from charges

[41] As noted above, internalisation can be achieved in many ways, with taxes being only one possible solution, and subsidies being another one.

[42] Of course, one response might be to increase the level of subsidies in order to maintain the regional policy objective. However, in this case there would be no price impact for end users, and thus no impact on behaviour. The only effect would be that the amount made available for subsidisation would have to increase, which might imply further efficiency losses if as a result distortive taxes were to increase.

aimed at internalising climate change costs on additional infrastructure, for example, would improve welfare rather than reduce it – and equally there is no reason to expect that limiting such investment to money raised from internalisation charges would be meaningful.

In summary, the CE study would not appear to be a sound and reliable basis for the evaluation of potential internalisation scenarios and policy options.

2 The polluter pays vs. the cheapest cost avoider principle

2.1 Introduction

The previous section shows that the CE study is in many respects flawed. Its fundamental weakness is however its failure to consider the cheapest cost avoider principle. Instead, it adopts a "polluter pays" approach. Using the example of road transport, this part of the study provides a theoretical analysis of the relative strengths and weaknesses of the "polluter pays" and the "cheapest cost avoider" principles.

Road transport causes many types of external costs: environmental pollution, road accidents, noise and congestion. It is generally held that these negative externalities give rise to a misallocation of resources which requires corrective measures from the government. The polluter pays and the cheapest cost avoider principles prescribe alternative ways for governments to react to this problem.

According to the polluter pays principle the polluter (i.e. the generator of the externality) should pay the bill of the external costs he produces. The government should impose a so-called "Pigovian tax" on the polluter in order to provide the necessary incentives for the latter to internalise the externality. Alternatively, in order to prevent damage from occurring, the government can follow a command and control approach and restrict activity levels (e.g. through speed limits, or prohibiting certain types of vehicles at specific times), or prescribe the installation of avoidance and abatement devices or alterations in the mode of operation.

On the other hand, the cheapest cost avoider principle requires that the party which can prevent the damage at the lowest cost take action[43]. According to this principle, if there are any preventive measures which cost less than the benefit of the damage that they avoid, then they should be undertaken, whether by the polluter or by the pollutee, and on the condition that they are the least costly means available to accomplish such a reduction. Means to reach this end can be financial charges, taxes, fines, liability or even command and control measures, such as regulated standards or zoning.

As will be outlined in detail below, the underlying economic logic of the polluter pays principle, i.e. the Pigovian way of economic thinking, turns out to be fallacious. As demonstrated by Ronald Coase in his 1960 seminal article, the mere existence of externalities does not, of itself, provide any reason for governments to induce polluters to take action, because the polluters might well be the highest cost avoiders. Thus the full internalisation of external costs is not always socially useful in that it does not necessarily maximise social welfare. There exist other institutional

[43] See Calabresi and Hirschoff 1972, p 1060 f.; Demsetz 1972, p 28.

solutions to which society can resort, for example a laissez-faire policy or business as usual (BAU).

Moreover, the polluter pays principle does not take into account the fact that externalities are caused jointly, and that the externality problem is of a reciprocal nature: to avoid harm to the pollutee would inflict harm on the polluter. The real problem is to avoid the most serious harm.[44]

As Donald Wittman puts it: "Until Coase, people talked about a driver causing harm to a pedestrian or factory smoke damaging laundry hung out to dry. After Coase's article, people realised that cause is an outmoded concept. Rather, the appropriate insight is to see that there are several inputs into the production of damage. In the automobile accident case, the care by the driver and choice by the pedestrian are both inputs into the accident. The question becomes, which combination of inputs is optimal? This, of course, depends on the marginal productivity of the input in reducing expected damage and the marginal cost of the input".[45]

We will see that in the cheapest cost avoider framework, "polluter pays" is one possible outcome of the analysis, but not a generally applicable principle.

Whereas Coase's teachings have had a considerable influence on the economic analysis of law, the theory of externalities and the New Institutional Economics in general, this does not apply to the fields of transport economics and transport policy, where the Pigovian way of thinking still dominates. As will be shown below, designing transport policies from this perspective can violate both efficiency and fairness.

The remainder of this part is organised as follows. Section 2.2 deals with the Pigovian tradition on which the polluter pays principle is based. This section also explains some fundamental terms and concepts. Section 2.3 presents the paradigm shift induced by the Coasean way of economic thinking and identifies the flaws of the Pigovian tradition. Section 2.4 discusses the cheapest cost avoider principle and its application. Sections 2.5 and 2.6 adopt a comparative institutions approach to compare the cheapest cost avoider and polluter pays principles.

2.2 Coase versus Pigou

2.2.1 The Pigovian tradition and the polluter pays principle

According to Pigou, when faced with actions that entail external costs, policymakers should develop mechanisms to internalise the externality. This implies forcing the polluter to consider in his individual cost-benefit calculus all the costs associated with his activity, including those that are imposed on others. Pigou believed that this internalisation could best be achieved by

[44] The reciprocal nature of the externality and the problem of joint causation will be discussed further on.

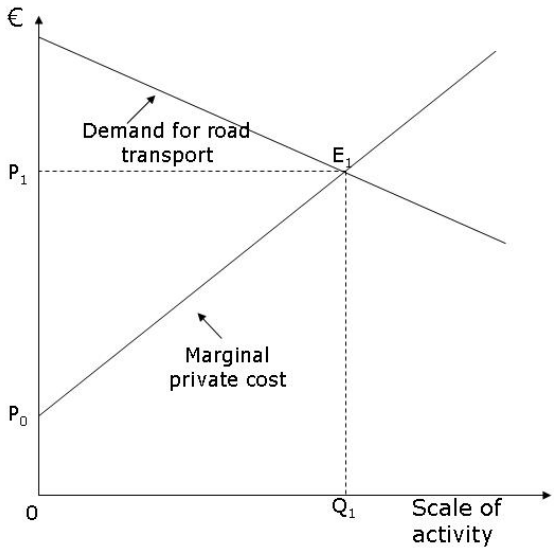
[45] Wittman 2006, p 54.

imposing a tax equal to the external costs on the polluter. The latter then has an incentive to conduct his activities up to the point at which his net benefit equals the tax. The external costs would thereby reach a Pareto efficient state: there exists no other feasible allocation which would make all individuals in the economy at least as well off, and at least one strictly better off.[46] Imposing a Pigovian tax is also considered as contributing to fairness as corrective justice, i.e. the polluter is held liable for the effects imposed on others.

To illustrate this point, consider the case of vehicles using a particular stretch of road, and assume that the noise caused by driving along this road harms local residents, which we will call the pollutees (or, taken as a group, the pollutee). Let the horizontal axis measure the scale of activity of road transport (for example the number of heavy duty vehicles using that stretch of road), and let the vertical axis measure the costs and benefits associated with various quantities (see fig. 2-1).

All actors are assumed to be risk neutral.

Figure 2-1 Demand and supply of road transport



In fig. 2-1, the demand curve slopes downwards: the higher the price of road transport, the less demand there is. The demand curve represents the marginal value of road transport to society: the more there is, the less one extra activity is worth to society.

The upward sloping curve represents the marginal private cost to the road transport industry, including wages and payments for petrol (we abstract from fixed costs). The marginal private cost curve slopes upward because

[46] See definition of Pareto efficiency in Lockwood 1987.

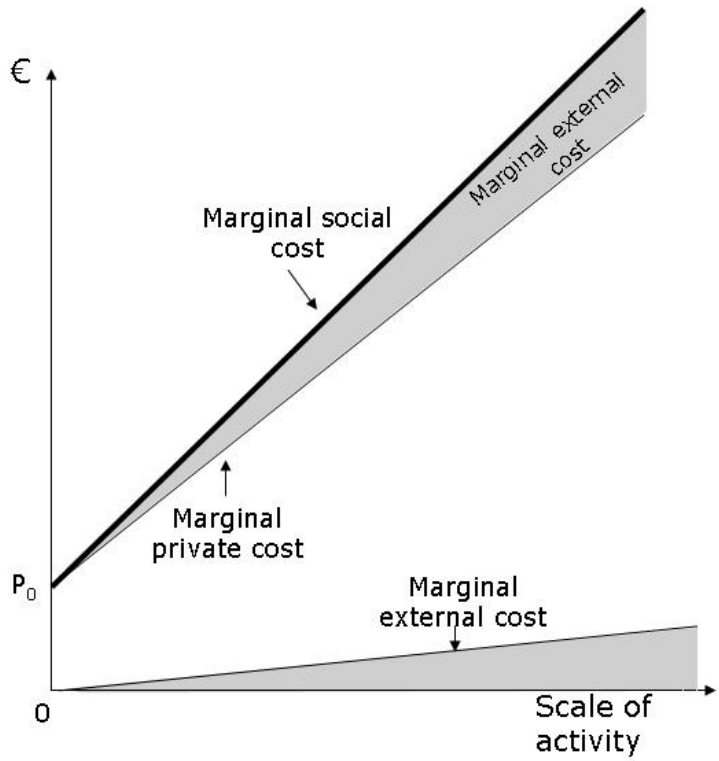
the more road transport takes place, the more one extra unit costs the transport industry.

Assume that the transport industry is perfectly competitive. In the short run, the firms' marginal private cost curves are their supply curves. Firms maximise profits by producing until private marginal cost equals marginal revenue (= price) in E_1.

If the industry were to produce more units, then the additional cost would be greater than its additional revenue and profits would fall. On the other hand, if the industry produced less than E_1, the lost revenue would be greater than the cost savings, also leading to a decline in profits. Hence, E_1 is the point at which the industry's profits are the greatest.

Thus, in equilibrium, the transport industry produces Q_1 units. Total revenues are price P_1 multiplied by Q_1, i.e. the rectangle $0Q_1E_1P_1$; total private cost is measured by the area under the private marginal cost curve, i.e. $0Q_1E_1P_0$. The total profits are total revenues minus total cost, i.e. the triangle $P_1P_0E_1$.

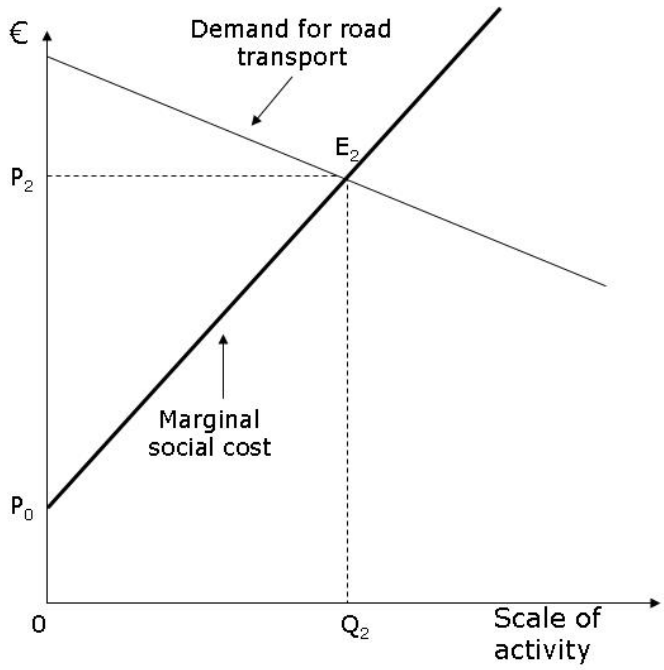
Figure 2-2 Marginal social cost and external cost



Road transport is a source of noise, which harms local residents.[47] The upward sloping marginal external cost curve in fig. 2-2 indicates the marginal cost to the residents of this pollution for each quantity measured on the horizontal axis. Since this cost is only borne by the residents, it is not part of the transport industry's cost calculation. It is a marginal external cost. In fig. 2-2 it is assumed that the marginal external cost increases with the quantity of road transport.

The bold printed upper curve represents the "marginal social cost" or "full marginal cost" of producing a supplementary unit of transport services. It is equal to the sum of the marginal private cost of the transport industry and the marginal external cost.

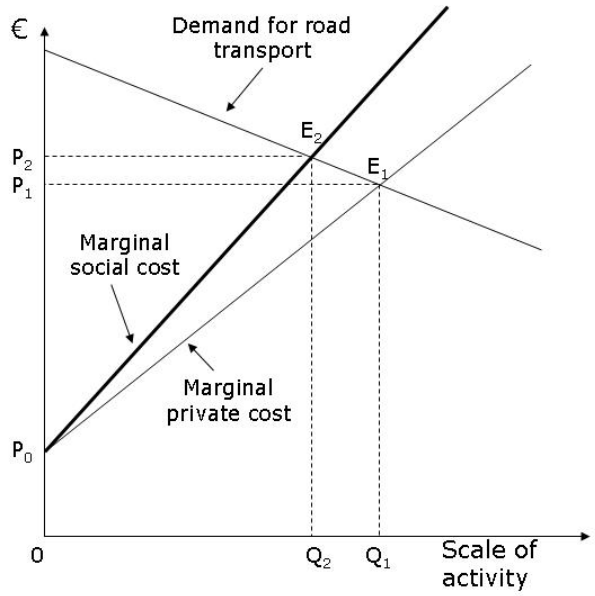
Figure 2-3 Social optimum



In fig. 2-3, the efficient outcome is at E_2, with Q_2 the efficient scale of activity, where the social marginal costs equal the marginal benefits (demand). Beyond E_2, the additional cost of transport is greater than its additional benefits for society; up to E_2, the extra benefit of a unit of transport is greater than its cost, and the quantity of transport should be increased in order to maximise welfare. Note that from society's point of view the optimal amount of pollution is not equal to zero.

[47] Noise is representative of any pollution causing harm. It is interchangeable with any other type of pollution.

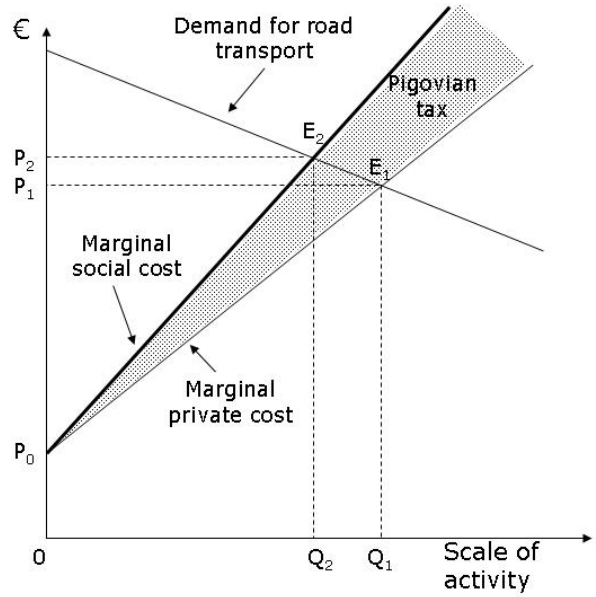
Figure 2-4 Private and social optima



However, as fig. 2-4 shows, the private optimum is not equal to the social optimum. From society's point of view, the scale of activity of the transport industry is too high when determined privately ($Q_1 > Q_2$). When the external costs are taken into account, it becomes clear that the industry should reduce its scale of activity from Q_1 to Q_2. This requires some kind of intervention, which can take the form of centrally fixing the scale of activity to Q_2. Alternatively, the regulator can modify the transport industry's incentives by means of a Pigovian tax.

The Pigovian solution to the overproduction depicted in fig. 2-4 is to make the firms operating in the transport industry pay a pollution tax (generally labelled Pigovian tax) which has the effect of internalising the externality. This tax is a marginal pollution tax in the sense that at the socially efficient quantity of transport services its amount equals the vertical distance between the social and the private marginal cost curves, i.e. the marginal external cost (see shaded area in fig. 2-5). The industry's total tax bill consists of the sum of marginal taxes for all infra marginal units, i.e. all units up to Q_2 - 1, plus the marginal tax for the last unit (at Q_2).

Figure 2-5 Pigovian tax



Assume that the government imposes a marginal tax that exactly matches the external costs for each quantity of road transport. Thus, the marginal tax curve tracks the marginal external cost curve. As a consequence, the industry's marginal cost is equated to social marginal cost; the industry's supply curve is now the social marginal cost curve. With such a pollution tax, the transport industry produces the amount Q_2, at which marginal cost (now including the tax) equals marginal revenue[48]. Below Q_2, the additional revenue is greater than the additional cost to the industry – the scale of activity expands, and beyond Q_2, the additional revenue is lower than the extra cost – the scale of activity diminishes. The Pigovian tax thus solves the externality problem.[49]

[48] Note that the same efficient quantity Q_2 can be realised if the government pays subsidies to the transport industry in order to make it reduce its production of emissions. From the point of view of the decision maker paying a marginal tax of, say, 10 is equivalent to foregoing a marginal subsidy of 10.

[49] The government can reach the same result with measures belonging to a command and control system. For example, one solution is to have the transport industry install noise reducing devices which would increase the private marginal costs such that the industry's marginal cost curve intersects the demand curve at quantity Q_2.

Robert Cooter pointed out a conceptual difference between taxes and regulatory measures. Whereas taxes are prices, regulatory prescriptions are accompanied by sanctions. According to Cooter, the price is the amount of money – here the tax – required by law for an allowed activity, while the sanction is a negative consequence associated with a prohibited activity (see Cooter 1984, p 1523). For Cooter only prices reflect the logic of an internalisation of external cost.

2.2.2 The Coase theorem: optimality without government intervention

Consider the following case: a rancher's cows stray onto a neighbouring farmer's land and trample the farmer's corn. If the rancher is liable for the damage to the farmer, will there be less damage to the farmer's corn than if the rancher is not liable, i.e. implying that he is entitled by law to let his cows stray? This question was addressed by Coase in what later became known as the Coase theorem.

Two initial property rights assignments can be distinguished: the rancher can be entitled to let his cows stray without being held liable by the farmer. Alternatively, the farmer can have the right not to have his corn damaged by acts or omissions of others.

If the rancher holds the initial property right, he can make a contract in which the farmer agrees to pay the rancher for not using his initial right to let his cows stray into the corn field.

Conversely, if the farmer holds the initial property right, in exchange for a payment, he can allow the rancher's cows to stray onto the farmer's land without holding him liable for the damage. Again, such an exchange is based on a contract between the rancher and the farmer.

Before moving onto the Coase theorem, it is necessary to define the crucial concept of transaction costs. A transaction can be considered as an exchange of property rights. Transaction costs are the costs of making and enforcing a transaction.

Transaction costs are differentiated into:

1) The costs of finding the holder of a property right.
2) The costs of negotiating an agreement. It costs the actors to bargain, since the efforts invested in the negotiation cannot be used for an alternative purpose (opportunity costs of time). If more than two people are involved (many participants on both sides) and unanimous agreement is required, then the bargaining costs increase dramatically and the likelihood of a successful agreement falls towards zero. This in turn produces costs in form of the lost opportunity to reap mutually beneficial gains from trade.
3) Since people can try to avoid respecting the contract, it is necessary to oversee their behaviour. This involves monitoring costs.
4) If a party to an agreement fails to honour the agreement, the other party has to find a way to enforce the contract. This generates arbitration or litigation costs.
5) Finally, costs can be incurred when one or both parties try to renege on the agreement.

In his seminal article "The Problem of Social Cost", Ronald Coase demonstrates that the Pigovian analysis is fallacious. In this section we will discuss two of Coase's arguments and apply them to the diagram.

If transaction costs are zero, then the upper curve in fig. 2-5 is the transport industry's "implicit" marginal cost curve, even if the transport sector is given the right to pollute[50]. The harm inflicted on third parties by a subject's activity constitutes a cost for the actor since it represents a lost benefit (gain) obtainable through a bargain struck with the victim. The victim is willing to pay a sum equal at the most to the harm he would otherwise suffer if the harmful action were carried out. If the actor carries out the harmful action, he loses this benefit. The activity thus comes at an opportunity cost. A rational actor will add this cost to his out of pocket costs when deciding whether to carry out the action or whether to come to an agreement with the victim.[51] There is no Pareto relevant externality[52], because the cost to the residents is internalised by the transport sector. There will be an efficient solution even without a tax. This is stated in what is called the "Coase theorem".

We will take up both points in turn.

One can state the Coase theorem as follows:

If transaction costs are negligible, then whatever the initial allocation of property rights,

a) The outcome will be efficient and

b) The outcome will be the same provided that the changing distribution of wealth does not affect consumption patterns.

Put differently, "with zero transaction costs, private and social costs will be equal" and "the value of production would be maximised".[53]

The logic behind the Coase theorem is very simple. If there are no transaction costs or no costs of exchange, an agreement will be found and all mutual gains from exchange will be exploited. The property right ends up with the person who values it most highly. More precisely, the right is "double sided": "One may have the right to perform a certain activity, or the right to keep that activity from being performed. Coase's theorem states that, in the absence of transaction costs and regardless of the initial attribution of claims, the right will always appear in the form that has the greatest value".[54] The outcome will be Pareto optimal, i.e. there is no way to make one person better off without making another worse off.

[50] See also Wittman 2006, p 51.

[51] See Baffi 2007, p 9.

[52] For this concept see section 2.3.4.

[53] Coase 1988, p 158. For a discussion of the Coase theorem, see also Cooter 1987 and Medema and Zerbe 2000.

[54] Baffi 2007, p 10.

2.2.3 A simple illustration

Let us illustrate the essence of the Coase theorem with a simple example. Suppose Adam has a theatre ticket but cannot go. Consider the following three scenarios:

1) Adam auctions off the ticket to Eve, who outbids Abel;
2) Adam gives the ticket to Eve;
3) Adam gives the ticket to Abel.

In scenario one the ticket ends up with the person who values it most. It is unlikely that a bargain could be struck in which Abel pays Eve to give up the ticket. The same holds for the second scenario. As for the third scenario, given that Eve would have outbid Abel in an auction, Abel should be able to sell the ticket to Eve.

This simple example shows that, whatever the original property right, the person with the highest willingness to pay gets the ticket. The final allocation is the same (part (b) of the theorem) and it is Pareto efficient regardless of the initial allocation. The total surplus is maximised – the ticket goes to the person who is willing to pay the most for it (part (a) of the theorem).

2.2.4 The zero transaction costs scenario

Why analyse a world of zero transaction costs? Several pertinent answers to this question can be given.

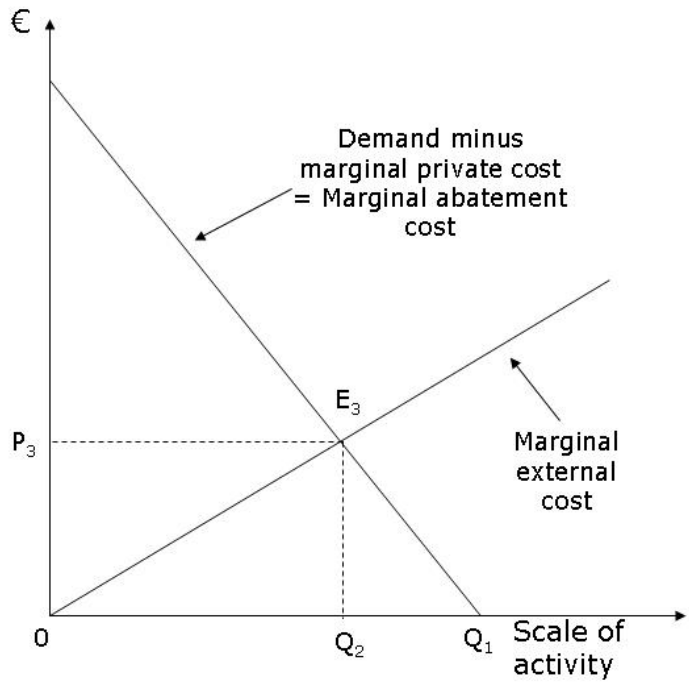
- "Economists, following Pigou whose work has dominated thought in this area, have […] been engaged in an attempt to explain why there were divergences between private and social costs and what should be done about it, using a theory in which private and social costs were necessarily always equal. It is therefore hardly surprising that the conclusions reached were often incorrect. The reason why economists went wrong was that their theoretical system did not take into account a factor which is essential if one wishes to analyse the effect of a change in the law on the allocation of resources. This missing factor is the existence of transaction costs."[55]
- In order to understand the crucial role played by transaction costs, one should analyse a world with zero transactions costs. Studying this mirror image of reality provides valuable insights into the functioning of a world in which this assumption does not hold.
- Policy proposals are usually based on a theory implicitly assuming zero transactions costs. However, if this assumption were true, Pigovian taxes would be superfluous. As Coase put it: "If there were actions that could be taken which cost less than the reduction in damage that they would bring, and they were the least costly means available to

[55] Coase 1988, p 175.

accomplish such a reduction, they would be undertaken."[56]Suppose that the transport industry has the property right of the use of the environment. Suppose further that both the transport industry and the residents who are negatively affected by the pollution it causes can be organised into groups that can be considered as unitary actors. This is possible if the costs of organisation – in essence these are also transaction costs, but transaction costs within a non market setting – are assumed to be low.

Consider again the figures presented in section 2.2.1. Fig. 2-6 is a slightly different representation of the same case.

Figure 2-6 Trading rights



The marginal external cost curve describes the extra cost incurred from a one unit increase of the scale of activity (see also fig. 2-2). The other curve describes the abatement costs, i.e. what it would cost the industry to reduce the scale of its activity by one unit. At Q_1 it incurs no extra cost: the industry is at its optimum. A reduction of the scale of activity from Q_1 however implies opportunity costs because the industry is not at its optimum. In other words, the area below the abatement cost curve describes how much of its rent the industry has to give up for a given scale of activity. His rent consists of the difference between what a factor of production earns in a given activity and what it could earn in the best alternative activity.[57]

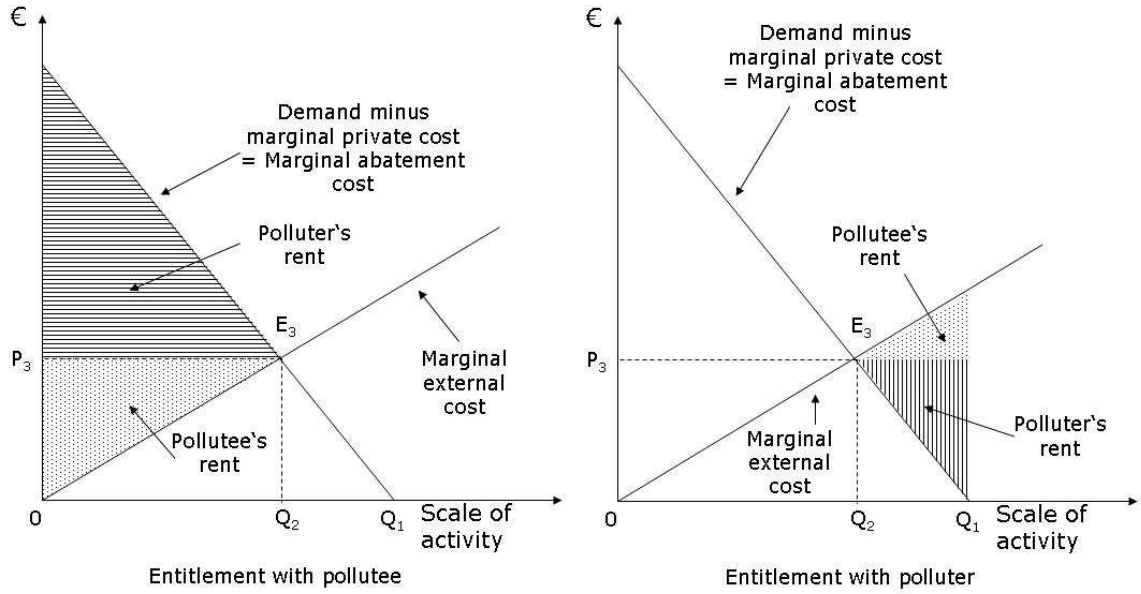
[56] Coase 1988, p 175.

[57] See Coase, 1988, p 163 – 170.

This curve is derived by subtracting the marginal private cost from the demand curve. Where demand equals marginal cost, i.e. at the private optimum Q_1, this difference is equal to zero. At a lower scale of production, demand is superior to marginal cost. The industry could increase its rents by satisfying the extra demand; if it does not do so, this implies opportunity costs (foregone rents and profits).

At E_3, the marginal external costs equal the marginal abatement costs: the extra cost of abatement equal the extra avoided external cost. Under the condition that transaction costs are negligible, this is a market solution, and it leads to the socially optimal scale of activity Q_2. Indeed, any solution above Q_2 would imply that the external costs suffered are higher than it would cost to abate. Below Q_2, the abatement costs would exceed their benefit, i.e. the reduction in external costs. The price of the trade in rights that leads to Q_2 is P_3, which equals $P_2 - P_1$ and also corresponds to the Pigovian tax.

Figure 2-7 Trading different rights



Suppose that the inhabitants hold the right to noise (left hand side, fig. 2-7). Then the marginal external cost is a supply curve, because it represents the marginal costs of the pollutee allowing pollution. The abatement cost curve represents the demand by the polluter to be allowed to pollute. Pollutee and polluter agree on the socially optimal quantity Q_2, where demand equals supply, at price P_3.

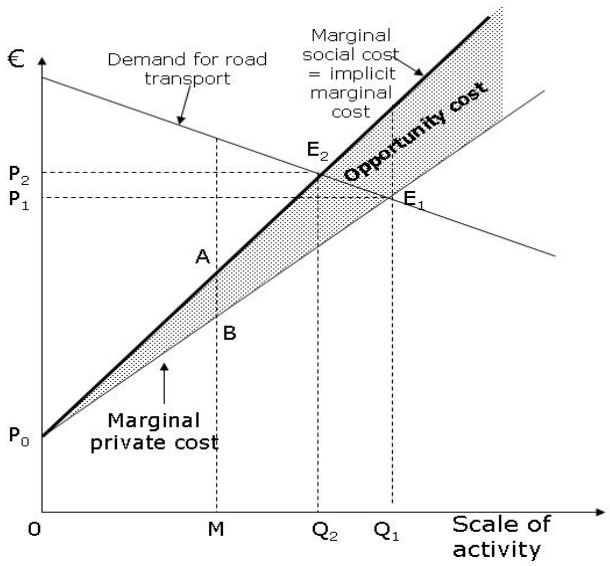
The pollutee gains the dotted area on the left hand graph in fig. 2-7. His rent equals his benefit, minus his cost, i.e. the external cost. The polluter gains the striped area in the same graph. Indeed, instead of paying a unitary price for the right to pollute, he would have been willing to pay up to his marginal abatement cost. Any price he pays that is inferior to the marginal abatement cost implies a rent.

Now assume that the entitlement lies with the polluter (see right hand graph, fig. 2-7). Now the abatement cost is the supply curve, to be read from right to left: it shows the marginal cost of supplying a reduction in the scale of activity. The marginal external cost curve is the demand by the pollutee for pollution reduction. Again, pollutee and polluter agree on the socially optimal quantity Q_2, where demand equals supply, at price P_3.

The pollutee's rent, which corresponds to the dotted area, is the difference between what he would have been willing to pay and the price he has to pay for a reduction in the scale of activity. The polluter's rent (striped area) is equal to his income from the sale of his rights $(Q_1 - Q_2)*P_3$, minus the cost of giving up the right, i.e. the abatement costs.

Note that whoever holds the entitlement, both parties gain from the trade. If the polluter holds the right, he can anticipate the trade, and will adjust to Q_2. The opportunity cost of renouncing from the trade is added onto his private marginal cost, leading to a quasi, or implicit, private marginal cost curve that equals the social cost.

Figure 2-8 Opportunity cost from renouncing the trade



To illustrate this point, consider output M in figure 2-8. The pollutees are willing to pay up to an amount of AB to prevent the transport industry from producing the Mth unit. Not accepting this payment implies incurring an opportunity cost which has to be added to the out of pocket costs of MB, and so on up to Q_2.

We have given a diagrammatic presentation of the Coase theorem, showing that when there are no transaction costs, the final allocation of the property right will be the same whatever its initial allocation. There is no Pareto relevant externality because the cost to the residents is internalised by the transport industry. Thus, allocative efficiency (maximum welfare) can be realised even without a tax.

2.2.5 The high transaction costs scenario

Road transport involves a high number of both polluters and pollutees. Thus, the analysis should not be based on the assumption that two unitary actors try to strike a bargain. Rather, we have to consider a multi party bargaining situation. Haggling about the division of the gains from trade and about sharing the costs of payment is likely to prevent a solution to the problem. This implies that transaction costs can be prohibitively high. Hence, the participants cannot easily sell and buy property rights. The initial distribution of property rights is likely to be also the final distribution.

Giving the transport industry the entitlement results in too many transport services being produced (Q_1), causing too high levels of pollution. Giving the residents the right to a pollution free environment results in too few transport services being produced (0), and in an inefficiently low level of pollution. Both property rights endowments fail to reach the efficient solution to the problem at hand.

In this case, the use of a Pigovian tax can make sense. Transforming the private marginal cost curve into the marginal social cost curve would lead the profit maximising transport industry to produce output Q_2, which is the socially optimal amount of transport services and pollution.

However, the graphical representation of the Pigovian tax solution can be very misleading, since there is nothing in the diagram to suggest that not only the polluter can take action, but that the pollutee might also be able to reduce pollution damage. Of course, the transport industry can reduce the external costs by decreasing production or by investing in noise avoiding devices, but the pollutees can also make investments to fight noise, like installing double glazing or moving elsewhere. The diagram does not illustrate that it takes two to "produce" damage. It concentrates on the behaviour of the polluter by indicating his costs, his benefits, and the costs that he imposes on the pollutees. The behaviour of the pollutees is only implicitly assumed in the cost curve of the pollutees[58]. Only if the residents' cost curve is calculated correctly does the Pigovian diagram yield an efficient solution. However, the diagram does not focus on the possibility that the pollutees might be in a better position than the polluters to reduce or avoid the external costs. Instead, it puts the burden on the behaviour of the polluters. Modelling a symmetrical situation in a way that hides this aspect can induce policy makers to commit serious mistakes.

It is now a generally accepted view that the most important message to be derived from Coase's seminal paper is not the so-called Coase theorem but rather its insistence on the fact that to reach efficiency it can be necessary to encourage the optimal behaviour of all parties involved in the production of damages. This insight is at the heart of the cheapest cost avoider principle.

[58] See Wittman 2006, p 53.

Box 2-1 offers an example which illustrates that one *possible* solution to an externality problem can be for the residents to move rather than for the polluter to reduce the pollution. Of course, before studying all the data of a case in question, it is not possible to say that this would be the optimal solution in a different case.

Box 2-1 Buying out a town

"According to the February 8, 2004 edition of the New York Times, sulphuric acid emissions from the massive coal burning power plant in Cheshire, Ohio, caused sore throats, burning eyes, and blisters. Sometimes the smog was so thick that cars drove through the streets at noon with the headlights on. In a series of town meetings in the spring of 2002, lawyers presented an offer from American Electric Power to buy all of Cheshire for $20 million. The 200-odd residents would have to move, their houses would be razed, and their community would cease to exist – and in exchange, they would each receive about three times the assessed value of their property. Though a few dissenters stood up and said they would rather fight than leave, they could not sway their neighbours. In the end, everyone accepted the offer and waived their right to sue."[59]

2.2.6 The Coase theorem, Pigovian taxes and welfare with zero transaction costs

In this subsection, we will demonstrate that a Pigovian tax can decrease or increase welfare depending on the circumstances. There is also a third possibility, i.e. to leave welfare unaffected. Remember, a rent consists of the difference between what a factor of production earns in a given activity and what it could earn in the best alternative activity. The latter amount is called the opportunity cost of the engagement of the factor of production in the activity under consideration. In fig. 2-1 the opportunity costs of the transport industry at producing Q_1 are the total private costs of production Q_1. The rent corresponds to the area between the y-axis, the marginal cost curve and P_1, i.e. $P_0E_1P_1$. In the words of Coase: "The factors engaged in an activity would be willing to pay an amount of money up to slightly less than the sum of their rents to allow their employment in that activity to continue. Even after taking this payment into account they would be better off than if they had to move to their best alternative. Similarly, they would be willing to abandon an activity in return for any payment greater than the sum of the rents, since, including this payment, they would be better off by moving to their best alternative than by continuing in this activity."[60]

The concept of economic rent can also be applied to residents. Their rent of living in the location is the difference between the value of the current location and the value of the best alternative location. If necessary, they would be willing to pay an amount of money up to slightly less than the sum

[59] Wittman 2006, p 53.

[60] Coase 1988, p 165.

of their rent to be allowed to continue living in that location. Even after taking this payment into account, residents would be better off than if they had to move to their best alternative. Note that the residents' best alternative is not necessarily to move: they could for example change their profession from the tourist industry to cow farming and thus avoid a specific damage from pollution. In what follows, we will stick to the alternative location story.

Similarly, residents would be willing to move in return for any payment greater than the sum of their rents (see box 2.1). For simplicity, we call welfare the sum of the rents properly defined of all factors engaged in an activity.[61] Welfare is maximised if the sum of the rents of all factors involved in an activity is maximised. As Coase put it: "Since the rents represent the increase in welfare (and therefore in income) from undertaking a particular activity rather than the best alternative, it follows that welfare, as measured on the market, is maximised when rents are maximised".[62]

If residents live in their particular location and there is no road transport, welfare corresponds to the rents of the residents. If the transport industry operates and there are no residents living at the particular location, then welfare resulting from road transport is measured by the rents of the transport industry. If there are both the transport industry and residents, but no damage to the residents, welfare is measured by the sum of the rents of the transport industry and the residents. Finally, if both the transport industry and the residents are present and if there is damage to the residents, welfare is measured by the sum of the rents of both the transport industry and the residents, minus the external costs.

In what follows, we define several scenarios and show for each scenario, firstly, that with zero transaction costs the Coase theorem holds, and secondly, what impact on welfare can be expected from the imposition of a Pigovian tax. We define the different scenarios by drawing on figures used by Coase[63]. Note that while the *chosen value* of the variables is arbitrary, the *relative values* of the rents follow a logical structure. Indeed, they reflect all possible constellations of all crucial variables. As a consequence, the following scenarios are exhaustive. Depending on the relative level of the rents of the transport industry and the residents, damage and the costs of other measures, it is shown that the maximisation of social welfare (the sum of the rents) can require action on the side of *either* the transport industry *or* the residents. Which of the two parties should hold the right to continue their activity (pollution or residence) depends on who can obtain the highest rent.

Table 2-1 summarises the rents, or the net benefits, of the actors. Emphasised rents designate that the rent is indeed obtained, i.e. the

[61] See also Coase 1988, p 165.

[62] Note that a situation in which welfare is maximised is also Pareto efficient. See Coase 1988, p 165.

[63] Coase 1988, p 166-170.

transport industry continues operating, or the residents continue to live in the same place. Damages (d) and the cost of erecting a wall (w) (modified scenario) always cost 50 and 5, respectively. A wall reduces the damage to 20. Highlighted scenarios are those in which a Pigovian tax can, under certain circumstances, be inefficient.

Table 2-1 Rents

Scenario/ Rents	1	**2**	3	**4**	5
Transport Industry (r^T)	**100**	100	25	**40**	30
Residents (r^R)	**100**	25	**100**	30	**40**
Structure	$d < r^T = r^R$	$r^R < d < r^T$	$r^T < d < r^R$	$r^R < r^T < d$	$r^T < r^R < d$

Scenario 1: $d < r^T = r^R$

If the transport sector is liable for the damage to the residents, it could compensate the residents and continue its operation and still be better off (by an amount of 100 - 50 = 50) than if it abandoned production. If the transport sector is not liable, the residents are unable to induce the transport industry to stop operating. Whatever the property right allocation, the transport industry continues to operate and the residents remain at their location. Welfare is 100 + 100 - 50 = 150, which is more than the 100 if either the transport industry discontinued their operation or the residents moved to another location.

Consider now a slightly modified version of scenario 1. Suppose that it costs 5 a year to build and maintain a wall alongside the road, thereby reducing damages to residents by 30 a year. If the transport industry is liable for the damage to the residents it would pay for building and maintaining the wall. If the transport sector holds the property right, the residents pay for building and maintaining the wall.

It follows that whatever the property right, the transport industry continues to operate and the residents remain at their location. Welfare is 100 + 100 - (5 + 50 - 30) = 175, which again is more than the 100 if either the transport industry discontinued or if the residents would want to relocate.

Next consider the imposition of a Pigovian tax. In the first variant of scenario 1 the tax would amount to 50. The transport industry would pay the tax and continue operating, and the residents would suffer 50 from the damage, but they would not move to another location. The tax would not have any real effect on the economy. It would simply redistribute rents from the transport industry towards the government. The same holds for the second variant of scenario 1. Here the Pigovian tax would be 20 (again a pure transfer of rents from the transport sector to the government), and the residents would pay for building and for maintaining the wall.

Scenario 2: $r^R < d < r^T$

Assume first that the transport industry is liable for the damage it causes. Under these circumstances, a bargain would be struck by which the residents would be induced to move to another location in exchange for a payment by the transport sector that is greater than the residents' rents, but less than the damage. The residents would be better off since the payment would be greater than the rent from their original location. The payment being lower than its damage liability, the transport sector would also gain.

Now assume that the transport industry holds the property rights on the environment, which means no liability on its side. Since the rents of the transport industry from continuing its activities are greater than the residents' rents, the latter would be unable to make a payment that would induce the transport industry to cease operating. Therefore, the residents would move to their next best location. It follows that a change in the legal position has no effect on the allocation of resources.

The resulting allocation maximises welfare. If both sides do not change their behaviour, welfare is 75 (100 + 25 - 50). If the transport industry discontinues operating, welfare is equal to the rents of the residents amounting to 25, while if the transport industry alone continues to be present welfare would be 100 (the rents of the transport industry).

Again, consider a slightly modified scenario that allows for a wall that protects the residents from part of the noise. By investing 5 a year for setting up a wall alongside the road, external costs can be reduced by 30. Whatever the property right, the wall will be set up and welfare will increase to 100 (100 + 25 - (5 + (50 - 30))), making both sides better off compared to the first variant of scenario 2.

Finally, consider imposing a Pigovian tax. In the first variant, the tax would amount to 50. As in scenario 1, the tax results in a simple transfer of rents from the transport industry to the government. The tax does not have any allocative impact; the situation is equivalent to a legal position in which the transport industry is not liable. The residents move to their next best location and welfare is maximised.

In variant 2 of scenario 2 the value of the production is maximised by a tax of 50. If the tax equalled the damage after having invested in the wall, the transport industry would pay the investment cost of 5, since this would make it better off compared to a situation in which tax equals 50: 100 - 5 - 20 = 75 > 50 = (100 - 50). The residents will not move to their next best location because their payoff is not negative: rent = 25 and damage = 20; 25 - 20 = 5. Welfare is: 100 - 5 + (25 - 20) = 100.

Scenario 3: $r^T < d < r^R$

If the transport industry is liable for the damage, the amount that the industry would have to pay to compensate the residents is higher than its rents. The transport industry would cease operation and the residents would remain at their location. Welfare would be 100. If the transport industry were not held liable for the damage, i.e. if it had the entitlement to pollute, an agreement would be found, according to which the transport industry

would cease operation and the residents would remain at their location. It is willing to do this in return for a payment that is higher than its rents. The residents would be willing to offer such a payment, provided that it is lower than the damage (which happens to be the case). Welfare would be 100. It follows that whatever the legal position, the outcome remains the same. Welfare is maximised, as the following reasoning reveals:

If the transport industry maintains its operations, and if the residents do not move to their next best location, welfare is 75 (25 + 100 - 50). If the residents do move, the increase in welfare would be 25 (rents of the industry), while if residents do not move and the transport industry ceases operating, welfare would be 100 (rents of the residents).

Again, consider a slightly modified scenario 3, in which a wall alongside the road – costing 5 – would reduce damage by 30. With liability on the side of the transport industry, it would be willing to pay 5 for a reduction of damage liability to 20. With no liability on the side of the transport industry, it would be in the interest of the residents to make this investment. Compared to the first variant of scenario 3 welfare remains the same: 100 (100 + 25 - 5 - (50 - 30)).

Finally, consider the imposition of a Pigovian tax. In the first variant it would amount to 50. Since the rents in the transport industry are 25, operation would cease. Welfare is maximised. In the second variant the tax would be 20. The transport industry would invest 5 in setting up the wall, thus reducing the tax bill from 50 to 20. Its payoff would amount to 25 - 5 - 20 = 0. Welfare is maximised: 100 (100 + 25 - (5 + 20)).

Scenario 4: $r^R < r^T < d$

If the transport industry is liable for the damage caused by its operation, it does not cease operating. Instead, it would be willing to pay the residents an amount greater than their rents (but less than its own rents) to induce the residents to move away from the damage to their next best location. This payment would leave both the transport industry and the residents better off. If the transport industry were not liable for the damage, the residents could try to induce the transport industry to cease operation. But since they can offer at the maximum slightly less than their rents and since their rents are lower than those of the transport industry, they cannot induce the transport industry to stop operating. It is thus best for the residents to leave their initial location.

It follows that the outcome would be the same whatever the legal position. Furthermore, the outcome would maximise welfare. If both the transport industry and the residents are present, welfare is 20 (40 + 30 - 50). If the transport industry stops operating, welfare would be of 30 (the rents of the residents). If the transport industry continues operating while the residents leave their location, welfare would be 40 (the rents of the transport industry).

Again consider a slightly modified case, in which a wall alongside the road – costing 5 - would reduce the damage by 30. With liability on the side of the transport industry, it would be willing to pay 5 for a reduction of its damage liability to 20. Since (30 - 20) > 0, the residents would not move.

With no liability on the side of the transport industry, it would pay the residents to make the investment and to remain in their current location. Whatever the legal position, welfare would be 45 (40 + 30 - 5 - (50 - 30)).

Again, consider the imposition of a Pigovian tax. Confronted with a tax of 50, the transport industry would cease operating (40 - 50 < 0), the residents would remain in their location, and welfare would be 30 (rents of the residents). It is inferior to the social rent of 40 that is obtained by giving the rights to the transport industry.

With an option to finance a wall, the transport industry would do so, thereby reducing its tax bill to 20. The residents would remain in their current location. The increase in welfare would be of 45 (40 + 30 - 5 - (50 - 30)).

Scenario 5: $r^T < r^R < d$

Assume first that the transport industry were liable for the damage caused by its operation. Since it would be willing to offer the residents to leave their current location (thereby reducing damage to zero) a maximum sum of slightly less than its rent, and since the residents would not be willing to leave unless they received slightly more than their rents (which are greater than the rents of the industry), the transport industry would be unable to compensate the residents for a move to their next best alternative. Thus the transport industry should compare its rents of 30 to the damage liability of 50. Obviously, it would decide to stop operating, which is efficient.

Now assume that the transport industry were not liable. In this case, the residents could avoid the damage by making a payment to the transport industry for ceasing to operate which would be slightly higher than the industry's rents, but below the value of the rents of the residents. Residents would still be better off than if they decided to leave.

It follows that whatever the legal position, the outcome would be that the transport industry would cease operating and that the residents remain in their location. A calculation similar to that in scenario 4 would reveal that the outcome would be such as to maximise welfare.

Again consider a slightly modified case, in which a wall alongside the road – costing 5 – would reduce the damage by 30. With liability on the side of the transport industry, it would be willing to pay 5 for a reduction of its damage liability to 20. Since (40 - 20) > 0, the residents would not move to their next best location.

With no liability on the side of the transport industry, it would pay for the residents to make the investment and to remain in their current location. Welfare would be 45 (40 + 30 - 5 - (50 - 30)) whatever the legal position.

Finally, consider the imposition of a Pigovian tax. Confronted with a tax of 50, the transport industry would cease operating (30 - 50 < 0), the residents would remain in their location, and welfare would be 40 (rents of the residents) .

With an option to finance a wall, the transport industry would do so, thereby reducing its tax bill to 20. The residents would remain in their current location. Welfare would be 45 (40 + 30 - 5 - (50 - 30)).

Summing up, the examination of these five scenarios leads to conclusive results. In each scenario the allocation of resources remains the same whatever the legal position. Furthermore, the outcome in each scenario maximises welfare, i.e. the sum of the rents of the transport industry and the residents minus the damage to the residents. In the originally defined scenarios, damage will only persist if it is valued at less than the rents of both the transport industry and the residents. If the damage is greater than the rents of either the transport industry or the residents, but not of both, the activity in which the rents are lower than the damages will be stopped. If the damage is greater than the rents of both the transport industry and the residents, the activity which yields the lower rent will not be undertaken. In the modified scenarios the damage will never persist; however, this result is due to the figures assumed in the analysis. Pigovian taxes can also maximise welfare, but they do not do necessarily so. They are imposed on the polluter, inducing him to take action to prevent the damage to the pollutees. This is inefficient and reduces welfare if the pollutees are the cheapest cost avoiders. We will address this problem in more detail later on.

It should be noted that the results derived from the preceding examples remain essentially unchanged if, instead of the question being solely whether the transport industry will operate or not or whether the residents will move to their next best location or not, one also allowed for the possibility that there could be more or less road transport activity or that there could be more or fewer residents living near the stretch of road.

2.2.7 Pigovian taxes and welfare with high transaction costs

When transaction costs equal zero, it makes no difference for society whether the law makes the transport industry or the residents liable, because the same efficient outcome will result. However, we have seen that when transaction costs are high, the initial assignment of property rights is likely to be the final allocation. Since not all allocations maximise welfare, law makers must be very careful when deciding how to allocate property rights in the first place. This is – as Coase explicitly stated – the most important message that his seminal article was intended to provide: "The world of zero transaction costs has often been described as a Coasian world. Nothing could be further from the truth. It is the world of modern economic theory, one which I was hoping to persuade economists to leave. I argued that in such a world the allocation of resources would be independent of the legal position."[64]

With high transaction costs, the law should endow the party that is likely to end up with the property right if there were no transaction costs. We have seen that transaction costs are usually high if more than two people are involved in a transaction and if unanimous consent is required. In these scenarios it would be virtually impossible for each firm in the transport industry to purchase the right to emit noise from each resident, even if the

[64] Coase 1988, p 174. See also ibid. p 15-16.

firm valued it more than the residents. If unanimous agreement is required, every resident would try to extract all the value for himself (what economists call the hold up problem). On the other hand, if the transport industry has the right to pollute without being held liable, agreements would most likely break down even if the residents valued a damage free situation sufficiently high as to compensate the transport firms for ceasing operation. Each resident would hope that the others would pay the transport firms to cease operating and would adopt a free rider position.

Thus, if it is efficient to avoid damage to the residents (and if there were no technical devices like walls, sound proof windows and so on), it would make sense to give the right to no damage to the residents (an approach called mimicking the market). In accordance, the Pigovian tax assumes that the property rights lie with the pollutees.

On the other hand, if the socially optimal decision is not to limit the transport industry, then we would want to allow the transport industry to inflict damage on the residents without holding it liable. In this case, the imposition of a Pigovian tax is no longer equivalent to the optimal allocation of the entitlement.

This is the basic flaw in the polluter pays principle. As will be shown below, in all situations in which welfare is greater if the polluter is not liable, the imposition of a Pigovian tax diminishes social welfare. To illustrate this point, we analyse the above scenarios assuming high transaction costs. What the analyses will show is that whether or not welfare is maximised through the imposition of a Pigovian tax depends on the circumstances of the particular case.

Table 2-2 summarises the scenarios. Bold scenarios are those in which a Pigovian tax would lead to an inefficient outcome.

Table 2-2 Scenarios summary

Scenario / Rents	1	2	**2a**	3	**3a**	**4**	**4a**	5	**5a**
Transport Industry r^T	100	100	100	25	25	40	40	30	30
Residents r^R	100	25	25	100	100	30	30	40	40
Damage d (with wall)	20	20	20	20	30	20	20	20	20
Cost of wall w	5	5	10	5	5	5	25	5	20
Structure									

Table 2-3 Relative rents

Scenarios	Relative rents
Scenario 1	$w < d < r^T = r^R$
Scenario 2	$w < d < r^R < r^T$
Scenario 2a	$w < d < r^R < r^T$
Scenario 3	$w < d < r^T < r^R$
Scenario 3a	$w < r^T < d < r^R$
Scenario 4	$w < d < r^R < r^T$
Scenario 4a	$d < w < r^R < r^T$
Scenario 5	$w < d < r^T < r^R$
Scenario 5a	$w = d < r^T < r^R$

Scenario 1

Assume that the transport industry's rent is 100, the residents' rent is 100, and the damage inflicted on the residents is 50. Welfare reaches a maximum of 150 in the status quo. Whether the polluter is held liable or not, and whether there is a Pigovian tax, the outcome is the same.

In the slightly modified scenario in which damage can be reduced to 20 by setting up a wall costing 5, the imposition of a Pigovian tax would induce the transport sector to invest in the wall, thereby increasing welfare to 175 (100 + 100 - 5 - (50 - 30)).

However, doing nothing would result in the same increase of the value of production, since the residents would make the investment.

Scenario 2

Suppose that the damage to the residents were valued at less than the rents of the transport industry, but at more than the rents of the residents. The following figures represent the case: rents of the transport industry are equal to 100, damage to the residents is 50, and the rents of the residents are 25.

We have seen that the residents should move to their next best alternative location in order to maximise welfare. They do this whether or not a Pigovian tax is imposed on the transport industry. The welfare, i.e. 100, would be realised.

In the slightly modified scenario, the imposition of a Pigovian tax would induce the transport industry to install the wall. The residents would not move, and welfare would remain at 100 (100 + 25 - 5 - (50 - 30)) which is the maximum. Without a tax the same result would be obtained since the residents would set up the wall, making them as well off as if they moved to their next best alternative location. Thus, whether or not the Pigovian tax is imposed, the outcome remains the same.

However, note the inefficiency of a Pigovian tax if the figures are slightly modified (scenario 2a). Assume that setting up the wall costs 10 instead of 5. Welfare would amount to 100 without a tax. With a tax, a wall would be set up, reducing welfare to 95. (100 + 25 - 10 - (50 - 30)). Thus, the imposition of a Pigovian tax clearly violates efficiency.

Scenario 3

Assume that the rents of the transport industry are 25, the residents' rents are 100, and the damage is 50. We have seen that in order to maximise welfare, the transport industry should cease operating. This would lead to welfare of 100. With a Pigovian tax of 50, the transport industry would decide not to operate. Thus a Pigovian tax contributes to efficiency. However, mimicking the market by allocating the property rights to the residents would produce the same result.

In the slightly modified scenario, the imposition of a Pigovian tax can induce the transport industry to set up the wall on the condition that the industry acts according to its group interests. Welfare would be 100 (100 + 25 - 5 - (50 - 30)). But again, mimicking the market by allocating the property rights to the residents would lead to the same outcome. Why, then, impose a tax?

Note the possibility of an inefficient outcome from imposing a tax if the figures are slightly modified (scenario 3a). Assume that the damage is reduced to 30 instead of 20. The investment in the wall is made, but welfare is reduced compared to allocating the property rights to the residents: 90 (100 + 25 - 5 - (50 - 20)). Thus it depends on the circumstances of a particular case whether or not the Pigovian tax is in accordance with the goal of efficiency.

Scenario 4

Assume that the rents of the transport industry are 40, the residents' rents are 30, and the damage is 50. Mimicking the market would require

allocating the property right of the environment to the transport industry. The residents would move to their next best alternative location. Welfare is 40, which is a maximum.

A Pigovian tax, however, induces the transport industry to cease operating. This is clearly inefficient, since welfare is only 30.

In the slightly modified scenario with a Pigovian tax, the transport industry would set up the wall and continue to operate, and the residents would not move. Welfare amounts to 45 (40 + 30 - 5 - (50 - 30)), which is a maximum. Note that, whatever the legal position, the same welfare would be realised.

Again, there is a possibility of a Pigovian tax creating an inefficient outcome if the figures are slightly modified (scenario 4a). Assume that the reduction of the damage by way of setting up a wall cost 25. In this case, the transport industry would stop operating, leading to welfare of 30. Here, it depends on the circumstances of a particular case whether the Pigovian tax furthers efficiency or violates it.

Scenario 5

Assume that the rents of the transport industry are 30, the residents' rents are 40, and damage is 50. From an efficiency point of view, the market should be mimicked, and residents' rents should be protected by allocating them the property rights. The transport industry would cease operating, which leads to welfare of 40, which is a maximum. Obviously, a Pigovian tax would lead to the same result. But then, why use a tax instead of an entitlement?

In the slightly modified scenario, whatever the entitlement, the same amount of welfare would be created, i.e. 45 (40 + 30 - 5 - (50 - 20)). A Pigovian tax would induce the transport industry to invest in the wall, leading to welfare of 45 (40 + 30 - 5 - (50 - 30)). But again the question arises, why use a Pigovian tax?

A tax can also result in inefficiency (scenario 5a). Assume that the wall cost 20. The transport industry would cease operating, which implies welfare of 40.

In summary, we find that the real effects of a Pigovian tax can be neutral in the sense that they can be equivalent to an adequate allocation of the property right. Depending on the circumstances of a particular case, a Pigovian tax can lead to an inferior result in terms of social welfare compared to an adequate allocation of property rights. In no scenario is the Pigovian tax superior to an allocation of property rights that mimics the market, i.e. which would result if transaction costs were zero.

Consider the case in which the Pigovian tax neither violates nor improves efficiency. Why use a Pigovian tax if designing an optimal tax requires the same information as necessary for optimally allocating property rights? An answer can be found in the literature on public versus private enforcement of law. Tax laws can be cheaper to enforce than private laws.

The preceding analysis has identified some flaws in the application of the "polluter pays" principle. To avoid this and other flaws of the "polluter pays"

principle, a general cost-benefit methodology has to be applied. That is exactly what the "cheapest cost avoider" principle amounts to.

2.2.8 The joint cause of damage

Consider the following situation[65]: The transport industry earns a rent of 100 and generates noise, but there is no damage, since no residents are located near the stretch of road. Of course, because we have no damage, the Pigovian tax would be equal to zero.

Now suppose that residents move to locations in the vicinity of the stretch of road, generating rents of 100, and that as a consequence the value of the damage created by the noise becomes 50 per annum. Residents may move because they count on the transport industry to install a noise prevention device costing 40 per annum. The noise prevention device may also consist in changing transport routes. The transport industry would install the noise prevention device, since this would enable it to avoid a tax of 50. Residents would not suffer any damage from noise, because it will be avoided. But this situation may not be efficient.

Suppose that for an additional cost of 20 the residents had been able to settle at another location that is equally satisfactory and noiseless. If there were no tax, the transport industry would continue to emit noise and welfare be greater by 20 (180 - 160). Welfare with a tax is 160 (100 + 100 - 40). Welfare without the tax is 180 (100 + 100 - 20).

Note that the residents generate a harmful effect of 40 on the transport industry (the costs of installing the noise protection device). In such a case, the institution of a double tax system would be desirable[66]: If the transport industry is to be made to pay a tax equal to the damage caused, the residents should be made to pay a tax equal to the additional cost incurred by the transport industry. Thus, the transport industry would have to pay 50 if the damage occurred, and the residents would have to pay 40.

Clearly, the residents would move to the alternative location, earning a rent of 80 (100 - 20). Consequently, there would be no installation of the noise prevention device (saving 40 per annum), no damage and no tax (to be paid by the transport industry).

Note that a double tax system would be required only if the Pigovian tax imposed on the polluters were based on damage.[67] However, if the tax is based on the fall in welfare elsewhere occasioned by the noise, the Pigovian tax would result in a maximised value of production.[68]

As we will see in the following section, this joint cause of damage approach is at the heart of the Coasean approach to externalities.

---

[65] For the basic idea, see Coase 1988, p 180 – 183.

[66] See Coase 1960, pp 151-152; Coase 1988, p 181.

[67] Coase 1988, p 181.

[68] Coase 1988, p 183.

2.3 Paradigm shift

In order to better understand why the cheapest cost avoider principle is superior to the polluter pays principle, we will outline up to date economic reasoning in more detail. The cheapest cost avoider principle approach will offer a more general view on the problem of the internalisation of external costs.

Modern economic theory on the efficient treatment of negative externalities goes back to Coase's 1960 seminal article where the problem created by externalities is not perceived as one of a polluter imposing costs on pollutees, but rather as a consequence of two or more actors competing for the same scarce resource. This reciprocity is now acknowledged to be at the heart of the problem of negative externalities.

In his 1960 article, Coase also undermines widely accepted notions about causation and corrective justice. Moreover, we are lead to question the generally held view that state action in form of regulation, financial charges or taxes is necessary for solving the externality problem.

As Ogus aptly puts it: "There is no a priori reason for assuming that, because the polluter's activity involves a physical interference, her claim on the environment is less valuable than that of the pollutees"[69].

Taking full account of the fact that the predicament of the parties involved in a negative externality is of a reciprocal nature, two further propositions can be derived.

- *First*: Any conflict about the use of the same scarce resource should be resolved according to which of the two conflicting uses has a greater social value. If, for example, the transport industry values the emission capacity of the environment higher than the residents value an emissions free environment, then the transport industry should have the right to emit pollutants. Vice versa, if the residents place higher value on their use of the environment than the transport industry, then they should have the right to an emissions free environment. In section 2.4 we will show that this prescription is an alternative formulation of the cheapest cost avoider principle.
- *Second*: The perception that, for every harmful action, a mechanism or institution is needed to ensure that the inflictor considers external costs as his own, i.e. to force the actor to internalise the externality, is fallacious. Instead, it can be preferable for social welfare to authorise a harmful activity without using devices to internalise external costs.
- It follows that negative externalities can be the result of the allocation of rights. The actor who holds the right is allowed to inflict harm on third parties. For example, he is allowed to drive, although the polluting effect of this activity is proven to cause harm. Thus, the whole logic of the necessity of internalisation must be questioned.

[69] Ogus 2006, p 167

In this section we describe the fundamentals of the paradigm shift in modern economic reasoning.

The new paradigm is organised around eight principles:

1) There is no negative externality without rivalry over the use of a scarce resource.
2) The problem of negative externalities is of a reciprocal nature.
3) Some externalities are the result of exclusive rights: only one of the parties has the right to the resource.
4) Only Pareto relevant externalities are socially desirable to eliminate.
5) Harm is not caused unilaterally by the polluter but is always jointly caused.
6) Efficiency requires double responsibility at the margin.
7) If transaction costs are zero, no mechanism is needed to ensure the internalisation of external costs. The solution of the problem can be left to the involved parties. The bargain they strike will be efficient.
8) If transaction costs are high, property rights should be assigned to the party who values them highest, which implies that harm is inflicted on the cheapest cost avoider.

It should be noted that these principles are mostly neglected by the polluter pays principle, but fully taken account of in the cheapest cost avoider principle. We now discuss principles 1-6 in greater detail. Principles 7 and 8 do not need to be elaborated upon; they have already been treated in section 2.

2.3.1 There is no negative externality without rivalry

As already mentioned in section 2.2, the Pigovian tax diagram (see fig. 2-5) misleadingly suggests that the origin of the externality problem lies in the unilateral infliction of costs by one economic actor on another. The action of the polluter is considered as the cause of the negative externality. Inflictors of damage should be required to compensate victims, or to pay a tax or financial charge, thus internalising the externality.

The polluter pays principle is based on the paradigm that since the polluter is the unique cause of the damage, he must be held liable. However, as Coase demonstrated in 1960, this view is based on an obsolete concept of cause. Since Coase's article we perceive the problem created by negative externalities not to be that of a polluter imposing costs on pollutees, but instead as the consequence of two or more actors competing for a resource. It is the existence of conflicting demands on the use of the same scarce resource which causes negative externalities. There can be no negative externality without this rivalry of use.

To illustrate: as a by-product of her smoking activity, a smoker sharing a room with a non smoker needs the air to deposit smoke, whereas the non smoker wants to breathe smoke free air. A rancher wants to use a piece of land to allow the cattle to stray, while a farmer wants the land free from cattle to grow crops. The transport industry wants to use the environment to deposit noise and other pollutants as by-products of its activity, whereas residents want the environment to be free from those emissions.

Again, one cannot grasp the concept of negative externalities without realising that they are caused by contradictory claims on the use of the same scarce resource. Each side wishes to use the scarce resource in ways which maximise its utility or profit.

2.3.2 The reciprocal nature of the externality problem

Probably the most important part of the Coase paradigm goes back to a criticism stating that the traditional Pigovian approach neglects the "reciprocal nature of the problem".

In the words of Coase: "The traditional approach has tended to obscure the nature of the choice that has to be made. The question is commonly thought of as one in which A inflicts harm on B and what has to be decided is: how should we restrain A? But this is wrong. We are dealing with a problem of a reciprocal nature. To avoid the harm to B would inflict harm on A. The real question that has to be decided is: should A be allowed to harm B or should B be allowed to harm A? The problem is to avoid the most serious harm."[70]

To illustrate, consider the previous example (see also the model in part I of this study).

Giving the transport industry the right to emit noise (or more generally, to pollute) allows it to inflict harm on the residents. But avoiding this damage would necessarily inflict harm on the transport industry.

It follows immediately that if there are competing claims on the use of a scarce resource, it is impossible to avoid all harm. Society is confronted with a situation usually referred to as a *tragic choice*. Whatever the decision, i.e. the allocation of property rights, harm will be produced. Thus, the problem of dealing with negative externalities becomes more complicated. The solution to the problem cannot simply be to prohibit harmful activities, since this would necessarily result in creating harm, measured as the loss in benefits that such activities yield.

As Coase points out, because it is inevitable to generate harm, the solution to the problem of correcting harmful externalities can only be to avoid the most serious harm.

As the analysis of the different scenarios in section 2 reveals, this can mean not to interfere with the activities of the parties involved, but instead to invest in devices suitable to reduce the harmful effect.

2.3.3 Some externalities are the result of exclusive rights

We normally think of a negative externality as the result of an activity that the actor does not have the right to carry out. But this is not correct.

Taking the insight of the "reciprocal nature of the problem", i.e. that externalities are the result of exclusive rights[71], seriously, implies that

[70] R. Coase, 1960, p 2.

[71] See Baffi 2007.

authorising the transport industry to emit noise *allows* it to harm the residents. On the other hand, recognising the residents' rights to be free from noise *allows* the residents to harm the transport industry.

The difference lies only in the fact that, in the first case, the externality is a by-product of action taken by the transport industry, whereas in the second case, the externality derives from the right to prevent the transport industry from carrying out activities that are beneficial to it.

Thus, Coase's reflections highlight flaws in the conception of externalities that scholars and politicians usually adopt, i.e. that externalities only arise when an actor invades the physical space of another person's property (physical things, life, and mental condition (Hume)), or the "protected domain", as Hayek calls it. However, there is a tendency to neglect the negative externality imposed on those whose actions are prevented. This is all the more surprising since the theory of property rights tells us that property, which is the very right to exclude others from using an item, always involves a negative externality in the form of a sacrifice borne by the subjects who might have otherwise utilised the item.[72]

2.3.4 It is socially desirable to eliminate only the Pareto relevant negative externalities

One occasionally encounters petitions in favour of the absolute prohibition of actions causing harmful effects, or of the internalisation of all external costs. Even from the point of view of Pigovian tradition, these proposals are flawed.

Consider Fig. 2-5. A Pigovian tax leads the transport industry to adopt the activity level Q_2. However, at Q_2, there remain Pareto-irrelevant external costs amounting to the shaded area up to Q_2. Moreover, note that a Pigovian tax equal to the marginal external costs at the activity level of status quo, i.e. Q_1, would lead to an inefficiently low activity level of the transport industry. A Pigovian tax should equal the marginal external costs at the optimum.

In order to correct this problem, one should distinguish between "Pareto relevant externalities" and "Pareto irrelevant externalities".[73]

According to Buchanan and Stubblebine, a Pareto relevant externality exists when the marginal costs inflicted on third parties are greater than the net marginal benefit obtained by the actor. An externality is called Pareto irrelevant if the opposite holds. Put differently, a Pareto relevant externality exists whenever its removal results in a Pareto improvement. It would be inefficient to eliminate Pareto irrelevant externalities.

This lesson also applies to optimal action in the Pigovian tradition. However, the Pigovian tradition focuses only on one side of the medal and neglects the other. As Baffi puts it:

---

[72] See Baffi 2007; Nicita et al. 2007.

[73] See Buchanan and Stubblebine 1962.

"Where the law gives a subject the right to prohibit a certain activity, the problem of externalities arises as it would in any other case, and it is possible that the private cost which the holder of the veto power manages to avoid by exercising his power is in fact lower than the lost utility for the subject who sought to carry out a certain action. Pigou seems to be referring in fact only to activities that cause harm to third parties."[74]

However, as we have seen in the previous sections, giving someone the right to prevent harmful activity can result in an inefficient outcome. Since exercising this right implies imposing harm on another person, this right should only be bestowed if the costs that the holder of the right otherwise would have to bear are greater than the net benefit forgone on the side of the party whose action is prevented.

2.3.5 All harm is jointly caused

The polluter pays principle is tied to a naive predetermined concept of causation which assumes that only activities can cause harm. However, harm is jointly caused.[75] We will call this approach the enlightened concept of causation, as opposed to the naive principle of causation.[76]

The consequences of the enlightened concept of causality are far reaching: if it is correct that the amount of harm derived from a certain activity also depends on the activity carried out by the victim, then an efficient solution would also require the victim to take precautionary measures.

In the example used for the scenarios, both the transport industry and the residents can take action. Can it really make sense to say that the presence of the residents is as much a "cause" of their damage as the behaviour of the firms in the transport industry? At the first glance, this sounds odd. The traditional and naive analysis of causation, based for example on instinctive notions of corrective justice, would deny the responsibility of the resident. But as Ogus reminds us: "further reflection should lead us to recognise that there are many situations where traditional causation analysis and instinctive notions of corrective justice are insufficient to address the complexity of the issues and interests at stake. The determination of when

[74] Baffi 2007, p 8.

[75] See Coase 1960, p 11.

[76] In economic models the naive principle of causation is represented by a variable controlled by one person that appears in the utility or production function of someone else (see Cooter and Ulen 1995, p 266-267).

Let $U_A = U_A(S,X)$ be the utility function of A, where S denotes the amount A smokes, and X indicates all other variables affecting A's utility. B's utility depends upon his health (H) and wealth (W), written in functional form $U_B = U_B(H,W)$. Assume that B's health is affected by A's smoking, thus $U_B = U_B(H(S),W)$. As one can see, variable S, controlled by A, appears in both utility functions, which means that both functions are interdependent. Interdependent utility functions are a suitable example to represent both the idea of cause and externality formally. Similar interdependencies can exist between utility functions and production functions or between production functions.

individuals should be held liable for negligent omissions, that is a failure to act, rather than positive acts, provides an excellent specific example".[77]

That both the transport industry and the residents are simultaneously the cause of the externality can be proven by drawing on the "but for" test which is generally applied in tort law.

Let A be an action and B an event. To decide whether action A is the cause of event B, the test requires asking the following question[78]: "But for A, would B have occurred?" If the answer to this question is a No, then A is considered being the cause in fact of B. Is the answer to this question a Yes, than A is not the cause in fact of B.

To illustrate: If the residents were not present, the noise would not harm anyone; if the transport industry did not operate, there would not be any harm either. Hence, both parties satisfy the "but for" test: The externality would not have occurred but for the presence of both the transport sector and the residents.[79]

Legal scholars use a limited version of the "but for" test in order to avoid endless causal chains. We will see in what follows that there is an economic equivalent to the "but for" test that defines the limits of the application of the concept.

2.3.6 Efficiency requires double responsibility at the margin

Cases in which both the inflictor of a harmful effect and the victim are required to take action for getting an efficient outcome are usually termed "cases of bilateral precaution".[80] In these cases the efficiency condition is called double responsibility at the margin.

Robert Cooter puts it this way: "When each individual bears the full benefits and costs of his precaution, economists say value is internalised. When an individual bears part of the benefits or part of the costs of his precaution, economists say that some social value is externalised. The advantage of internalisation is that the individual sweeps all of the values affected by his actions into his calculus of self interest, so that self interest compels him to balance all the costs and benefits of his actions. According to the marginal principle, social efficiency is achieved by balancing all costs and benefits. Thus, the incentives of private individuals are socially efficient when costs and benefits are fully internalised, whereas incentives are inefficient when some costs and benefits are externalised. In situations when both the injurer and the victim can take precaution against the harm, the internalisation of

[77] Ogus 2006, p 8; see also p 168-178, dealing with the issues of liability for negligent omissions and coming to a nuisance.

[78] See Cooter and Ulen 1995, p 265.

[79] In some cases the "but for" test to determine causation can be useless or misleading. Pertinent examples are situations with multiple causes, redundant causes and sequences of events (see Cooter and Ulen 1995, p 265-266; Miceli 1997, p 9, 22-25; Friedman 2000, p 191-197.)

[80] See Baffi 2007, p 12.

costs requires both parties to bear the full cost of the harm. To illustrate, suppose that smoke from a factory soils the washing at a commercial laundry, and the parties fail to solve the problem by private negotiation. One solution is to impose a pollution tax equal to the harm caused by the smoke. The factory will bear the tax and the laundry will bear the smoke, so pollution costs will be internalised by both of them, as required by social efficiency. In general, when precaution is bilateral, the marginal principle requires both parties to be fully responsible for the harm. The efficacy condition is called double responsibility at the margin."[81]

The problem is how to make this efficacy condition work. For those who think along the lines of corrective justice, the solution is simply to hold the inflictor of the harm liable, for example by following a rule of strict liability. This rule states that the inflictor of external costs should fully compensate the victim for the damage, without reference to a level of care[82]. However, in cases of bilateral causation, the application of such a liability rule does not produce the optimal incentives for the victim and gives rise to what Cooter called the "paradox of perfect compensation". This paradox refers to a problem known in the literature on insurance as "moral hazard": perfect compensation means that the victim is fully insured against any damage. He does thus not have any incentive to take suitable actions in order to mitigate the damage. On the contrary, he has incentives to act carelessly. Only if he victim has no influence on the occurrence of the damage, i.e. in the unilateral case, can perfect compensation lead to an optimum.

Double liability at the margin can be achieved through various mechanisms[83]: a system of fault liability or strict liability with a defence of contributory or comparative negligence. However, these systems do not achieve the complete internalisation of external costs[84].

Interestingly, Pigovian taxes do achieve the complete internalisation. They create a system of *decoupled liability*, which is characterised by the fact that the payment the generator of the harmful activity must make differs from the payment received by the victim.[85] In the Pigovian system the injurer pays a tax and the injured receives nothing. This asymmetry solves the problem efficiently, since both the injurer and the injured bear the costs of their harmful activity: the injurer compares his marginal tax with his marginal avoidance costs, whilst the injured bears the harm, and thus has an incentive to take action in order to mitigate the damage if the costs of

[81] Cooter 1985, p 3.

[82] We use the expressions "care" and "precaution" indeterminately to designate actions that can be undertaken by the actors in order to reduce the damage caused by the externality.

[83] See Baffi 2007, p 14.

[84] See Baffi 2007, p 14.

[85] See Baffi 2007, p 15, n. 25.

these actions are less than the harm.[86] Thus, a Pigovian tax can solve the two-sided moral hazard problem.

2.4 The logic of the cheapest cost avoider principle

When damages can be avoided or mitigated by either of two parties, the "cheapest cost avoider" principle (CCAP) suggests that the party which could have prevented or mitigated the damage at the lowest cost should take action, on the condition that the cost of preventing or mitigating harm is lower than the benefit. This is the ex ante version of the cheapest cost avoider principle. The ex post version refers to a situation in which a damage has already occurred. According to the cheapest cost avoider principle, if either of two parties can reduce the occurred damage, the party which is able to do so with the lowest cost should act, as long as this cost is lower than the benefit.[87]

The application of the cheapest cost avoider principle requires four steps:

- The first step consists in identifying the possible actors who can influence the outcome. Possible actors can be the polluters, the pollutees, or a third party, like government.
- The second step identifies alternative ways in which the outcome can be altered.
- In the third step the minimum costs of the various methods figured out in step two are calculated.
- In the fourth step the least cost method and the actor connected to it is chosen.[88]

Note that if the benefits of taking action are lower than the costs of the least cost method, nothing should be done. To see why, refer to scenario 2a above.

One can distinguish between *joint care* and *alternative care*.[89] With joint care, the care expenditures of the parties are complements, i.e. both must incur costs, either in fixed or in variable proportions, to reach the appropriate care. This problem is known from team production theory.

In the *alternative care* case expenditures are substitutes, i.e. more care by one party makes care by the other party less productive.[90] In this model two

[86] See Baumol 1972.

[87] The cheapest cost avoider principle was proposed by Calabresi 1970, taking up ideas from Coase 1960 and Demsetz 1972, 1974; see also Calabresi and Hirschoff 1972.

[88] After the implementation of the chosen method it must be checked whether the results are as expected and whether a revision of the policy is required. Note that this four-step procedure fits well with the key analytical steps in impact assessment (see European Commission 2005, p. 4).

[89] See Dari-Mattiacci and Garoupa 2007, p 3.

[90] See Brown 1973, p 223 – 349.

care expenditures can be *perfect* substitutes or *imperfect* substitutes. In the case of *perfect* substitutability the socially optimal care level of one party is zero if the other party takes care.

Often the least cost avoidance problem is seen to be present exclusively in cases of *strict* substitution, i.e. both parties can take care but only one of them – the cheapest cost avoider – should do so.[91]

We do not follow this restrictive course, since a least cost avoidance problem also exists at the margin, i.e. in cases of imperfect substitution of care expenditures.

Identifying the cheapest cost avoider is one thing; it is another problem to make sure that the cheapest cost avoider rather than the highest cost avoider has the incentive to behave in an optimal manner, or, in case that both should take care, that the optimal combination is realised. We will take up this issue after having analysed the working properties of the cheapest cost avoider principle. We start with a discussion of cases in which the level of care is a strictly binary variable, followed by a case where it is continuous.

2.4.1 Cheapest cost avoidance: care as a strictly binary variable

To illustrate the implications of the cheapest cost avoider principle and at the same time to provide the necessary information for a comparison of the strengths and weaknesses of the polluter pays principle and the cheapest cost avoider principle, we take up the scenarios described earlier on.

Scenario 1

Avoiding the damage of 50 – the benefit from taking action – is less than the costs of 100, i.e. the rents foregone either on the side of the transport industry or the residents. Thus the cheapest cost avoider principle suggests to do nothing. Welfare reaches a maximum of 150 in the status quo.

In the slightly modified scenario in which damage can be reduced to 20 by setting up a wall costing 5, the wall should be set up. This would increase welfare to 175 (100 + 100 – 5 – (50 – 30)), which is a maximum given the circumstances of the particular case. It is of no importance who sets up the wall.

Scenario 2

Clearly, the residents are the cheapest cost avoiders. The maximum value of production, i.e. 100, would be realised. In the slightly modified scenario, there are two lowest cost methods to realise the benefit, i.e. avoidance of the damage of 50: moving the residents and setting up the wall. Both methods lead to the maximum welfare of 100.

In the slightly modified scenario 2 a, in which setting up the wall costs 10 instead of 5, moving the residents remains the lowest cost method.

[91] See Dari-Mattiacci and Garoupa 2007, p 3; Burrows 1999, p 227 – 244.

Of course, the benefit of having a wall is greater, i.e. 50, than its cost, i.e. 10. However, using this method would reduce welfare to 95.

Scenario 3

The analysis is symmetrical to that of scenario 2. Now, the transport industry is the cheapest cost avoider; it should cease operation.

In the slightly modified scenario 3, both setting up the wall as well as ceasing operation are lowest cost methods. If the cost of the wall were 10 instead of 5, ceasing operation would be the lowest cost method to avoid the damage of 50. The same conclusion holds for scenario 3 a, in which cost of the wall is assumed being 5, but the damage is reduced to 30 instead of 20.

Scenario 4

Here, the cheapest coast avoiders are the residents; they should move to their best alternative location. Welfare would be 40, which is a maximum.

In the slightly modified version of scenario 4, the least cost method is setting up the wall, resulting in a welfare of 45, which is a maximum.

If the cost of the wall were 25, as assumed in scenario 4 a, and the reduction of the damage is 30, the residents are the cheapest cost avoiders; they should move to their best alternative location. Welfare is 40, which is a maximum.

Scenario 5

This scenario is the mirror image of scenario 4. Now, the transport industry is the cheapest cost avoider; ceasing operation generates welfare of 40, which is a maximum.

In the slightly modified scenario, setting up the wall is the least cost method. Welfare is 45, which is a maximum.

If, as in scenario 5 a, it is assumed that the cost of the wall is 20, the transport industry is the cheapest cost avoider, and it should cease operation.

Table 2-4 summarises these results. The highlighted rents indicate the cheapest cost avoider.

Table 2-4 Summary rents

Scenario / Rents	1	2	2a	3	3a	4	4a	5	5a
Transport Industry r^T	100	100	100	**25**	**25**	40	40	**30**	**30**
Residents r^R	100	**25**	**25**	100	100	**30**	**30**	40	40
Damage (with wall) d	20	20	20	20	30	20	20	20	20
Cost of wall w	**5**	**5**	10	**5**	5	**5**	25	**5**	20

Table 2-5 Relative rents

Scenarios	Relative rents
Scenario 1	$w < d < r^T = r^R$
Scenario 2	$w < d < r^R < r^T$
Scenario 2a	$w < d < r^R < r^T$
Scenario 3	$w < d < r^T < r^R$
Scenario 3a	$w < r^T < d < r^R$
Scenario 4	$w < d < r^R < r^T$
Scenario 4a	$d < w < r^R < r^T$
Scenario 5	$d < w < r^T < r^R$
Scenario 5a	$w = d < r^T < r^R$

2.4.2 Cheapest cost avoidance: care as a continuous variable

In the previous section, only two levels of care are possible, i.e. zero precaution or precaution through abandoning one of the damage causing activities, thereby sacrificing the whole rent flowing from it. In contrast, we now consider the case in which care is a continuous variable.

This means that care, measured as forgone rents or the cost of the measures taken to prevent damage, can take values from zero to the maximum rent. Thus, the higher the level of care, the lower the rent that flows from the activity. Furthermore, we assume that for both parties the higher the level of care, the lower the expected damage. Note that a similar analysis can be made for a situation in which damage has already occurred. Just substitute care (or avoidance) by abatement – the results would not be altered.

The residents as well as the transport industry can take care to reduce the expected costs of pollution. It makes sense to interpret the reduction of the damage as an output produced by two inputs, namely care by the transport industry and care by the residents. The question becomes which combination of inputs is optimal. The answer depends on the marginal productivity of the inputs to reduce expected damage, compared to the marginal cost of the input.

To minimise the total cost of damage and damage prevention requires that both the residents and the transport industry should increase spending on damage prevention until the last Euro they spend reduces damage by one Euro.

Figure 2-9 Efficient avoidance of an externality

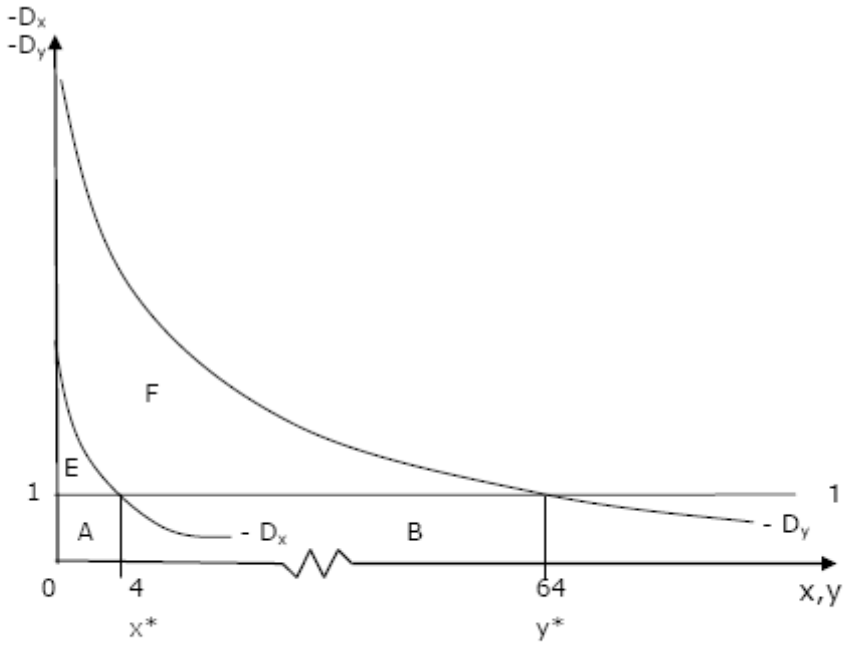


Figure 2-9 represents this idea. Let x and y be the costs (or sacrificed rents) of the transport industry and the residents, respectively, and let D be damage. The more the transport industry spends on damage prevention, the lower the expected damage, i.e. $D_x<0$. We assume that marginal productivity decreases, which means that additional amounts spent on damage prevention reduce expected damage at a decreasing rate. The same applies to the damage prevention activities of the residents. Under the assumptions made, the transport industry is not as productive in damage prevention as the residents.[92]

[92] Of course, in the case in which the figures for y belong to the transport industry and those of x to the residents, the opposite would hold.

The horizontal line represents the marginal costs of damage prevention. For simplicity, we assume that the marginal costs are constant and equal to one for both the transport industry and the residents. The areas below the downward sloping curves $-D_x$ and $-D_y$ measure the amount of damage reduced by investing respectively X and Y. Thus, these curves depict the marginal benefit of care. In equilibrium, parties should contribute to damage prevention to the point where marginal cost equals marginal benefit: x* for the transport industry, and y* for the residents. Both parties contribute to the prevention of the damage. The transport industry reduces damage by the sum of areas E + A. Since A measures the costs of damage reduction, E is the net benefit of the transport industry's investment in damage reduction. The residents reduce damage by the sum of areas A + B + E + F. Since A + B measures the costs of damage reduction, E + F is the net benefit of the residents' investment in damage reduction. The net benefit created by the investments of both the transport industry and the residents amounts to E + (E+F) = 2E + F, which represents a maximum increase in welfare.

In this case, efficiency requires that the residents should bear the greater burden; in other cases, it is possible that the opposite holds, or even that one of the two parties should not contribute at all. The latter is the case when the marginal costs of prevention for one party are higher than the marginal benefits for all levels of care.[93]

2.4.3 Putting the cheapest cost avoider principle into practice

In this section we develop a formal analysis of the efficiency of a number of methods for controlling externalities according to the cheapest cost avoider principle. After having identified the cheapest cost avoiders, the question arises how to create an incentive structure that induces optimal care by the parties in a Nash equilibrium.

In Nash equilibrium, each party[94] maximises its own objective function, taking the decision of all the other actors as given. Let x denote the cost of prevention measures to the transport industry and y to the residents. In equilibrium, given the other party's decision, no party has an incentive to change its decision. Thus, in a Nash equilibrium each party's behaviour is an optimal response (or "strategy" in game theory) to the behaviour of the other players.[95] We assume that the parties choose their precaution simultaneously. To yield efficiency, a Nash equilibrium requires $x^* \equiv x^*(y^*)$ and $y^* \equiv y^*(x^*)$.

[93] For an a simple formal treatment of the continuous case, see appendix.

[94] Also called "player" in the game theory literature.

[95] See Watson 2002, p 82-86.

We discuss the following four mechanisms with regard to their ability to induce optimal behaviour:

- The tax–subsidy approach
- The strict liability rule
- The no-liability rule
- The negligence rule

The tax–subsidy approach

According to this approach, the government imposes a tax on the polluter (or it pays him a subsidy) in order to make him internalise the externality and choose the optimal level of care. Let t_x be the marginal tax. The polluter is given optimal incentives when the marginal tax is equal to the marginal external damage, i.e. $t_x = D_x*$. The same outcome can be accomplished by paying a marginal subsidy denoted s' to the polluter. s' is equal to the negative of the marginal external damage, i.e. $s_x = -D_x^*$ at the optimal care level.[96]

Damage is D(x,y), where x continues to be the avoidance cost of the transport industry and y the avoidance cost of the residents. Now, the tax (or subsidy) faced by the transport industry is t(x,y). As long as the marginal tax is such that $t_x = D_x*$, the transport industry will choose the optimal level of precaution x^*, given the level of precaution chosen by the residents. The residents will also choose their optimal level of precaution y^*, provided that they do not receive compensation for their damages (or that the compensation is independent of their choice of y).[97] The reason why such a tax produces an equilibrium outcome (x^*, y^*) is that both the transport industry and the residents are fully liable at the margin.[98]

Consider fig. 2-9. With a marginal tax of $t_x = 1$, the transport industry is induced to choose $x^* = 4$. The residents still suffer the damage of area (A+E +F+B). Up to y = 64, the marginal gain from investing in damage reduction, i.e. $-D_y$, is greater than the marginal cost of 1. Thus, residents will choose $y = y^*$, which is efficient.

If the optimal solution is one in which only the residents should take action, i.e. $x^* = 0$, and $y^* > 0$, the tax should be zero $t_x = 0$. The residents have the incentive to ensure that the choice is $y = y^*$, which is efficient. If on the other hand $x^* > 0$, and $y^* = 0$ were optimal, the tax should be set at $t_x = D_x*$.

The Strict liability rule

Under a rule of strict liability, the transport industry is held liable for all the damage suffered by the residents. It is well known that the victim has no

[96] See Miceli 1997, p 119 and Polinsky 1979.

[97] See Miceli 1997, p 122.

[98] See Miceli 1997, p 122.

incentive to undertake any precaution when a polluter is liable for all the damage he generates. Thus, if optimality requires the victim, here the resident, to undertake precautions, the rule of strict liability is inefficient.

To prove this assertion, a little formalism is necessary. Let C^{xs} and C^{ys} respectively stand for the avoidance costs of the transport industry and the residents, respectively. Under strict liability, these costs are characterised as follows:[99]

$C^{xs} = y+D(x,y)-D(x, y) = y.$

The residents only pay for their own measures of damage prevention. Of course, they suffer damages, D(x,y), but they are fully compensated for them. The residents want to minimise their costs, which, implies $y^* = 0$.

The transport industry also wants to minimise its costs. The first order condition for a minimum is

$C_x^{xs} = 1+D_x(x,y) = 0,$

which is equivalent to

$- D_x(x,0) = 1.$

The transport industry chooses $x = x^*(0)$.

If the optimal amount of prevention by the residents is greater than zero, i.e. $y^* > 0$, then the Nash equilibrium under a rule of strict liability is inefficient. This result holds whether there is independence of care, i.e. $D_{xy} = 0$, or substitutability, i.e. $D_{xy} > 0$. In the case of substitutability, lowering y would result in an outward shift of the curve labelled $-D_x$ in fig. 2-9, thereby increasing x^*. If the transport industry invested x = 64 in precaution, the social cost would increase to 132.

As is well known from the law and economics literature, adding a defence of contributory negligence to the strict liability rule can achieve the optimum solution in equilibrium.[100] A defence of contributory negligence would allow the transport industry to avoid liability if the residents are negligent, i.e. if they choose $y < y^*$.

The simple rule of strict liability achieves efficiency in equilibrium if the optimal amount of damage prevention by the residents y is zero, i.e. $y^* = 0$. Damage from a nuclear power plant explosion or damage from an airplane crashing onto a house roughly approximate these situations; it is foreseeable that damage prevention by the residents is too costly to be efficient. In fig. 2-9, this implies an outwards shift of the $-D_x$ curve, resulting in $x > x^*$ (= 4).

[99] See Wittman 2006, p 137.

[100] See Miceli 1997, p 19.

The no-liability rule

Under this rule the residents are liable for their damages and the transport industry would face prevention costs. It is in effect a rule of strict liability for the residents. Obviously, the rule of no liability is the mirror situation to strict liability for the transport industry. The result is straightforward: the transport industry chooses $x = 0$ for all y. Therefore the residents' problem is to minimise $y + D(0,y)$, which yields $y^*(0)$.

If $x^* = 0$, then the equilibrium of the transport industry under no liability is efficient. However, if $x^* > 0$, then the rule of no liability yields an inefficient equilibrium. Both with independence of precaution, i.e. $D_{xy} = 0$, and with substitutability, i.e. $D_{xy} > 0$, the residents invest too much in damage prevention, i.e. $y^*(0) > y^*$, and the transport industry too little, i.e. $x < x^*$. In terms of fig. 2-9, substitutability would imply an outwards shift of the $-D_y$ curve, resulting in $y > y^*$ (= 64).

The negligence rule

We will restrict our attention to one sided negligence, which means that only the transport industry will be held liable, and this only happens if the precautions undertaken are less than a predefined due standard of optimal precaution.[101]

Assume that the due standard of care, denoted z, is set equal to the transport industry's optimal care level, i.e. $z = x^*$. It can easily be shown that (x^*,y^*) is a Nash equilibrium.[102]

Suppose that $y = y^*$. Then the transport industry can choose between two options:

$$x \geq z = x^* \text{ and}$$

$$x < z = x^*.$$

With $x < z$, the transport industry is held liable. Thus its cost of this option is $x + D(x, y^*)$. With $x \geq z$, the transport industry is not held liable. Its cost is x. Rationality would require it to set $x = x^*$. The optimal decision is $x^*(y^*) = x^*$, since

$x^* < x^* + D(x^*, y^*) \leq \min_{(x < z)} x + D(x,y^*)$.

Let us now consider the residents' problem and let $x = x^*$. If the transport industry chooses $x^* = x^*(y^*)$ it is free from liability; however, there are still damages for which the residents are strictly liable in the sense that they are responsible for the damages they suffer. Thus, the residents choose y to minimise $y + D(x^*,y)$, which yields $y^*(x^*) = y^*$ and which is efficient. To illustrate, consider fig. 2.9. With $z = x^*$, the transport industry chooses $x^* = 4$. It is optimal for the residents to choose $y^* = 64$.

[101] For a general discussion of negligence rules, see Wittman 2006, p 142-151 and Miceli 1997, p 18-20, 123.

[102] See Miceli 1997, p 18-19.

Summing up, we can say that the tax subsidy approach induces both the transport industry and the residents to choose the optimal actions, i.e. which induce cheapest cost avoidance, to control externalities, since both are fully liable for the damages at the margin. A simple rule of strict liability is only efficient in the special case where the residents should not undertake any level of precaution. This setting corresponds to making the polluter pay. However, the reasons are different from those in the polluter pays principle: the polluter is made to pay not because he causes the harm, but because he can reduce it most efficiently. The polluter pays outcome is thus a special outcome of the cheapest cost avoider principle. The existence of other possible outcomes shows that it should not be made a principle.

A rule of no liability is only efficient in the corner case where the transport industry should not undertake any precaution. If however both the transport industry and the residents should undertake precautionary measures, the simple rules of strict liability or no liability are in general inefficient, i.e. they do not implement cheapest cost avoidance. However, negligence rules encourage the optimal levels of precaution if the due standard of care is chosen appropriately.

At which level should the cheapest cost avoider principle be applied?

The cheapest cost avoider principle is applied in court rulings; if not explicitly, then implicitly.[103] In part III, we will give two examples where courts balance the interests of the involved parties.[104]

However, even in the presence of a properly working judiciary, there are a number of arguments which speak in favour of centralisation, i.e. regulation. As Esty (2001) puts it: "(C)ommand and control regulations, based on government defined pollution control technologies or emissions standards, represent a way to overcome individual analytic incapacity, achieve scale economies in the technical dimensions of environmental protection, and fill by fiat the information gaps (and related doubts about causal connections)"[105]. Also, "by mandating pollution control rules and requirements, regulatory strategies limit the institutional obstacles to recovery that individuals of harms face and the risk that strategic behaviour in negotiations (by either polluters or victims) will derail agreement on compensation."[106] Esty (2001) concludes that "in at least some circumstances, command and control regulation offers a promise of lower technical, political/legal, and strategic costs yielding outcomes that

[103] See the Law and Economics literature, and especially Posner 1986. Even if court judgements do not use the language of economics, but legal terms and concepts, judicial reasoning is always concerned with the balancing of antagonistic interests.

[104] For a general discussion of the role of economics in courts see Ogus 2006, p 299-310; see also Breyer 1993 and the Symposium "Economists on the Bench" 1987.

[105] Esty 2001, p 13. See also Ogus, who discusses similar issues under the heading "From private governance to (modifiable?) public governance" (Ogus 2006, p 79-83).

[106] Esty 2001, p 13.

internalise externalities, and protect property rights at lower social cost than a tort and contract based regime."[107]

Indeed, insights of the economic analysis of law, which is largely efficiency orientated, are increasingly used in policy making and legal reforms. In part inspired by the American model that requires a fully fledged cost-benefit analysis to be undertaken for major regulatory proposals[108], government departments in Europe as well as the European Commission, when making proposals, have to prepare what has become known as a "regulatory impact analysis", a "regulatory impact appraisal", or a "regulatory impact assessment", which includes some form of cost-benefit analysis.[109] The United Kingdom is a good example. To quote Ogus: "The current UK regime requires a regulatory impact assessment (RIA) to be undertaken by government departments making any proposal for regulation which has an impact on business, charities or the voluntary sector. The document must include an identification of, *inter alia*, the risks that the proposal is addressing and an attempt to quantify them; the different regulatory options for meeting them; the benefits and costs of these options and the sectors which will bear them; equity and fairness issues; and any distributional impacts of the proposals. It should conclude with a recommendation regarding the preferred option, giving reasons based on the elements of the assessment, in particular the analysis of the benefit and costs."[110]

In case C-320/03, discussed in part III, Commission of the European Community vs. Republic Austria, concerned with the sectoral prohibition of the movements of lorries of more than 7.5 tonnes carrying certain goods, one can read in the summary judgement: "Such a prohibition obstructs the free movement of goods and, in particular, their free transit (…). Such legislation cannot be justified by imperative requirements in the interests of environmental protection where it has not been demonstrated that the aim pursued could not be achieved by other means less restrictive of freedom of movement." In the vein of a cost-effectiveness assessment, the court states in §87: "(…) the Austrian authorities were under a duty to examine carefully the possibility of using measures less restrictive of freedom of movement, and discount them only if their inadequacy, in relation to the objective pursued, was clearly established."

2.5 Comparative Institutions: the polluter pays principle

In order to provide the relevant information needed for a comparison of the polluter pays principle and the cheapest cost avoider principle, this section discusses their strengths and weaknesses. One advantage of the cheapest

[107] Esty 2001, p 15. See also Shavell 1987, p 277-290, who suggests several advantages that pertain to a statute-based system compared to a tort system, and Rose-Ackerman 1992 expanding on this theme.

[108] See Ogus 2006, p 282.

[109] See Ogus 2006, p 279-292, European Commission 2001, 2005.

[110] Ogus 2006, p 283.

cost avoider principle relative to the polluter pays principle, derived in the previous sections, is that the polluter pays principle cannot guarantee efficiency, whereas the cheapest cost avoider principle does. This result cannot be overstated since efficiency, in the static as well as in a dynamic sense implies the avoidance of a waste of valuable resources, an increase in the value of production, positive growth rates and an increase in average income. Efficient transport furthers the international division of labour, improves the functioning of the internal market and is a necessary precondition for reaching the "growth and jobs" goals of the Lisbon strategy. There can be no doubt that an efficient transport industry furthers the common good (i.e. social welfare) and, analogously, that it is necessary to maximise the wealth of a society. Wealth can be used as a proxy for social welfare.[111] Thus, the cheapest cost avoider principle clearly beats the polluter pays principle as far as efficiency is concerned. However, one can ask whether efficiency is all that matters. What about fairness and the costs of administering and implementing both principles?

For example, there may exist an efficiency-equity tradeoff. The question is how to deal with it. Wealth maximisation is not the only goal of state activities: Moral values such as human autonomy and dignity and distributional justice are of great importance. In many areas policy makers need to practice a tradeoff between "economic" and "non-economic" goals. How should such tradeoffs be carried out?

In this section we address possible arguments for and against the implementation of the polluter pays principle as well as the cheapest cost avoider principle under "non economic" aspects. We will study whether the polluter pays principle contributes to the notion of fairness in the sense of corrective justice, in the sense of distributive justice (equity), and the sense of a level playing field in the transport sector. In addition, we will deal with the question whether the costs of administering the polluter pays principle relative to those of the handling of the cheapest cost avoider principle could make a case for a preference of polluter pays principle over cheapest cost avoider principle.

Regarding the cheapest cost avoider principle, we will take up the fairness and administering costs issues as well as comments on the breadth and soundness of its underpinning concept of welfare economics in section 2.6.

2.5.1 The polluter pays principle and corrective justice

In this subsection we will deal with three questions:

- What does corrective justice mean?

[111] Wealth stands for the sum of all goods and services in a society weighted by their values. The term "value" refers to value in exchange, value as measured or at least measurable in a market, whether explicit or implicit (see Posner 1981, p 60). This concept of value is broader than of a price in the sense that it takes all valuations into account, and not only the marginal value in equlibrium. For the relation between wealth and social welfare, see section 2.6.4 below.

- Why is corrective justice so appealing?
- Why the polluter pays principle cannot be founded on corrective justice.

What does corrective justice mean?

The classical principle of corrective justice requires a person who wrongfully inflicts harm on others to pay for the harm that he has caused.[112] Paying for the harm can mean one of two things. Either the injurer pays and the payment goes to the victim (the victim is fully compensated), or the injurer pays and the payment goes to a third party, say the government. Whereas in the second interpretation the notion of fairness is one involving punishment motivated by the idea of retribution, the first one is driven by the desire to rectify the outcome created by the injurer's action[113]. The first mentioned position, according to which a person who wrongfully harms another should compensate the victim, is usually advanced under the heading of corrective justice.[114] Here we take a broader view on the term and subsume both positions in which the injurer has to make amends, whether to society or to the victim, under the term "corrective justice".

Note that it is the physical interference or invasion into another person's protected domain (or property, widely defined) which is the significant factor to study. Thus, the notion of causation underlying the principle of corrective justice is that of the unilateral infliction of harm; there is only one "causer" of harm who is identified in reference to a purely physical relationship.

Note further that the notion of corrective justice does not require ascribing responsibility to those who inflict harm on others in general, but rather it assigns responsibility only if a person *wrongfully* harms another.

Finally, note that a fair treatment of individuals based on a notion of corrective justice only depends on the situational character of an event, and does not depend on how the treatment influences individuals' behaviour and, in turn, on how such behaviour affects individuals' well-being.[115] Only two questions have to be asked: Is a person's conduct wrongful, and does it cause harm – if this is so, then the principle of corrective justice requires that person to pay compensation to the victim or to pay a tax. This is not to deny that, in reality, a fair treatment in the sense of penalising wrongful, harmful conduct will influence behaviour. But as Kaplow and Shavell rightly mention:

"(A)ssessing such effects is not part of an analysis based on the notion of fairness (…), because such normative analysis is avowedly independent of how the pursuit of fairness will influence the well-being of individuals." [116]

---

[112] See Kaplow and Shavell 2002, p 41, 87.

[113] See Kaplow and Shavell 2002, p 87.

[114] See Kaplow and Shavell 2002, p 89.

[115] See Kaplow and Shavell 2002, p 41.

[116] See Kaplow and Shavell 2002, p 43.

It follows from this characterisation of fairness that the polluter pays principle as a fairness based normative analysis is potentially in conflict with the cheapest cost avoider principle. The cheapest cost avoider principle is concerned exclusively with effects on people's well-being, whereas notions of corrective fairness underpinning the polluter pays principle are not at all concerned with such effects.[117]

At the first glance, the polluter pays principle seems to fit the notion of corrective justice nicely. The generators of pollution, the polluters, should pay for the costs inflicted on other persons by their activities – the residents in the examples used in the previous sections. Thus, striving for corrective justice seems to require an adherence to the polluter pays principle rather than to the cheapest cost avoider principle.

In what follows, we will first deal with the question of why corrective justice as a goal of public policy is so appealing. Consequently, we will show that corrective justice is a flawed concept, and that adherents of the polluter pays principle cannot hope to use it as an argument to defend the polluter pays principle against the cheapest cost avoider principle.

Why is corrective justice so appealing?

There are several answers to the question why corrective justice is appealing. We offer some explanations for the apparent fairness based analyses.

Cognitive biases: Notions of corrective justice are typically used "to reach conclusions based upon situational characteristics of events". The assessment is frequently made from an ex post perspective and often ignores important aspects of ex ante behaviour: "This tendency to focus on what is salient … is related to familiar and prevalent cognitive biases", such as the "hindsight bias" or the "tendency of subjects to evaluate the probability of events on the information that is most available or salient".[118] Thus, "insights from cognitive psychology seem to offer at least a partial explanation for the apparent attractiveness of fairness based analysis".[119]

Fairness proponents often appeal to intuitions or instincts. If the intuitions and instincts are grounded on the promotion of individuals' well-being, reliance on these sources seems to be adequate. But if intuitions and instincts can lead us astray, reliance on these sources of notions of fairness is self defeating.[120] It is an important purpose of explicit normative analysis of public policy to identify and avoid those situations.

Social norms: Notions of corrective justice seem to be closely related to various social norms that guide ordinary individuals in their everyday lives,

---

[117] See also Kaplow and Shavell 2002, p 42, who compare a fairness based normative analysis with welfare economics.

[118] See Kaplow and Shavell 2002, p 50, n. 69.

[119] See Kaplow and Shavell 2002, p 50.

[120] Kaplow and Shavell 2002, p 60.

such as not harming others, and being held responsible when one does so. Social norms can be inculcated or are the result of evolution in the biological sense. They play an important role in channelling individuals' behaviour in a socially desirable manner[121], thereby reducing the costs of private and public enforcement of the law.

A second reason why social norms tend to be valuable regulators of everyday behaviour is that "they may serve as useful proxy principles, heuristics, or rules of thumb that promote individuals' welfare".[122] In this case, notions of corrective justice are not invoked as evaluative principles in their own right[123] but as a proxy principle to help identify legal rules that increase social welfare. Accepting norms or following rules is valuable since it saves decision making costs and reduces the probability of costly errors[124]. There are however two preconditions: norms must be followed quasi automatically, and norms should be simple and general in application.[125]

Why the polluter pays principle cannot be founded on corrective justice

We propose three reasons why notions of corrective justice cannot support the polluter pays principle.[126]

1) Corrective justice is based on an incomplete notion of causation. The implication is that there is only one causer of damage (the polluter). But, as we have seen in previous sections, all damage is jointly caused. Damage is not the result of the unilateral invasion into somebody's protected domain, but instead arises from the friction which occurs when two or more actors compete for the same scarce resource.

   We know since Coase (1960) that the position of polluters and pollutees is symmetrical and reciprocal: "each wishes to use the environment in ways which are utility maximising. There is no a priori reason for assuming that, because the polluter's activity involves a physical interference, her claim on the environment is less valuable than that of the pollutees. The normative economic proposition which flows from this is not necessarily that the interference be abated or paid for by the

[121] Kaplow and Shavell 2002, p 62-65.

[122] Kaplow and Shavell 2002, p 88.

[123] Kaplow and Shavell 2002, p 44.

[124] See Heiner 1983.

[125] Kaplow and Shavell 2002, p 69.

[126] Kaplow and Shavell nicely describe why it should not be applied even if it were possible: "As a normative matter, however, if the appeal of notions of fairness (…) derives from what amounts to mistakes in judgment, there is no basis for giving the notions weight as independent evaluative principles, to be pursued at the expense of individuals' well-being". Since an effects based welfare economic approach to policy assessment reflects a complete consideration of factors that plausibly seem relevant, while the approach based on notions of fairness does not, the former approach would seem superior to the latter, a priori. Consequently, "furthering notions of fairness, whenever they favor policies different from those endorsed under welfare economics, leads to reductions in individuals' well-being". Kaplow and Shavell 2002, p 48-59.

inflictor, but rather that the friction between the conflicting resource uses be relieved at lowest cost, taking account both of the value of the resource uses and the costs of adapting behaviour".[127] It follows that the costly interaction between the transport industry and the residents cannot be attributed to the actions of either party individually; instead, it is "caused" by resource scarcity.

Consequently, as Demsetz reminds us, the use of words such as "blame", "responsible", and "fault" must be treated with care "because they have no useful meanings in an economic analysis of these problems other than as synonyms for the party who could have most easily avoided the costly interaction. Whether the interaction involves crop damage, accidents, soot, or water pollution, the qualitative relationship between the interacting parties is symmetrical. It is the *joint* use of a resource, be it geographic location, air, or water that leads to these interactions. It is the demand for scarce resources that leads to conflicting interests".[128]

And most importantly, referring to court decisions (which can be generalised to policy actions), Demsetz points out: "If courts are to ignore wealth, religion, or family in deciding such conflicts, if persons before the courts are treated with regard only to the cause of action and available proof, then, as a normative proposition, it is difficult to suggest any criterion for deciding liability other than placing it on the party able to avoid the costly interaction most easily".[129] Notions of corrective justice are insufficient to address the complexity of the issues and interests at stake.[130]

2) Wrongful harm cannot be determined by notions of corrective justice

Instinctive notions of corrective justice and naive causation analysis are insufficient to address the complexity of the issues and interests at stake[131]. For example, nearly all legal systems, whether in civil or in common law countries, contain rules determining when individuals should be held liable for negligent omissions, that is, failures to act, rather than positive acts. In particular cases, wrongdoers get off scot-free if the other party is proven to have been negligent.[132]

The Coasean view requires an expansion of the parameters to be taken into account. As Bruce Ackerman puts it: "Instead of sifting the facts in

[127] Ogus 2006, p 167.

[128] Demsetz 1972, p 28.

[129] Demsetz 1972, p 28.

[130]See Ogus 2006, p 81.

[131] See Ogus 2006, p 8.

[132] "The rule that contributory negligence is a complete defense to a negligence suit, the rule that there is no right of contribution among joint tortfeasors, and the substitution of heavy criminal penalties for lighter penalties imposed with a probability of one illustrate the common law's apparent willingness to allow wrongs to go uncorrected in many cases", Posner 1986, p 243.

search of *the* cause of the trouble, the lawyer-economist urges a conception of causation that recognises how a multiplicity of factors, operating over a lengthy period of time, contribute to our legal discontents (…) Rather than beginning with the moment at which the actors get into some form of obvious trouble, Coasean assumptions force the lawyer to start his story at a much earlier point in time: when the parties could have reorganised their activities in a way that could have avoided the trouble entirely"[133]. As a matter of fact, all known legal orders approach problems of causation and damage in a more sophisticated way than suggested by notions of corrective justice.[134]

It is worth noting that the fact that wrongs go uncorrected is misleading, since it results from a failure to distinguish ex ante (before the occurrence) from ex post (after the occurrence). Of course, ex post a tortfeasor who injures a contributory negligent victim gets off scot-free. But ex ante a contributory negligence system generates correct incentives to potential injurers as well as potential victims[135]. Of course, one could make sure that every wrong is corrected. But the question is whether corrective justice is a commodity for which society should be willing to pay an infinite price[136].

3) Corrective justice is typically a procedural rule rather than a complete substantive notion of justice[137]. The principle prescribes rectification of wrongful acts that cause injury, "regardless of the relative merit of injurer and victim considered apart form the act, but it does not define what acts are wrongful; this definition is not itself part of the concept of corrective justice."[138] Thus, corrective justice is incomplete in the sense that in order to make it operational, "one must look elsewhere for a substantive theory of what counts as wrongful injury"[139]. For example, following libertarian approaches, one could stipulate as wrongful all harm caused by a person's voluntary actions. Alternatively, one could deem all those acts as wrongful that are inefficient or that reduce social welfare[140]. Thus, the most important message is: since nearly any substantial principle could be embedded in corrective justice, this principle cannot be used to challenge the claim that policy analysis should be based on the cheapest cost avoider principle. Of course, the polluter pays principle can also be embedded in corrective justice.

---

[133] Ackerman 1984, p 52-53; as quoted in Ogus 2006, p 9.

[134] See for example ch. 6 in Ogus 2006, in which the difficulties of articulating principles of corrective justice in relation to omissions, i.e. failure to act, coming to the nuisance, the requirement of damage and different types of damages are outlined.

[135] See Posner 1986, p 243.

[136] See Posner 1986, p 244.

[137] See Posner 1981, p 73-74, Kaplow and Shavell 2002, p 93-95.

[138] Posner 1981, p 74.

[139] Kaplow and Shavell 2002, p 93.

[140] See Kaplow and Shavell 2002, p 95.

However, the polluter pays principle can provide the foundation for, but it cannot itself be founded on, the principle of corrective justice.

2.5.2 The polluter pays principle and distributive justice (fairness)

Typically, societies have distributional goals which they rate as least as high as efficiency. Thus, from the point of view of maximisation of social welfare, efficiency and distribution of wealth or income should be considered together for two reasons: firstly, a given distributional goal may involve a sacrifice of efficiency, or, secondly and conversely, the attainment of efficiency may imply an unacceptable distribution of wealth or income. Since the polluter pays principle affects the distribution of wealth or income when pollutees and polluters are legally distinct groups with different levels of wealth or income, it might appear that its deficits on the efficiency side can be outweighed by its contribution to distributive justice.

To illustrate this point, suppose that firms in the transport industry are owned by very rich persons and that the residents are very poor. Assume further that the residents are the cheapest cost avoiders and that transaction costs are high. Then by letting the polluters pay, the external costs will be borne by the transport industry and this may increase social welfare, even if efficiency is violated. Thus, a societal preference for a more equal distribution of wealth or income would suggest trading efficiency for justice.

But what about the case in which residents are very rich and the owners of the firms in the transport industry are poor? Care is required here because of what economists call a partial equilibrium analysis: the focus is only on the polluting activity and its direct consequences[141]. But there may be consequences for other parts of the economy. As Stephen reminds us, regarding a context similar to ours but assuming that a court grants an injunction in favour of the residents: "What if the residents live in rental houses and the removal of the pollution raises the value of property and allows the rich landlord to raise rents and acquire all the benefits of the court's decision? The higher rents may not even be paid by the original residents who perhaps cannot afford them. They may therefore move out and be replaced by other, richer tenants whose interests might not have weighed so heavily against those of the factory owner in the original decision."[142]

Distributional consequences are likely to be more important in cases where the output of the transport industry will be affected by paying for damage: "Reducing the output (perhaps to zero) will affect levels of employment and incomes of (possibly) relatively poor workers … and the pollution conscious residents may in fact be relatively wealthy."[143]

[141] For this argument see Stephen 1988, p 62.

[142] Stephen 1988, p 63.

[143] Stephen 1988, p 63.

In addition, given the importance of the transport industry for the national and international division of labour and for the functioning of the internal market, we can reasonably expect that there are spillovers to other parts of the economy, resulting in patterns of redistribution which are not in accordance with the distributional goals of a society.

The upshot of this argument is that distributional effects are relevant under the aspect of social welfare maximisation. However, it is a generally held view among economists that distributional issues may often best be addressed directly, through tax and transfer schemes.[144]

2.5.3 The polluter pays principle and fairness between transport modes

According to the CE study, the internalisation of external costs is an important precondition to "guarantee fairness between transport modes, that means fair prices considering the overall performance and potentials of the different transport modes"[145]. Fairness in this context means to provide a level playing field for all modes of transport[146] and implies the elimination of distortions of competition between transport undertakings. At the first glance, the goal of levelling the playing field for modes of transport seems to strongly support the implementation of the polluter pays principle. However, closer scrutiny raises doubts. Three arguments are pertinent here:

- Why can we not also have a level playing field between modes of transport without any internalisation? Each mode would have to compete on the basis of its direct operating costs. Of course, whether this would encourage sustainable transport – an important goal of directive 2006/38/EC – is another question.
- If sustainable transport in the EC contributes to maximal welfare, we may well wonder whether an internalisation policy concerning all modes of transport is adequate, given that maximal welfare should be the ultimate goal of public policy. Since the modes of transport are heterogeneous in terms of the type and scale of external costs, the transport routes taken and the abatement and avoidance technologies, it is by no means clear that it is sustainable that each mode of transport should internalise its external costs.

Consider figure 2-9. Remember that x and y represent the abatement activities of the transport industry and the residents, respectively. For some cases, the situation as depicted in figure 2-9 might be relevant; in other cases, x should refer to the residents and y to the transport industry. There might also be situations in which it is efficient to abstain from an internalisation at all, since the residents are the cheapest cost avoiders. Which scenario is relevant is an empirical question.

[144] See Kaplow and Shavell 2002 p 33-34, 86, 460.

[145] CE study, p 1.

[146] See CE study, p 36.

Thus, we would agree with the quoted statement of the CE study that "the overall performance and potentials of the different transport modes" must be taken into account when considering the issue of fairness between transport modes from the point of view of maximal welfare. However we do not accept that the internalisation of external costs is always an important precondition for fairness. In a sense, this conclusion is in accordance with the basic message of second best theory that deviations from marginal social cost pricing may be appropriate from a social welfare point of view.[147]

2.5.4 The polluter pays principle and administrative costs

All policies addressing the problem of external costs affect peoples' well-being not only to the extent that they influence their incentives and the allocation of risks, but also on account of administrative costs. These costs include the setup and operating costs incurred by both the private and the public sector (legislation, administration, courts). The administrative cost issues should be addressed using a comparative institutions approach. In reality, all institutions – here broadly understood as rules and norms – are imperfect in the sense that they do not operate without costs: opportunity costs in terms of a misallocation of resources and risks, setup and operating costs. Consequently, rationality requires taking all costs into account when making an institutional choice. One could be inclined to make a case for the polluter pays principle out of this argument.

We know that the cheapest cost avoider principle is superior to the polluter pays principle as far as efficiency is concerned. But if the setup and operating costs of the polluter pays principle were much lower than those resulting from the cheapest cost avoider principle, such that the cost differential stemming from the efficiency side is outweighed, society should prefer the polluter pays principle from a welfare point of view.

Indeed, at first glance the polluter pays principle seems to beat the cheapest cost avoider principle as far as the administrative costs are concerned. Whereas the cheapest cost avoider principle requires some form of a cost-benefit analysis to be undertaken in order to identify the cheapest cost avoider, the polluter pays principle simply requires information about the polluter.

But having identified this polluter, the question arises how to determine what he should pay. In order to solve this problem, policy makers need information on the external costs and on the abatement (or avoidance) costs. But if policy makers have all this information, they know all that is necessary to identify the cheapest cost avoider. If they do not, they can

[147] It is well known from the economic literature that the internalisation of external costs can reduce welfare if firms have market power. In this case, output is too low from a welfare point of view; an internalisation policy would lead to a further reduction of output, which is – over all – detrimental to social welfare. See Perloff 2004, p 635-637; Just et al. 2004, p 537-538.

apply the polluter pays principle; however, whether their policy is efficient depends on luck and not on economic reasoning.[148]

However, if the taxes and charges or the non market instruments of regulation are implemented without this knowledge, the policy variables will be set wrongly, resulting in welfare losses. It is an empirical question whether the savings on the information cost side outweigh the costs of wrong policy decisions in terms of welfare losses due to the misallocation of resources. We will return to this issue when discussing criticism raised against the cheapest cost avoider principle in the next section.

Before moving to this discussion, a final argument weakening the information problems mentioned above needs to be studied: We do have some information concerning the amount and the sources of external costs (environmental costs, costs of climate change etc.). We do not necessarily need information about the abatement (or avoidance) costs. Instead, we can rely on a mechanism, the operating of which creates incentives such that the knowledge possessed by the polluters regarding abatement (or avoidance) costs is used efficiently in a decentralised manner. Such a mechanism exists in the form of trading emission rights. Economists and policy makers are in favour of this mechanism, since it can be expected that a given goal of emissions can be reached with least costs. In fact, this mechanism appears to be based on cheapest cost avoider principle reasoning: the market ensures that the marginal abatement costs are equalised among all participants in the market. The given goal of emissions will be reached efficiently, i.e. with the least costs overall.

This result is well in line with what is called cost-effectiveness analysis. However, there are two problems associated with such an analysis: the first one is to determine the maximum of emissions to be allowed. There is a danger that this maximum does not coincide with the optimum. The optimum is defined by the equimarginal principle: marginal abatement cost equal marginal external cost. If the maximum is higher than the optimum, welfare losses occur due to a too high scale of activities generating emissions. If it is too low, then welfare losses result due to a too low scale of activities. Thus, cost-effectiveness analysis does not guarantee maximum welfare. However, it might be helpful, given a status quo, to apply an incremental policy asking, in the status quo, what are the incremental costs and benefits of a policy trying to restrict pollution.

The second problem with a cost-effectiveness analysis is to determine who should participate in this mechanism. Basically, polluters from all sectors of the economy should be included. Because the externalities concerned are of a world wide order, the list of the participants in the market has to be adapted to this frame. The problem is to define the relevant market in terms of location and the types of the participants. As far as pollutees can be considered as the cheapest cost avoiders, they should also be included.

[148] Only if per chance the pollutee was already at an abatement optimum would a policy based only on information on the polluter be efficient.

2.6 Comparative Institutions: concerns about the application of the cheapest cost avoider principle

In this section, we discuss a number of issues that appear to pose difficulties for the application of the cheapest cost avoider principle, ranging from problems of evaluation, omission of "soft" variables (intangibles) and fairness considerations to the problems of indeterminacy, efficiency-equity tradeoffs and the problems involved in predicting individuals' behaviour. We will also deal with concerns that the value of the protection of nature as such seems to be forgotten, and take up the questions whether welfare maximisation is the final target of human activities, and whether one can compare the "right to pollute" to the "right to a clean environment". In line with Kaplow and Shavell's defence of the welfare economics framework[149], we will argue that many of these issues present practical challenges to public policy analysis, but that these challenges are met by the cheapest cost avoider principle.

To see the difficulties which we examine below as a criticism of the cheapest cost avoider principle involves a misunderstanding. Making other criteria, for example the polluter pays principle, the basis for policy assessment turns out to be easier than a cost-benefit analysis only if these alternatives *ignore* complexities of the real world concerning the effects of public policy on individuals' well-being discussed below.[150]

2.6.1 Valuing non pecuniary factors such as life, pain and suffering

One might object that the application of the cheapest cost avoider principle requires placing a Euro value on life, pain, suffering and other non pecuniary factors that do not seem readily convertible into such a common denominator. Indeed, economists quantify everything that is of value to us: money, of course, but also friendship, shade, nature, silence etc. Money is simply used as a unified measure in order to compare gains and losses. The translation into money values is not arbitrary: it is possible to find out what something is worth to a person via his willingness to pay or to accept. People pay money in order to live in peace and quiet or in order to obtain a seat in the shade. This indicates the value of these benefits. People also give up certain privileges in exchange for money: for a sufficiently high salary, they are willing to move from their neighbourhood, they are even willing to risk their lives (this is the reason why security personnel in Iraq is paid exorbitant salaries). There is nothing that cannot be translated into monetary terms by this method.

Far from assuming that "money is everything", economists recognise the value of non monetary benefits and losses. They use money simply as a scale of comparison.

[149] See Kaplow and Shavell 2002, p 449-450.

[150] See Kaplow and Shavell 2002, p 404.

Critiques of this practice of finding a common denominator fail to realise that it is a prerequisite to a coherent policy assessment: "As a matter of logic, it has long been understood that, if any tradeoffs are to be made among factors of concern, implicit prices can be used to signify the tradeoffs".[151] Kaplow and Shavell explain the tradeoff logic in this context this way: "We begin by observing that any complete theory that can be used for the evaluation of policy must be able to tell us whether it is desirable to spend $1,000, $1,000,000, $1,000,000,000, or half the GDP to do so. Furthermore, we presume that any plausible theory will answer affirmatively if the cost is sufficiently low (only $1) and negatively if the cost is sufficiently high (half the GDP). (…) Under any such theory, therefore, all else being equal, there will be some point – some dollar cost – below which it is deemed appropriate to make the expenditure and above which it is deemed inappropriate to do so. The concept of a monetary value of life in the present context simply refers to that point, whatever it happens to be.

Moreover, if we do care about saving lives and reducing pain, there is a virtue in formal policy analysis being explicit about the valuations to be used. If tradeoffs are to be made consistently – which is necessary if one wishes, for example, to save more rather than fewer lives – one must know what those tradeoffs are. The familiar example is that some government regulations save a statistical life at costs of a hundred thousand dollars and others at costs in the billions; if the regulations were rationalised to use a consistent implicit valuation and if they continued to require the expenditure of the same amount of resources, *many* more lives would be saved. The reason is that reallocating expenditures from places where they have a very low payoff in terms of saving lives to places where they have a very high payoff will greatly increase the number of lives saved."[152]

The widely held belief that economists use a cold hearted approach or that they adopt a narrow view of what truly matters, and that the "antiseptic" language of policy analysis obscures what is really at stake[153] is incorrect. As Viscusi (1992)[154] indicates, using a common denominator does not deny that life, pain and suffering are what really matters. The numbers assigned to costs and benefits are merely analytical constructs in order to represent the choice that individuals would make. Despite the public discussions about the priceless nature of life, policy makers cannot help but accept the basic point that logically consistent choice involves choosing between alternatives that, accordingly, have to be compared to each other.[155]

[151] Kaplow and Shavell 2002, p 450.

[152] Kaplow and Shavell 2002, p 450-452.

[153] See the literature discussed in Kaplow and Shavell 2002, p 452 n. 116.

[154] Viscusi 2002, chapter 2.

[155] Note that even the Stern report applies this method.

2.6.2 "Soft" variables

The objection here is that economists tend to speak about tradeoffs in quantitative terms, and ignore factors that are difficult to quantify.[156] Even if this claim has some truth, it should be seen as an argument in favour of undertaking a more comprehensive analysis, rather than for substituting an inferior principle such as the polluter pays principle, which does not even in principle pay attention to the correct variables.[157]

Note the somewhat ironic fact that the objections to the omission of "soft" variables stand in tension to the previously mentioned criticism of economists' attitudes to quantify factors that seem least amenable to measurement.

2.6.3 Empirical data concerns

Policy conclusions derived from an application of the cheapest cost avoider principle depend on empirical data. Even if the analyst makes coherent judgements, when the available information is insufficient, the judgements are bound to be wrong. That is true, but does not necessarily require substituting the polluter pays principle for the cheapest cost avoider principle, since the polluter pays principle is – as we have seen – affected by conceptual shortcomings. Substituting the polluter pays principle for the cheapest cost avoider principle would mean to substitute a systematic error for a practical challenge in public policy.

There remains the question of how to use the cheapest cost avoider principle in those cases in which it yields uncertain conclusions. There are two answers to this question, one addressing the short run and one the long run.

When, in the short run, decisions must be based on incomplete empirical information, policy makers should realise that even a proper analysis yields tentative conclusions in which they may have little confidence and take into account that a subsequent revision of policies may be required.

In the long run, a proper use of the cheapest cost avoider principle is helpful to the formulation of scholarly agendas and government policy supporting this research.

2.6.4 Fairness issues

Our analysis proposes the pursuit of efficiency rather than focussing on equity. Indeed, there need not be a tradeoff between efficiency and equity. To illustrate: "Given a particular social welfare function, *society might prefer an inefficient allocation to an efficient one*. We can show this result by comparing two allocations. In Allocation *a*, you have everything and everyone else has nothing. This allocation is Pareto efficient: We can't make others better off without harming you. In Allocation *b*, everyone has an

[156] See Kaplow and Shavell 2002, p 454.

[157] See Kaplow and Shavell 2002, p 454, who discuss this point in relation to welfare economics in general.

equal amount of goods. Allocation b is not Pareto efficient: I would be willing to trade all my zucchini for just about anything else. Despite Allocation *b*'s inefficiency, most people would probable prefer *b* to *a*.

Although society might prefer an inefficient Allocation *b* to an efficient Allocation *a*, according to most social welfare functions, society would prefer some efficient allocation to *b*. Suppose that allocation *c* is the competitive equilibrium that would be obtained if people were allowed to trade starting from Endowment *b*, in which everyone has an equal share of all goods. By the utilitarian social welfare functions, Allocation *b* might be socially preferred to Allocation *a*, but Allocation c is certainly socially preferred to *b*. After all, if everyone is as well off or better off in Allocation *c* than in *b*, *c* must be better than *b* regardless of weights on individuals' utilities. According to the utilitarian rule, however, *b* is preferred to c because only strict equality matters. Thus by most of the well known social welfare functions, but not all, *there is an efficient allocation that is socially preferred to an inefficient allocation*."[158] Note that the European Union's institutions also consider efficiency, and not fairness, as one of its main goals.[159]

The cheapest cost avoider principle focuses on efficiency or wealth maximisation and interprets these measures as proxies for the maximisation of social welfare. Indeed, fairness issues, such as corrective justice, distributive justice or the level playing field for competitors do not play any role. The neglect of notions of corrective justice cannot be judged a serious deficit, since these concepts are not well defined. Distributive concerns are more effectively addressed directly, through the income tax and transfer system. Justice in the sense of having a level playing field will be realised as a by-product of the application of the cheapest cost avoider principle, since it makes sure that all competitors in a market are confronted with that level of abatement (or avoidance) costs which is optimal from society's point of view.

In his book "The Economics of Justice", Richard Posner develops the idea of wealth maximisation as an ethical concept.[160] He shows that "the wealth maximisation principle encourages and rewards the traditional 'Calvinist' or 'Protestant' virtues and capacities associated with economic progress" and points out that "wealth maximisation is a more defensible moral principle also in that it provides a firmer foundation for a theory of distributive and corrective justice".[161]

---

[158] Perloff 2004, p 345-346, emphasis added.

[159] Conference of the representatives of the governments of the Member States, Draft Treaty amending the Treaty on European Union and the Treaty establishing the European Community, 5 October 2004, CIG 1/1/07 REV 1, "Article 9 : 1. L'Union dispose d'un cadre institutionnel visant à promouvoir ses valeurs, poursuivre ses objectifs, servir ses intérêts, ceux de ses citoyens, et ceux des États membres, ainsi qu'à assurer la cohérence, *l'efficacité* et la continuité de ses politiques et de ses actions." Emphasis added.

[160] See Posner 1981, p 60-87.

[161] Posner 1981, p 68-69.

As long as mankind does not live in paradise but rather has to struggle with the problem of scarcity, it seems morally and ethically imperative to avoid wasting resources; that is exactly what the maximisation of wealth means. Maximising the size of the pie is morally required, since with a bigger pie distributional goals and equity can be realised more easily than with a smaller pie. In this sense, it is beside the point to talk about an efficiency-equity tradeoff. Indeed, as can be shown more formally, if there are no wealth effects, and there is the possibility of side payments (taxes and subsidies), everybody in a society is better off with wealth maximisation.[162] If there are wealth effects, framing effects or a status quo bias, the analysis will be much more complicated, but welfare economics does show how to deal with these effects.[163]

If one can make offsetting adjustments to the income tax and transfer system in order to ensure that the overall distribution of income remains the same, "the pure efficiency test will be determinate, and, more importantly, it will indicate which reforms raise individuals' well-being and thus increase social welfare".[164]

2.6.5 Rational individual welfare maximisation

The analysis of sections 2.2-2.5, which leads to the conclusion that the cheapest cost avoider principle can guarantee efficiency whereas the polluter pays principle cannot, is based on the assumption that all actors are rational maximisers of their own welfare. One may find economic assessments implying this assumption problematic, since individuals do not always behave strictly rationally. We all know that individuals may be compulsive, myopic, inconsistent, confused by uncertainty, irrational in their reactions to risk; they often have a preference for fairness and are motivated, at least in part, by the desire to adhere to social norms.[165] Accepting this, however, does not force us to accept the implications which are sometimes supposed. It is often argued that public policy cannot rely on models (scenarios) based on the assumption of rationality because it is not realistic.

At least four arguments speak against such a conclusion:

First, the standard assumption of rational maximisation has proven to be useful in a wide range of settings, in particular in studies of the behaviour of private enterprises. Rational maximisation is often a good approximation of decision making behaviour, and especially in situations characterised by the anonymity of interacting parties. Assuming a type of bounded rationality would not change the results much, since even here it is assumed that

[162] See Milgrom and Roberts 1992 p 35-39.

[163] See Kaplow and Shavell 2002, p 460-461.

[164] See Kaplow and Shavell 2002, p 460.

[165] See Kaplow and Shawell 2002, p 461.

individuals try to do their best: economic actors are assumed to be "*intendedly* rational, but only *limitedly* so".[166]

Secondly, one cannot reject the cheapest cost avoider principle and favour the polluter pays principle on the grounds that the analysis leading to the superiority of cheapest cost avoider principle assumes strict rationality; indeed, the analysis used to recommend the polluter pays principle also relies on this assumption.

Thirdly, it is clear that human behaviour is complex and not always easy to predict. However, if one seeks to make problems tractable enough to provide for some illumination, we need to make simplifying assumptions which do not capture reality. The optimal scale of a map depends on the context that is going to be used in; a map with a scale of 1:1 is clearly of no use at all. The same applies to modelling human behaviour.

Fourthly, as it turns out that predicting the effects of a policy is very sensitive to the behavioural assumptions made, and because behavioural economics, cognitive psychology, evolutionary biology, sociology, or anthropology yield valid insights, they should be incorporated into public policy analysis.[167]

2.6.6 The value of the protection of nature

Is the intrinsic value of the protection of nature neglected? Before assessing the validity of this criticism it seems useful to repeat the fundamentals of external costs:

- External costs are the result of conflicting interests in the use of a scarce resource: nature (environment).
- Without rivalry there are no external costs.
- Consequently damage (= external cost) is jointly caused and the treatment of the issue should take the reciprocal nature of the problem into account.
- Damage is always a loss of value to *somebody* from a change in the state (or quality) of the environment.

Economists take an anthropocentric stance. The reason why the environment (including animals) is in need of protection is simply this: a change in the state (or quality) of the environment is the loss of value to somebody. Due to the reciprocal nature of the externality, preventing this loss of value is necessarily accompanied by the creation of a loss of value to somebody else. Thus, nature is not taken into account as such. Rather, we include environmental considerations via their impact on persons.

Consider the problems of giving nature a separate status, similar to that of the actors. How to evaluate a change? The environment changes through

[166] Simon 1961, p xxiv, quoted in Williamson 1985 p 45.

[167] See Kaplow and Shawell 2002, p 462.

natural and man made factors. If, say, an oak dies and a birch grows in its place, would one consider this as damage to the environment? Surely not. If an oak is felled and a birch is planted instead, would that be damage? It could be, if oaks are considered as more valuable than birches. If not, there is no reason to consider the change as damage.

Consider another example. Man Friday is alone on his island. His actions have no effect on other islands. He decides to fell all the trees and to plant a flower garden instead. If he does so, this implies that he gains from living on an island covered in a flower garden instead of trees. No damage is caused by the replacement of the trees. Now consider the same situation when Robinson Crusoe has arrived on the island. Robinson values the shade of the trees to protect him from the sun. If Friday fells the trees, this comes at a cost to Robinson: damage is done.

The environment is a scarce resource. There are competing uses for it: some benefit from trees, other benefit from felling them; some benefit from clean air, others benefit from emitting exhausts into the air. As a consequence, a change in the status quo of the use of the environment (or another scarce resource) is bound to harm somebody and to profit another.

Our framework makes it possible to incorporate environmental considerations into the analysis without making an *a priori* judgement on whether it is more important to protect the environment or the economy, for example. As long as people value a clean environment, pollution causes damage, and this damage has an impact on the result of the considerations.

2.6.7 The interests of future generations

In the analysis of sections 2.2-2.5 (see especially figures 2-1 to 2-8) it is implicitly assumed that the benefit function of the polluters and the damage function of the pollutees are known. In addition, the relevance of the cheapest cost avoider principle is described in terms of the interests of present generations.

There are two major difficulties with this approach.

Firstly, the benefit and cost functions for all affected parties cannot always be specified. For non market outcomes, shadow prices must be derived. This is not an easy task. Moreover, many externalities, for example ecological externalities, are largely unforeseen. Human beings and other species exist in a complex interaction with their environments. Changes in the environment will have repercussions on living species, existing generations and future generations, which are not easy to isolate.

Despite the current campaign to convince the public worldwide that global warming and climate change will endanger mankind, the precise nature of the many ecological links between humans and their environment and the interdependencies between environmental parameters are still not well understood. Indeed, one can question whether even the theory of complex systems allows predicting the state of the environment, say in 2050 or

2090. This is not to say that we should neglect these possibilities. However, it seems preferable to treat these externalities in the context of uncertainty.[168] In any case, cost-benefit analysis is capable of addressing problems of uncertainty. A policy governed by the cheapest cost avoider principle can draw on these insights.

Many ecological changes will also affect future generations. There are a number of competing ethical positions that one can hold on how to take their interests into account, some of which we list here:[169]

- Intergenerational neutrality represented by a near-zero time discount rate[170];
- Each generation should have at least as much societal capital (tangible, natural, human and technological) as it inherited. A wide array of time discount rates would be admitted;
- Societies should maximise the well-being of the poorest generation. As a consequence, the current consumption should *increase* sharply to reflect the projected future improvements in productivity;
- The minimax or precautionary principle implies that societies should maximise the minimum consumption along the riskiest path.

However, none of these approaches can resolve the fundamental problem in taking future generations into account, which is that they do not yet exist, and that therefore their preferences cannot be known. As Dasgupta and Pearce mention, judgements must therefore be made on two bases: "Either the current generation *alone* must count, *or* the decision-maker must judge *on behalf* of future generations, guessing as to their likely preferences. Cost-benefit analysis tends either to limit society to present generations, or it implicitly assumes that future generations will have a want structure very much like the existing one. As such, the two approaches tend to produce the same answers".[171]

Four problems need to be addressed before undertaking a cost-benefit analysis:

- Is it reasonable to expect that future generations have a want structure very much like the existing one?
- How can the future generations be committed to want what we want them to want?
- What is the weight that the present generation attaches to the interests of future generations? The question is that of determining the "correct"

[168] See Dasgupta and Pearce 1978, p 174-198, Dasgupta 1982, Pearce 1983, p 73-89.

[169] See Nordhaus 2007, p 692.

[170] This is the approach adopted in the Stern report.

[171] Dasgupta and Pearce 1978, p 129.

social discount rate.[172] Who is to decide upon this rate? Scientists cannot do so, since the decision on the value of the social discount rate is necessarily based on value judgements. In a democracy, voters should possibly make this choice.[173] However, individuals and politicians tend to be notoriously myopic. In this respect, it is of a particular importance to realise that ecological externalities frequently take on the attribute of "irreversibility", i.e. they cannot be altered.

- How to take account of the fact that future generations are probably much richer than the present generation?

The upshot is that the cheapest cost avoider principle, properly defined, neglects neither uncertainty nor the interests of future generations. To define the cheapest cost avoider principle properly means to give it an inter-temporal dimension. This raises the question: who is the cheapest cost avoider: the present or the future generations?

2.6.8 Costs of the cheapest cost avoider principle

Whether a governmental decision procedure is welfare maximising not only depends on its theoretical accuracy in tracking overall welfare, but also on its decision costs. It might well be the case that a decision procedure is in principle welfare maximising compared to available competitor theories, but that its decision-making costs are so high that, at a first glance, an inferior procedure should be chosen after all.[174]

As a form of cost-benefit analysis, the decision costs of the cheapest cost avoider principle are those of the cost-benefit analysis. These include both direct costs, such as wages for agency staff, the cost of information gathering, processing costs of carrying out the procedure, overhead costs, fees for the work of analysts and advisors, and the costs of the delayed undertaking of beneficial policies.[175] Cost-benefit analysis can entail relatively high costs.[176]

To evaluate the soundness of this complaint one would need concrete figures. Morgenstern and Landy have collected estimates of the direct costs of preparing cost-benefit documents for an Office of Management and Budget (OMB) review.[177] As Adler and Posner put it: "These data suggest

[172] What can go wrong in this respect can be exemplified by the errors committed in the evaluation of the costs and benefits of the Californian policy. See Stavins et al. 2007.

[173] See Sinn 2007.

[174] Of course, whether a governmental decision procedure is welfare maximising also depends on "the degree to which the political and institutional context in which the procedure is embedded prevents opportunism or mistakes by the decision-makers who are supposed to be implementing it" (Adler and Posner 2006, p 63). For a development of this issue, see Adler and Posner 2006, p 101-123.

[175] See Adler and Posner 2006, p 80 and Morgenstern/Landy 1997, p 455, 461-462.

[176] See Driesen 1997, p 601-605; Glicksman 2003, p 120-146; McGarity 1998, p 7.

[177] See Adler and Posner 2006, p 83, Tables 3.2 and 3.3.

that cost-benefit analysis has substantial direct costs, in the vicinity of $1 million - $2 million on average. More precisely, cost-benefit analysis' direct costs, as compared to a decision procedure with zero direct costs – one that involves no information gathering, computation or analytic effort by the agency, seems to be of $1 million - $2 million on average."[178]

One can imagine administrative decision procedures with nearly zero direct costs, for example if a statute instructs an agency to issue particular rules, with the subject matter specified in detail. Of course, such a statute does not leave an agency any discretion.[179] As Adler and Posner rightly put it: "although the agency itself incurs no direct costs, the process of information gathering and analysis is simply shifted from administrative officials to legislative staff".[180]

Furthermore, the comparative institutions approach is pertinent. As long as the competing procedures, such as the polluter pays principle, are general devices to structure agency choice in cases where agencies retain statutory discretion, "(a)ll such procedures will involve nontrivial information gathering, computation, and analytic effort by agencies"[181]. In addition, the costs of committing errors of type I (false positives), and type II (false negatives), have to be taken into account.

Even if it turned out that the costs of the cheapest cost avoider principle are higher than those of the polluter pays principle, this does not mean that the cheapest cost avoider principle should be rejected. For sufficiently large projects, the added value of the cheapest cost avoider principle compared to the polluter pays principle can justify high analytical and information gathering expenditures.[182]

However, this is only one side of the medal when evaluating cost-benefit analysis. There is another side: "Assuming cost-benefit analysis is more accurate in practice than competitors (taking into consideration not just intrinsic accuracy but also agency mistakes and opportunism), this direct cost will be swamped by the expected benefits of cost-benefit analysis".[183]

The other common objection to cost-benefit analysis is that it is bound to substantially delay or block agency decision making. As leading opponents of cost-benefit analysis put it: cost-benefit analysis means "paralysis by analysis"[184]. We cannot discuss the delay issue here, which is linked to a more general worry about the structure of administrative decision making, in detail. For a short, but careful analysis we refer to Adler and Posner (2006),

---

[178] Adler and Posner 2006, p 80.

[179] See Adler and Posner 2006, p 80.

[180] Adler and Posner 2006, p 80.

[181] Adler and Posner 2006, p 80.

[182] See Adler and Posner 2006, p 83.

[183] Adler and Posner 2006, p 87. For the issue of intrinsic accuracy, see Adler and Posner 2006, p 88-98; for the issue of opportunism see Adler and Posner 2006, p 101-123.

[184] McGarity 1998 p 50.

who point out "that existing evidence and the political economy of rulemaking call into question the claim that cost-benefit analysis produces substantial incremental delay, as compared to non welfare focussed, narrow welfare focused, or hybrid procedures (…)(and that) cost-benefit analysis can be expected to have small incremental decision costs – both direct and delay costs – as compared to its 'wide' competitor, namely, intuitive balancing."[185]

2.6.9 Social constraints

If the cheapest cost avoider principle is clearly the better policy approach, why is the polluter pays principle dominant in practice? Three arguments seem particularly pertinent:

1) An important part of the answer involves social norms that guide well socialised members of society in everyday life. It seems fair that he whose action is the source of damage should be held liable. The attachment to social norms, whether due to socialisation or to evolution, which tend to have the function of promoting individuals' well-being in ordinary interactions, does not imply that analysts and policy makers elevate them to the status of independent evaluative principles for assessing public policy.[186] As Kaplow and Shavell rightly put it: "Indeed, it would be ironic to treat social norms as the basis for giving weight to notions of fairness if in fact the purpose of the social norms is to promote individuals' well-being in the contexts in which the norms have arisen – because the consequences of treating notions of fairness as independent principles for policy analysis can only be to reduce individuals' well-being."[187]

   Policy makers and analysts should be aware that justifications for relying on notions of fairness, depending on our instincts and intuitions and on our general sense for what is appropriate, which make sense for many choices that individuals confront in everyday life, may be inadequate for the design of socially optimal policies.

   Given our instincts and intuitions, it is reasonable to expect that scholars, analysts and politicians will continue to advance notions of fairness as principles which should guide public policy, and in particular environmental policy. If a notion of fairness is to be taken seriously, five issues need to be addressed.[188]

   - "First, proponents of a notion of fairness must state the principles they are defending with some degree of precision and in a manner that is reasonably complete."

[185] Adler and Posner 2006, p 87.

[186] See Kaplow and Shavell 2002, p 467, discussing notions of fairness that correspond to social norms.

[187] Kaplow and Shavell 2002, p 467.

[188] See Kaplow and Shavell 2002, p 470-472.

- "Second, adherence to notions of fairness often leads to consequences that seem to conflict with the underlying motivations of the notions, a tension that the notions' proponents need to resolve."
- "Third, because notions of fairness sometimes result in a reduction in individuals' well-being – and in certain cases lead to a reduction in everyone's well-being – when they are given weight as independent evaluative principles, the manner in which the notion of fairness sacrifices welfare should be identified clearly so that it will be possible to appreciate what is at stake in adopting the principle."
- "Fourth, the rationale for giving weight to notions of fairness should be made clear."
- "Fifth, it is necessary to consider possible alternative explanations of the source of the underlying attraction of notions of fairness."

2) Another reason for the tendency to think along the lines of the polluter pays principle in the public arena is a one-sided view of the problem of external costs. One can observe much confusing rhetoric in this field. The environment is seen as a subject that is able to suffer from pollution or to be grateful for a reduction in pollution; it is anthropomorphised, giving it a status of dignity and attaching intrinsic value to it. This rhetoric seems convincing; however, it ignores fundamentals of the notion of external cost.
3) The last explanation for the dominant role that the polluter pays principle plays in the way of thinking about environmental problems that we want to mention is the self interest of government decision makers, especially politicians and public officials. Of course, as the rhetoric goes, their task is to choose policies that best advance the citizens whom they serve. As the economic theory of democracy[189] teaches, public officials who compete for votes to be elected or re-elected have incentives to promote the policies which serve their purpose best. Of course, their task is complicated by the fact that their constituents may often not be able to understand what really matters.

   Consider the scenarios used in section 2.2. There is "the transport industry" on the one hand, which is supposed to be able to pass additional costs on to its customers, who in turn behave in a similar way. Through a trickling down effect, the burden is distributed onto the whole economy. On the other hand, there are thousands of residents who are potential voters. In this case, which kind of policy probably has more appeal to a politician: one based on the polluter pays principle or one based on the cheapest cost avoider principle? Politicians will clearly prefer the polluter pays principle.

   Nevertheless, in the long run, government decision making may be able to make better policy decisions if those who analyse and advise decision

[189] See the literature on rent-seeking, for example, Buchanan et al. 1980, Mueller 2003.

makers devote themselves to identifying and recommending the rules and principles which best serve the common good.

In summary, part II compares the polluter pays principle with the cheapest cost avoider principle. The underlying logic of the polluter pays principle, i.e. the Pigovian way of economic thinking, is fallacious because the mere existence of externalities does not, of itself, provide any reason for governments to induce polluters to take action. Indeed, the polluters might well be the highest cost avoiders. Thus, full internalisation is not always socially useful in that it does not necessarily maximise welfare. Moreover, the polluter pays principle does not take into account the fact that externalities are caused jointly, i.e. that both the polluter and the pollutee "cause" the damage. Finally, the basic insight is that of Ronald Coase's seminal 1960 article which states that the externality problem is of a reciprocal nature.

The real question that needs deciding is: Should the polluter be allowed to harm the pollutee, or should the pollutee have the right to restrain the polluter? The problem is to avoid the most serious harm, and its solution is the cheapest cost avoider principle.

The cheapest cost avoider principle requires policy-makers to undertake some form of cost-benefit analysis for major regulatory proposals. They have to prepare what has become known as a regulatory impact analysis, regulatory impact appraisal, or regulatory impact assessment. Maximising the welfare of society, i.e. the sum of the welfare of its members, requires a comparison of the costs and benefits of different regulatory options and the choice of the option which promises the highest net benefit to society.

The cheapest cost avoider principle presents a number of clear advantages over the polluter pays principle:

- It guarantees efficiency, i.e. no waste of resources, which is in turn fundamental in the pursuit of the Lisbon goals of jobs and growth.
- It is a better means to achieve fairness than the polluter pays principle.
- It studies a broader set of options. In contrast to the polluter pays principle, it can lead to the choice of innovative projects.
- Its use of cost-benefit analysis in a welfare economics framework makes it take into account a much broader range of relevant variables, such as administration costs or values.
- The logic of the cheapest cost avoider principle helps to avoid regulatory failure and contributes to the success of the Commission's Better Regulation Agenda at the heart of which is regulatory impact assessment.

The cheapest cost avoider principle clearly beats the polluter pays principle in terms of efficiency. However, one should ask whether efficiency or wealth maximisation are all that matters. Of course, notions of corrective justice, moral values, distributive justice, fairness between transport modes as well as the administration costs should be taken into account when choosing the principle to guide public policy. However, we find that the efficiency deficits

of the polluter pays principle cannot be outweighed by the mentioned factors.

There are also a number of concerns about the applicability of the cheapest cost avoider principle. However, many concerns mentioned present practical challenges to all types of public policy analysis instead of being a criticism of the theory underpinning the cheapest cost avoider principle. There is no criticism that can be made of the cheapest cost avoider principle that is not also valid for the polluter pays principle. Thus, the cheapest cost avoider principle should guide public policy.

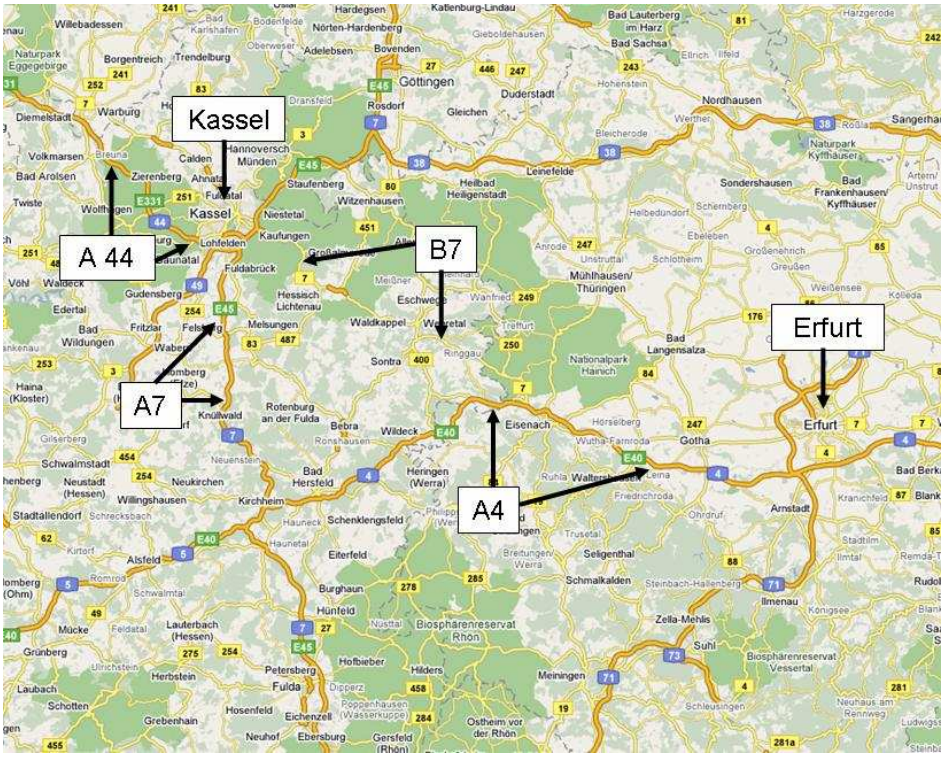
3 Case Studies

3.1 Introduction

In this part, we will apply the cheapest cost avoider principle to two real world cases: the question of the construction of the missing part of the motorway A44 near Kassel in Germany, and the problem of the sectoral limiting of the use of the Inn valley motorway in Austria. These case studies, although sketchy, show that it is possible to apply the cheapest cost avoider principle to such cases, that a clear answer can be obtained, and that judicial decisions are based on the same methodology.

In the first of the case studies, which we will call the A44 case, there exist plans for the motorway A 44 connecting Dortmund to Kassel to be extended in the direction Erfurt, Chemnitz and Dresden, via Eisenach (see fig. 3-1). However, except for a small section, these plans have never been realised. Instead, there is a direct road B7 between Kassel and Eisenach along the route of the planned motorway. This route is closed to trucks, who are obliged to drive a detour of 42 km via the motorways A7 and A4. In what follows, we will make a regulatory impact assessment in order to determine whether this is efficient, or whether it would be better for trucks to use the B7 or to build the missing motorway link between Kassel and Eisenach.

Figure 3-1 Map of the A44 case

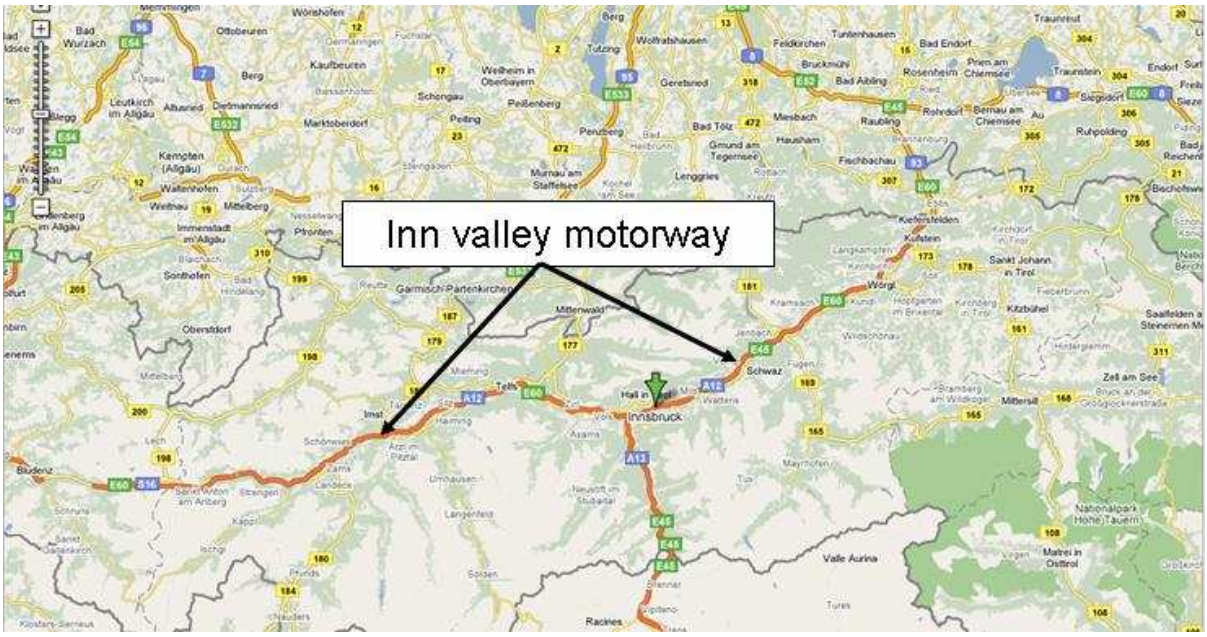


Source: Google Maps

Case 2 concerns the prohibition of a 46 km long section of the Inn valley motorway for trucks transporting goods belonging to a number of sectors (see fig. 3-2). Following the decision of the European Court of Justice, this

prohibition is not in force. We will apply the cheapest cost avoider analysis in order to determine whether closing the motorway to certain trucks is indeed the efficient solution, or whether the externalities can be reduced at a lower cost by other means.

Figure 3-2 Map of the Inn valley motorway



Source: Google Maps

In the first case, the transport industry loses from the status quo compared to alternative scenarios. It would save costs if trucks were allowed to use the B7, or if the A 44 were extended to Eisenach. In the second case, the concerned sectors would suffer losses if they were prohibited from driving on the relevant section on the Inn valley motorway.

3.2 Methodological considerations

Transport traffic imposes external costs on residents: noise, pollution, an increased number of accidents etc. The polluter pays principle prescribes that the transport industry should be made liable for the damages caused. This can take the form of a Pigovian tax or, as is the case here, of a regulation forbidding trucks to drive through the affected area. This measure reduces the externality to zero; in the case where certain trucks can use the motorway during limited time periods, the externality is reduced but not eliminated.

The preceding part has shown that it is not necessarily the most efficient solution to make the polluter pay. Only an analysis following the cheapest cost avoider approach can determine where the optimal outcome lies.

We have also seen that the cheapest cost avoider methodology shares some of the polluter pays principle's downsides, i.e. the difficulty of applying a first best solution and incomplete information. There is no infinity of possible regulations: regulators can decide on whether to allow transport traffic or not, they can limit it to daytime hours but they cannot limit it to 16 hours 32

minutes a day, even if this were the theoretical optimum. Even if they could, the enforcement of such a rule might be too costly. The number of scenarios is thus limited.

The same applies to data. It is not always possible to correctly estimate the exact costs and benefits of possible scenarios before they are actually realised. Even the real world status quo cannot be correctly evaluated: from a data point of view, it is impossible to take into account all externalities to every possible person, as well as all costs of reducing the externalities. For example, it is sufficient for a resident to move away from the concerned area to reduce the externality. A reduction in the price of a good in a distant market can make the concerned transport sector give up using the motorway because the goods can no longer be sold at a price that covers the transport costs, thus reducing the cost of closing the motorway. These are only examples of the huge number of possible changes that can impact on the costs. All cost estimations must therefore be treated with caution. What is more, some of the necessary data is simply not available, or it is flawed. The same arguments hold for the benefit side.

This does not, however, mean that it is impossible to make a decision. Not doing anything, i.e. maintaining the status quo (BAU), is a decision in itself. Any decision should be justified by applying the cheapest cost avoider approach to the available data. Even incomplete data can indicate where the best solution lies.

Which data is needed to be able to proceed with the analysis? We need to establish an estimate of how much polluter pays measures cost or would cost the transport industry. This is compared to the externality costs generated in the different scenarios, and to the costs of other actors, i.e. the residents, of reducing the externalities. Additional costs, such as public investments, are taken into account where necessary. We need to know the time frame, the social discount rate and possible future development (risk and uncertainty).

The social discount rate indicates the value of future benefits and costs. "Global climate change unfolds over a time scale of centuries and, through the power of compound interest, what to do now is hugely sensitive to the discount rate that is postulated. In fact, it is not an exaggeration to say that the biggest uncertainty of all in the economics of climate change is the uncertainty about which interest rate to use for discounting."[190] The less the future is valued, the higher it is, and the less the same face value in the future is worth now. The social discount rate is case specific. It is habitually set at 5-6%.[191] However, much lower discount rates apply to projects that

[190] Weitzman 2007, p 705.

[191] HEATCO 2005, p 15 quotes EC DG Regional Policy 2002 as "suggesting the use of a European discount rate equal to 5%". Weitzman 2007, p 707, calculates a reasonable discount (= interest) rate at 6%.

affect future generations in order to be able to take their interests into account (see section 2.6.7).[192]

It is also possible to take growth into account. Growth can compensate the discount effect. A social discount rate that is only slightly higher than the growth rate leads to a low net discount rate.[193]

3.3 The case of the missing A44

As stated in the introduction to part III, the case of the missing A44 is one in which trucks are forced to make a considerable detour. They are not allowed to use the direct road B7. Except for a short stretch, plans to extend the A44 motorway to create a direct link between Kassel and Eisenach have not been realised.

3.3.1 Cheapest cost avoider analysis

The Kassel chamber of commerce (hereafter IHK Kassel) conducted a series of interviews with companies. It discovered considerable adverse consequences of the detour, both economic and environmental.

The economic consequences for the transport industry result from the extra time that it takes to make the detour. A truck takes one hour longer to drive along the A7/A4 route than it would if the A44 were extended. Per year, this implies 780,000 extra driving hours, and 39 million € extra costs at 50 € per truck hour. Close to 100% of the interviewed companies estimate that the extension of the A44 is urgently necessary.

Concerning the externalities generated by the detour, the IHK Kassel estimates that 3,000 trucks drive the 42 km detour every day, driving 32.76 million extra kilometres a year (at 260 working days a year). 18 litres of diesel are needed to drive the 42 km; as a consequence, 14 million supplementary litres of diesel are burnt. The resulting extra emissions per year are given as 37.627 t of CO_2, 213 t CO, 63t HC, 445 t NO_X, 9 t particles and noise pollution on the 42 extra km. Note that the IHK Kassel study does not translate these figures into monetary costs.

Other studies do provide figures for the externalities. The CE study suggests an average load factor of 15t per vehicle. For 32.76 million extra kilometres, this implies 491,240,000 extra t per km per year. Using the CE estimates of the total externalities for inter urban HGV traffic, we calculate an externalities cost interval between 147,323,000 € and 589,488,000 €[194]. The HGV €/tkm are estimated between 0.3 and 1.2. The Infras study however

[192] See for example the Stern report 2007, which applies a discount rate of 1.4%. This low interest rate is however considered too low by many commentators; see for example Weitzmann 2007, p 705-709.

[193] In economics, the relation between the discount and the growth rate is summarised in the Ramsey equation: $r = \delta+\eta g$, where r is the interest rate, δ is the rate of pure time preference, η is the elasticity of marginal utility, and g is the growth-rate of consumption. See Weitzmann 2007, p 706 and HEATCO 2005.

[194] CE Study, p 32.

assumes only 0.0712 €/tkm. This figure implies the much lower cost of externalities of 34,878,040 €. Note that this figure represents a seventeenth of the upper estimate established with the CE study figures.

Based on these figures, let us construct 3 scenarios in order to determine threshold values that would justify the status quo.

Table 3-1 Extra costs of the status quo

In million €	**Extra costs of the status quo**		
	Infras (=low)	CE low (=medium)	CE high (=high)
Economic costs for the transport industry	39		
Externality costs	34.88	147.37	589.49
Total	73.88	186.37	628.49

The values in Table 3-1 summarise the extra costs compared to trucks driving along the A44.

However, the A44 still needs to be built. Motorway construction is financed by the tax payers. If it costs less than the lowest estimation of the extra costs of the detour to extend the motorway, then the cheapest cost avoider principle states that it would be socially optimal for the motorway to be built. Costs that are engendered in the status quo would be avoided at the lesser cost of extending the motorway.

If, on the other hand, the motorway link from Kassel to Eisenach cost more than the highest estimate of the total cost of the detour, then it is socially optimal not to build the motorway, in spite of the costs of the detour.

We have three estimations of the cost of building a kilometre of motorway. The German ministry for traffic estimates that 1900 km motorway can be built for a total cost of 15 billion €, or approximately 7.7 million €/km.[195] The estimates of the costs of the A44 are of 15 million €/km.[196] Finally, we assume a "worst case" scenario with high construction costs of 27 million €.

The length of the route of the A44 between Kassel and the A7 close to Eisenach is approx. 63km.[197] Table 3-2 compares the estimated costs of the detour to the costs of the construction of the motorway with a one year horizon.

[195] See Bundesministerium für Verkehr, Bau und Stadtentwicklung 2007.

[196] See Aktionsgemeinschaft Verkehr Nordhessen 2004.

[197] See Aktionsgemeinschaft Verkehr Nordhessen 2004.

Table 3-2 Cost-benefit analysis detour – A44, 1 year horizon (Million €)

Detour cost scenario	**Low**			**Medium**			**High**		
Motorway construction scenario	Low 7.7	Medium 15	High 27	Low 7.7	Medium 15	High 27	Low 7.7	Medium 15	High 27
Extra costs of the detour	73.88			186.37			628.49		
Motorway construction costs 63 km	485.11	495	17011	485.11	495	17011	485.11	495	17011
Difference detour – motorway construction	-411.22	-871.12	-1627.12	-298.73	-758.63	-1514.63	**143.69**	-316.21	-1072.21

Table 3-2 shows that all other things held equal, there is only one combination of figures that would make building the motorway beneficial from its first year. This is the highlighted case, i.e. where the costs of the detour are very high, and the costs of the motorway correspond to the lowest of the estimates. It would save 143.69 million € to build the motorway in this case.

In all following years, suppose there are no costs (the motorway is built), and only benefits from avoiding the detour. Assume that the decision maker considers a 10 year horizon with no discount. Externalities and economic losses are repeated every year, i.e. they are multiplied by 10. However, the motorway is only built once. We assume that its maintenance costs are negligible, and that the amount of traffic remains unchanged. Table 3-3 compares the possible different outcomes using this 10 year horizon.

Table 3-3 Cost-benefit analysis detour / A44, 10 year horizon (Million €)

Detour cost scenario	**Low**			**Medium**			**High**		
Motorway construction scenario	Low 7.7	Medium 15	High 27	Low 7.7	Medium 15	High 27	Low 7.7	Medium 15	High 27
Extra costs of the detour	738.8			1863.7			6284.9		
Motorway construction costs 63 km	485.11	495	17011	485.11	495	17011	485.11	495	17011
Difference detour – motorway construction	252.9	243	**-963**	1378.6	1368.7	162.7	5799.8	5789.9	4583.9

The situation is now inversed: if we consider a 10 year horizon, in all figure combinations but one, it is preferable to go ahead with building the motorway, because building the motorway is less costly than bearing the economic and environmental costs of the detour. Using the 10 year horizon makes us arrive at the opposite conclusion to the one year horizon: on the condition that we exclude the extreme case, i.e. low detour costs and high motorway construction costs, it is inefficient to let the polluter pay, i.e. to direct HGV on the detour. Instead, social welfare would be increased if the motorway was built. This result is robust although we included a scenario with very high motorway construction costs. It therefore takes into account and corrects the "optimism-bias", which is the systematic ex ante underestimation of costs.[198]

A number of additional considerations accentuate this result. For example, it is plausible that HGV traffic will increase as a consequence of economic growth and of increasing trade between Eastern and Western Europe. In this case, the detour costs would rise, and it would again clearly pay to construct the motorway. On the other hand, traffic might increase because of the construction of the motorway. Then the motorway would lead to extra emissions, causing externality costs which would possibly tip the balance in favour of the detour. Generally speaking, we can say that the longer the time horizon, and on the condition that the social discount rate of the detour costs is not too high[199], the clearer it becomes that under status quo conditions filling in the missing link in the A44 is the least cost option for society.

The status quo is to block truck traffic from using the B7, sending it on the A4/A7 detour. Given that the A44 does not yet exist, we use the cheapest cost avoider principle to test whether this solution is efficient. How does the detour compare to the B7?

Consider first economic costs. According to Google Maps, the detour takes 1h23, while the B7 route takes 1h25: while the B7 is more direct, it applies stricter speed limits. We therefore assume that there is no significant economic difference between the travelling time of the two. There are no construction costs: the only distinguishing variable concerns the externalities.

We know that the detour is approx. 42 km longer than the B7 route; we have already estimated the extra externality that is implied. The extra costs of the detour indicate that the B7 would be the cheapest cost avoider route. This result holds even if one sets the externality relative to the concerned population: the towns and villages along the B7 comprise a population of approx. 30.000. One of the many towns along the detour route, Bad Hersfeld, has the same number of inhabitants. As a consequence, all figures

[198] For a discussion of the optimism-bias, see HEATCO 2005.

[199] For simplicity, the social discount factor in Table 3-3 is zero. This assumption is not invalidating. It is discussed further on.

indicate that the detour involves higher externalities which concern a higher number of persons than the B7 route (see Table 3-4).

Table 3-4 Comparing detour to B7

In million €	**Detour**			**B7**
Population concerned	> 30.000			~ 30.000
Extra externality costs	738.8	1863.7	6284.9	0

Why, then, are the trucks redirected onto the detour? We consider three possibilities. The first is that we have neglected a significant variable in the above reasoning. If, say, the population close to the existing motorways does not suffer a high utility loss from the extra pollution caused by the detour, and if the population along the B7 suffered an extremely high utility loss from having trucks drive through their villages, then the result would be inversed and the detour would be the cheapest cost avoider solution.

It could also be that HGV traffic causes high repair costs on the B7.

An alternative explanation is that the damage imposed on the B7 residents by the trucks was treated according to the polluter pays principle, i.e. making the transport sector responsible for the reduction of the damage caused by it (in this case to zero), without taking into account that it might be more efficient to allow the trucks to impose the harm on the B7 residents, rather than to make them generate greater harm on the larger number of persons living along the detour.

Now let us compare the construction of the A44 to opening the B7 to HGVs. Assume that as the route is the same, so the externality costs and the concerned population do not differ between the two. The A44 needs building, but it has the economic advantage of reducing the time it takes to travel from Kassel to Eisenach.

Tables 3-5 and 3-6 show that it would only be better to build the A44 if one adopts a long term horizon. If traffic does not increase, the economic benefits from the faster A44 do not outweigh the costs of constructing the motorway section in any constellation of the figures. However, if the economic benefit increases, either because one adopts a longer time horizon as in Table 3-6, or because traffic has increased, it can become preferable to build the motorway, on the condition that the construction costs are not too high.

Table 3-5 Ten year horizon

	Cost construction A44		
	Low	Medium	High
	-485,1	-495	-1701
Economic benefit	390	390	390
Total	-95,1	-105	-1311

Table 3-6 Twenty year horizon

	Cost construction A44		
	Low	Medium	High
	-485,1	-495	-1701
Economic benefit	780	780	780
Total	**294,9**	**285**	-921

It would clearly favour the transport industry to construct the missing link of the motorway. From an efficiency point of view, this can also be society's best choice.

Table 3-7 Minimum number of years to be taken into account to make the motorway the least cost option

Mimimum no. of years	Low detour costs	Medium detour costs	High detour costs
Low motorway construction costs	7	3	0
Medium motorway construction costs	13	6	2
High motorway construction costs	/	10	3

Table 3-7 shows that in the example, for a social discount rate up to 1%, the number of years from which it becomes preferable to build the motorway depends on the chosen scenario. It exceeds 13 only in the worst case scenario of high costs for the motorway and low costs generated by the detour. The motorway construction becomes attractive relatively quickly and can, in the best case, save costs from the year it is built. Even the longest time horizon does not exceed 13 years, which is a gross underestimation of the length of time of the use of a motorway. The higher the social discount rate, the smaller the gain from building the motorway, and the faster the difference between the options becomes equal to zero.

3.3.2 Legal reasoning on the B7 ban

While not explicitly mentioned, the cheapest cost avoider principle is implicit in a great number of sentence justifications in all legal systems, including civil law countries like Germany[200]. Let us take a judgement concerning this case study.

In August 2005, the responsible legislator decided on a ban which prohibited trucks over 3.5 t from driving on certain parts of the B7 and the B27. Following a legal suit by a local transport firm, this ban was examined in court and found to be valid.[201] Interestingly, the explanation of this judgement follows the method of the cheapest cost avoider principle.

The court recognises that it has to consider and compare the interests of two parties: the transport industry, which has to make a detour, and the residents, who are subjected to the noise and pollution caused by the trucks. To be valid, the gain from the ban must be proportionate to the tort caused by it (in German, this is the principle of *Verhältnismäßigkeit*). The decision, it is explicitly mentioned, must be in the interest of social welfare (*Gemeinwohl*).

The ban is expected to reduce the number of trucks on the B7 and the B27 by 40-60%. A considerable reduction in noise pollution is obtained from a 40% reduction in the number of trucks. The ban is thus successful at reducing the externality. The claimant provided the court with no information on the costs of the ban for him or the transport industry.

The alternatives to the ban are also studied. It is found that neither a speed limit, nor "reduced noise" trucks lead to a significant reduction of the externality. An alternative route would lead to imposing the same externalities on a different population.

Also, measures taken by the pollutees are considered: sound protection windows are installed. They are taken into account in the calculation of the sound reduction.

The judgement in favour of the residents is based on the fact that the claimant has not proven that his damage is disproportionate compared to

[200] See section 2.4.

[201] Verwaltungsgericht Kassel 2005.

the proven gain for the residents. In other words, in this case, the transport industry is considered to be the cheapest cost avoider.

In economic terms, one can say that the court made a cost-benefit analysis on the basis of the information offered in the hearing. The benefits of the ban were compared to the benefits of other measures and found to be the most effective way of reducing the externality. No information on the costs of the measure being available, the ban can be considered to be the least cost (or most cost effective) measure to reduce the externality. Its benefits outweigh its costs. Therefore, upholding the ban seemed to the court to be in the interest of maximising social welfare.

3.4 The sectoral ban on the Inn valley motorway

In 2003, the Tyrol government decided to introduce a sectoral ban on part of the Inn valley motorway (A 12) on HGVs over 7.5 tons transporting the following goods:[202]

- Waste
- Corn
- Cork
- Ore
- Stones, earth
- Motor vehicles and trailers
- Steel for construction

The Commission of the European Communities successfully applied to the European Court of Justice (EJC) for the suspension of the sectoral ban.[203] In what follows, we will apply the cheapest cost avoider principle to this example, followed by a comparison between the findings of the ECJ and the cheapest cost avoider principle.

3.4.1 Cheapest cost avoider analysis

The sectoral ban on motorway driving on the 43km section of the A12 causes considerable costs to the transport industry, while leading to an increase in ambient air quality. We will discuss these effects in turn.

Baum et al. (2004) studied the effects of the ban on the German transport industry. They found that in Germany only, the costs would amount to 250 million €, plus the loss of 2500 jobs. It can be assumed that the losses for the Italian transport industry are also high. There are a number of causes for these costs.[204]

---

[202] See Baum et al. 2004, p 4.

[203] See European Court of Justice, Case C-320/03 R, 2 October 2003, and 15 November 2005.

[204] Baum et al. 2004.

Insolvencies Transport firms are highly specialised: 80% of firms make specialised investments in equipment for certain goods. A quarter of all firms concerned by the sectoral ban, all of which are small and medium size, see themselves threatened by insolvency. The sectoral ban would lead to a 210 million € reduction in turnover due to insolvencies.

Increase in costs 46% of the concerned companies see the possibility of shifting to a different transport route. Only 30% of the goods would be shifted, involving higher transport costs and longer transport delays. Only 3% of the goods can be transported by train. This mode of transport is not necessarily more costly, but it causes delays. 43% of the concerned firms indicate that they have no transport alternative to the A 12. Two thirds of the goods can no longer be transported. The increase in cost through the sectoral ban is estimated at 118 million €. Again, mainly small and medium sized transport companies are confronted with the increase in costs.

Reduced demand 28% percent of firms, 89% of which have fewer than 50 employees, will see their number of orders fall. Not counting insolvencies, this would lead to a loss of 43 million €.

Job losses Counting that a reduction by 1 million € turnover leads to the loss of 16 jobs, the reduction in turnover would cost another 690 jobs. Insolvencies due to the ban imply the loss of 820 jobs. Statistically, one job in the transport industry is linked to 0.7 jobs in related branches, such as vehicle production or the petrol industry. The total loss of jobs in Germany thus amounts to approximately 2500.

Let us now turn to the benefit of the sectoral ban in terms of the reduction of the costs of the pollution externality. NO_x (oxide of nitrogen and nitrogen dioxide) emissions exceed EU limit values in the status quo. Austria is compelled to take action in order to reduce NO_x emissions below the limit value to acceptable levels.

A 2004 study by the IFEU Heidelberg evaluates the reduction in pollution caused by the ban. It takes into account the nitrogen emissions from different types of HGVs and from passenger transport.

The sectoral ban is predicted to lead to a reduction by 6-12% of NO_x emissions on the concerned route. However, the study points out that supplementary emissions will be caused in different locations because of longer detours and the extra emissions caused by shifting goods onto trains.[205]

The same study considers three alternatives to the sectoral ban: a ban on technically out of date HGVs (according to their "Euro" class), a speed limit for all vehicles, including passenger cars, and the combination of the two.

Depending on how severe the ban on outdated HGVs is, the reduction of NO_x is estimated between 6% and 14%. This result is very similar to that of the sectoral ban. On the condition that transport is shifted onto newer models of

[205] IFEU Heidelberg 2004, p 13.

HGVs, this ban would not imply shifting the pollution to another location, as is the case for the sectoral ban.

A reduction of the speed limit for all vehicles to 100 km/h would lead to only 1% reduction of the emissions of HGVs, but to reductions of 36% and 8% respectively for diesel and petrol engine passenger cars. The total reduction in NO_x is also estimated to lie between 6% and 14%. There are no pollution effects on other locations. Further advantages would be a reduction in the consumption of fuel, in the noise pollution, and in the severity of accidents.

Finally, the combination of a ban on old HGV and a speed limit would lead to a much higher reduction of local NO_x emissions of 17% - 25%. For a summary of the effects, see Table 3-8.

Table 3-8 Pollution effects of measures

	Ban (1)	**Limit on type of HGV (2)**	**Speed limit (3)**	**Combination 2+3**
Reduction NO_x	6-12%	6-14%	6-14%	17-25%
Negative external effects	Increase in pollution (detour)	Yes if detour	Reduction in other polluter	Reduction in other polluter, negative if detour

Based on this information, we can now proceed to the cheapest cost avoider analysis. Note that we have information about the cost of the ban, but not on the cost of the pollution. We can however assume that the reduction in NO_x leads to a reduction in the cost of the externalities. While the costs of the ban seem very high, in a first step it is not possible to say a priori whether or not it is lower than the benefits from the reduction in pollution, i.e. whether the ban should be upheld.

However, we can use a cost-effectiveness analysis to achieve a result by comparing the effects of the different types of measures (Table 3-8). While the ban on old HGVs and the speed limit have very much the same NO_x abating effect as the sectoral ban, we find that the latter has high supplementary costs from the detour. The ban of old HGVs implies lower extra costs, and the speed limit has no extra pollution costs; instead, it is accompanied by benefits. We do not know the economic costs of the alternative measure; however, we can safely assume that they are below the costs of the sectoral ban and, in the case of the speed limit, perhaps even negligible: neither alternative measure would impact on a particular type of good, leading specialised firms to insolvencies, nor would they necessitate detours. Thus, if the aim is to reduce NO_x levels by approximately 6-14%, the most beneficial measure would be the speed limit on all vehicles.

If the aim is to reduce NO_x emissions as far as possible, the sectoral ban would also fail the test: the combination of a speed limit and a ban on old HGVs would lead to a higher reduction of emissions. The combination comes at a higher cost than the simple speed limit, because old HGVs would have to be replaced; the choice between the two measures thus depends on whether the extra benefits of the combination in terms of the reduction of the externality outweigh the costs.

Consequently, the sectoral ban should never be the chosen measure: it is neither the most effective at reducing NO_x emissions, nor is it the least costly concerning other pollution effects. Supposing that all measures have the same economic cost, we conclude that it is not the most cost efficient choice. Furthermore, as noted above, we can assume that the costs of the measures are not the same; the sectoral ban can be assumed to entail the most economic costs. Therefore, it cannot be the cheapest cost solution. It should be suspended.

3.4.2 The reasoning of the European Court of Justice

The European Court of Justice, seized by the Commission of the European Communities, ordered Austria to suspend the ban on 2 October 2003[206]. This order was confirmed by the judgment of 15 November 2005[207]. In its findings, the ECJ balances the interests of the concerned parties.

The ECJ establishes that the sectoral ban presents an obstacle to the free movement of goods (§62 - §69): "The contested regulation must therefore be regarded as constituting a measure having equivalent effect to quantitative restrictions, which in principle are incompatible with the Community law obligations under Articles 28 EC and 29 EC, *unless that measure can be objectively justified*." [208]

The free movement of goods needs to be realised *unless* there is a contrary interest that has even greater and contradictory consequences. Therefore, the ECJ considers the possible justifications of the obstacle. In §70 it recalls that "it is settled case-law that national measures capable of obstructing intra-Community trade may be justified by overriding requirements relating to protection of the environment provided that the measures in question are proportionate to the aim pursued". "In this case, it is undisputed that the contested regulation was adopted in order to ensure the quality of ambient air in the zone concerned and is therefore justified on environmental protection grounds" (§71).

Furthermore, the nitrogen levels measured in the concerned section of the A12 exceeded the limit values determined by the 1999 European directive. As a consequence, "in those circumstances, having regard to the provisions of Article 8(3) of Directive 96/62, the Republic of Austria was under a duty

[206] ECJ 2003 Case C-320/03 R

[207] ECJ 2005 Case C-320/03 R

[208] ECJ 2005 Case C-320/03 R, §69, emphasis added.

to act" (§80). However, the Member State is required to implement a plan or programme which "must contain a series of appropriate and coherent measures designed to reduce the pollution level in the specific circumstances of the zone concerned" (§81). §82 states that the measures of the sectoral ban "cannot be described as a 'plan' or 'programme' within the meaning of Article 8(3) of Directive 96/62, since they are not in any way connected to a specific situation in which limit values have been exceeded". Thus, "even if one were to concede that the contested regulation is based on Article 8(3) of Directive 92/62, it cannot be regarded as constituting a correct and full implementation of that provision" (§83).

The ECJ does however not "preclude the possibility that the obstacle to the free movement of goods arising from the traffic ban laid down by the contested regulation might be justified by one of the imperative requirements in the public interest endorsed by the case-law of the Court of Justice" (§84). §85 states that, "(i)n order to establish whether such a restriction is *proportionate* having regard to the legitimate aim pursued in this case, namely the protection of the environment, it needs to be determined whether it is *necessary and appropriate* in order to secure the authorised objective."[209] In other words, it is necessary that the gain (the protection of the environment) is pursued by the best means, and that it be proportionate to the losses caused by the sectoral ban.

"On that point, the Commission and the intervening Member States stress both the lack of any genuine alternative means of transporting the goods in question and the existence of many other measures, such as speed limits, or toll systems linked to different classes of heavy vehicles, or the ecopoints system, which would have been capable of reducing nitrogen dioxide emissions to acceptable levels" (§86). The "lack of genuine alternatives" entails the high costs of the restriction, as described in 3.4.1. The ECJ stresses the importance of studying other pollution reduction measures as to their cost-effectiveness.

Rather than an option, studying the alternatives is considered an *imperative* for the law maker: "… it suffices to say in this respect that, before adopting a measure so radical as a total traffic ban on a section of motorway constituting a vital route of communication between certain Member States, the Austrian authorities were *under a duty* to examine carefully the possibility of using measures less restrictive of freedom of movement, and discount them only if their inadequacy, in relation to the objective pursued, was clearly established" (§87).[210] Put differently, the ECJ requires Member States to make a cost-benefit analysis, including alternative scenarios.

Thus, "it must be concluded that, because it infringes the principle of proportionality, the contested regulation cannot validly be justified by reasons concerning the protection of air quality. Therefore, that regulation is incompatible with Articles 28 EC and 29 EC" (§91).

---

[209] Emphasis added.

[210] Emphasis added.

Note that, while the wording is different, the ECJ uses the same reasoning and finds the same results as the cheapest cost avoider principle approach: the ban's aim of reducing pollution could have been achieved by less costly measures. The sectoral ban can therefore not be considered proportionate: it leads to unnecessary costs. Limiting the overall costs is the very idea of the cheapest cost avoider principle, which promotes the measures that lead to the least costs.

In this case, it is well possible that a speed limit on all vehicles can be the cheapest cost avoider solution. Is this consistent with "polluter pays"? In a way, yes: all vehicles on the motorway contribute to the pollution. Passenger cars contribute to 51% of nitrogen emissions, compared to 43% for HGVs.[211] HGVs are thus certainly not the only polluters.

However, a large number of the passenger cars belong to Inn valley residents, who are also pollutees. In this case, the polluter pays principle would be tricky to apply: should polluters who are also pollutees be taxed at the same level as the transport industry? The cheapest cost avoider principle however leads to clear answers. These results as well as the method are validated by the European Court of Justice. The methods are even rendered obligatory for the decision making procedure of Member States.

Part III shows that the cheapest cost avoider methodology, i.e. a complete regulatory impact assessment, can indeed be applied to real life cases. Although the analysis can only be incomplete and explanatory in this framework, it provides clear answers on which decision should be taken. Court decisions concerning the case studies obviously apply the same reasoning as the cheapest cost avoider method, showing that this way of thinking is and even must be applied in practice. This finding is confirmed by the EU impact assessment guidelines.[212]

[211] See Ifeu Heidelberg 2004, p 11.

[212] European Commission 2005.

Conclusions

The directive 2006/38/EC requires the Commission to provide a model for the assessment of all external costs of road transport, which are to be internalised in the form of infrastructure charges. A study was commissioned from CE Delft and Infras to offer proposals on how external costs are to be defined and measured, and on why and how external costs should be internalised.

The present study offers a critical assessment of the CE study (part I), in which its narrow focus on the "polluter pays approach" is identified as a fundamental methodological flaw. Another serious weakness can be seen in its failure to discuss matters in the context of the EU Commission's Better Regulation Agenda, with its focus on regulatory assessment. Part II develops the theoretical foundations for the "cheapest cost avoider principle" which is superior to the "polluter pays principle" both methodologically and practically, in identifying the most appropriate policy for dealing with external effects.

The cheapest cost avoider approach takes into account the fact that external costs may be reduced by both the polluter and the pollutee. This insight has two fundamental consequences.

- First, external costs should be reduced no further than a socially efficient level, which takes into account the fact that reducing external costs itself is costly.
- Second, in order to minimise the welfare losses from external effects, it is the party that has the lowest cost that should make the investment; this party is not necessarily the polluter. It is identified by some form of cost-benefit analysis.

The cheapest cost avoider analysis incorporates "polluter pays" as one possible outcome, but does not mandate it without consideration of the alternatives. As a result, the cheapest cost avoider principle presents a number of clear advantages over the polluter pays principle:

- It guarantees efficiency, i.e. no waste of resources, which is in turn fundamental in the pursuit of the Lisbon goals of creating jobs and growth.
- It is a better means to achieve fairness than the polluter pays principle.
- It studies a broader set of options. In contrast to the polluter pays principle, it can lead to the choice of innovative projects.
- Its use of cost-benefit analysis in a welfare economics framework makes it take into account a much broader range of relevant variables, such as administration costs or moral values.
- The logic of the cheapest cost avoider principle helps to avoid regulatory failure and contributes to the success of the Commission's Better Regulation Agenda at the heart of which is regulatory impact assessment.

In this study, we advance the thesis that environmental policy should be evaluated in the light of the principle of wealth maximisation. Wealth maximisation is found to be a good proxy to achieve maximal social welfare. Properly interpreted, it not only reflects people's concerns about the size of GDP, but also their concerns about the distribution of income as well as environmental concerns. The idea that there is a tradeoff between efficiency and equality is incorrect: a bigger pie always makes it possible to distribute more income than a smaller one. Public policy should resist striving for distributional goals by using tools which can be inherently detrimental to wealth maximisation, such as the polluter pays principle. Rather, it should rely on direct measures, i.e. income tax and subsidy schemes.

A comparison of the polluter pays principle with the cheapest cost avoider principle from the viewpoint of wealth maximisation clearly favours the cheapest cost avoider principle over the polluter pays principle. Whereas the cheapest cost avoider principle can guarantee efficiency, the polluter pays principle does not. This result cannot be overstated. It is given additional support from what can be called the ethics of wealth maximisation: In a world struggling with the problem of scarcity, wealth maximisation is a moral imperative. We also deal with the question whether the inferiority of the polluter pays principle in comparison to the cheapest cost avoider principle can be compensated by factors such as corrective justice, distributive justice, fairness between modes of transport and administration costs.

A careful analysis reveals that they cannot. In a symmetrical vein, we discuss concerns which can be raised against the cheapest cost avoider principle despite it furthering efficiency. We find that the cheapest cost avoider principle can meet all concerns, i.e. valuing non pecuniary factors such as life, pain, and suffering; including "soft variables" (intangibles) such as environmental values, wildlife, mountainous scenery, data issues; fairness issues; rational individual welfare maximisation issues; questions on the value of the protection of nature; the interests of future generations; administrative costs; and social constraints.

Thus, we do not find any rationale that justifies following the polluter pays principle instead of the cheapest cost avoider principle.

The cheapest cost avoider principle is operational. As an offshot of cost-benefit analysis, this statement does not come as a surprise. The term "cost-benefit analysis" in its broadest sense refers to the measurement of the economic costs and benefits from any change in the resource allocation in the economy. A narrower, more conventional one refers to the evaluation of the net benefits to society from a specific investment project. Cost-benefit analysis is a well established approach applied in the evaluation of thousands of investment projects financed by public expenditures. Its use is often required by the law. However, the potential of cost-benefit analysis goes beyond the assessment of investment projects financed by public funds: all measures of public policy, and in particular measures devoted to the internalisation of external costs, should be based on a calculation of the net benefits to society. In part III of this study, we demonstrate the basic logic of cost-benefit analysis with the help of two practical examples: the

prohibition for trucks to use the B7 close to the city of Kassel in Germany, and the selective ban of the Inn valley motorway in Austria.

The case studies show that the cheapest cost avoider approach can be applied in practice. The scenarios analysed suggest that it would be beneficial to build the A44 motorway in order to avoid the detour caused by the ban for trucks on the B7. The detour is found to be inefficient. Legal reasoning in a related case implicitly applies the cheapest cost avoider principle. As for the sectoral ban on the Inn valley motorway, a cost-effectiveness analysis reveals that the ban cannot be the least cost option. The European Court of Justice implicitly uses a cheapest cost avoider analysis to come to the same result. Both theory and practice show that the cheapest cost avoider approach can be, is, and must be applied in political decision-making.

The methodology developed in this study applies to policy considerations concerning the transport industry. Its practical relevance however goes well beyond this sector of activity. For example, a related field to which it can be applied and for which it should be further developed is personal transport. Individual means of transport show many parallels to the analysis of the transport industry, while posing some extra challenges due to the fact that drivers are both polluters and pollutees.

This study has concentrated on the cheapest cost avoider methodology; for this reason, it has not exhausted the issues that policy-makers need to consider when taking decisions. For example, the question of the use of funds raised when taxing polluters or pollutees is of fundamental importance. We have seen that when pollutees can take action, taxes should not be used to compensate their damage, in order not to distort their incentives. But what should they be used for? Should they be earmarked for infrastructure projects or for environment projects? Should they be used for the cross-subsidisation of other modes of transport? Another question is that of the macroeconomic impact of policies. Will an increase in the transport industry's costs be fully transmitted to the consumers of the transported goods? This is the issue of the "trickling down effect" mentioned in section 2.6.9. What is the effect on the quantities sold? Further study is needed to address these questions.

Recommendations

- **It is not necessarily only the transport industry (i.e. the polluter) who should be made liable for externalities.** Other actors, such as the state or the pollutee may well be in a better position to take measures to reduce externalities, and they should do so in the interest of economic efficiency and fairness.
- **Not all harm caused by the transport industry should be internalised.** Internalising too much of the damage would cost society more than it would benefit it. An efficient level of damage should be accepted.
- **When introducing new measures to reduce pollution by the transport industry, all existing levies should be taken into account.** Among other taxes, the transport industry pays VAT and fuel tax. Environmental requirements for vehicles also present a cost to the transport industry. Optimal incentives can only be obtained by taking the effects of all these measures into account.
- **The type of measure taken to make the transport industry pay must be based on a broad impact analysis.** Tolls, taxes and speed limits have different impacts on incentives. For example, a km toll would not sufficiently reduce the harm caused by pollution if trucks could employ an alternative route with the same (or even more) emissions. Furthermore, making the polluter pay might not solve the pollution problem and it will give no incentive for the pollutee to solve the problem either
- **Policy-makers should rely on efficiency-maximising mechanisms** rather than on conceptions of fairness (i.e. the polluter should pay) that, on closer scrutiny, prove insufficient and contradictory to the well-being of society.
- **The polluter pays principle should not be used** because its underlying economic thinking is fallacious. It neglects the basic insight that external costs are caused jointly by all parties involved and that the externality is a problem of reciprocal nature.
- **The cheapest cost avoider principle,** for which Ronald Coase received the Nobel Prize for Economics, **should be used**, also for transport, because it guarantees efficiency. It is based on a complete cost benefit analysis (also called regulatory impact assessment), it is a better means to achieve fairness, and it finally leads to better incentives for all parties involved.
- **The cheapest cost avoider principle presents the fundamental tool in the pursuit of the European Union "Lisbon Goals" of growth, jobs and competitiveness.**

References

Ackerman, B. (1984): *Reconstructing American Law;* Harvard University Press, Cambridge/Mass.

Aktionsgemeinschaft Verkehr Nordhessen (2004): *A44: Klagen*. http://avn.cooltips.de/a44klagen.php.

Adler, M.D./Posner, E.A. (2006): *New Foundations of Cost-Benefit Analysis;* Harvard University Press, Cambridge/Mass.

Arrow, (2007): *Global Climate Change: A Challenge to Policy*; in: Economists' Voice. http://www.bepress.com.

Baffi, E. (2006): *The Problem of Internalisation of Social Costs and the Ideas of Ronald Coase;* Social Science Research Network, Working Paper Series. http://ssrn.com/abstract=920588.

Baum, H./Henn, A./Esser, K./Kurte, J. (2004): Auswirkungen des sektoralen Fahrverbots und des Nachtfahrverbots Österreichs auf der A12 Inntalautobahn für das Transportgewerbe und die produzierende Wirtschaft. www.uni-koeln.de/wiso-fak/verkehr/publikationen.html.

Baumol, W. (1972): *On Taxation and the Control of Externalities*; in: American Economic Review 62 (3), 307-322.

Breyer, S. (1979): *Analyzing Regulatory Failure. Mismatches, Less Restrictive Alternatives, and Reform;* in: Harvard Law Review 92 (3), 547-609.

Breyer, S. (1993): *Breaking the Vicious Circle. Toward Effective Risk Regulation;* Harvard University Press, Cambridge/Mass.

Brown, J.P. (1973): *Toward an Economic Theory of Liability*; in: Journal of Legal Studies 2 (2), 323 - 349.

Buchanan, J.M./Stubblebine, C. (1962): *Externality*; in: Economica 29 (116), 371-384.

Buchanan.J.M./Tollison R.D./Tullock G. (eds.) (1980): *Toward a Theory of the Rent-Seeking Society*; Texas A&M Press, College Station.

Bundesministerium für Verkehr, Bau und Stadtentwicklung (2007): *Investitionen in Bundesfernstraßen*. http://www.bmvbs.de/Verkehr/Strasse-,1442/Neubau-Erweiterung-von-Autobah.htm.

Burrows, P. (1980): *The Economic Theory of Pollution Control*; MIT Press, Cambridge/Mass.

Burrows, P. (1999): *Combining Regulation and Legal Liability for the Control of External Costs*; in: International Review of Law and Economics 19 (2), 227 - 244.

Calabresi, G. (1970): *The Costs of Accident. A Legal and Economic Analysis*; Yale University Press, New Haven.

Calabresi, G./Hirschoff, J.T. (1972): *Toward a Test for Strict Liability in Tort*; in: Yale Law Journal 81 (6), 1055 - 1085.

CE Delft/Infras (2007): *Methodologies for external cost estimates and internalisation scenarios.* http://www.ce.nl/eng/framerechts.html.

Coase, R. (1960): *The Problem of Social Cost;* in: Journal of Law & Economics 3 (1), 1-44.

Coase, R. (1988): *Notes on the Problem of Social Cost*; in: Coase, R. (Hrsg.): The Firm, the Market and the Law, University of Chicago Press, Chicago.

Cooter, R.D. (1984): *Prices and Sanctions;* in: Columbia Law Review 84 (6), 1523-1560.

Cooter, R.D. (1985): *Unity in Tort, Contract, and Property. The Model of Precaution*; in: California Law Review 73 (1), 1-51.

Cooter, R.D. (1987): *Coase Theorem;* in: Eatwell, J./Milgate, M./Newman, P. (eds.): New Palgrave. A Dictionary of Economics 1, 457 – 460, Macmillan Press, London.

Cooter, R.D./Ulen, T. (1997): *Law and Economics*; Second Edition, Addison – Wesley, Reading/Mass.

Dari–Mattiacci, G./Garoupa N. (2007): *Least Cost Avoidance. The Tragedy of Common Safety*; George Mason Law & Economics Research Paper No. 04-27.

Dasgupta, P. (1982): *The Control of Resources;* Harvard University Press, Cambridge/Mass.

Dasgupta, A.K./Pearce, D.W. (1978): *Cost-Benefit Analysis. Theory and Practice;* Macmillan Press, London.

Demsetz, H. (1972): *When Does the Rule of Liability Matter?*; in: Journal of Legal Studies 1 (1), 13 - 28.

Driesen, D.M. (1997): *The Societal Cost of Environmental Regulation. Beyond Administrative Cost-Benefit Analysis;* in: Ecology Law Quarterly 24 (545), 601-605.

Esty, D.C. (2001): *Environmental Protection in the Information Age;* INSEAD: European Institute of Business Administration, Working Paper 2001/87/EPS/CMER.

European Commission (2001): *White Paper: European transport policy for 2010: time to decide*. http.europa.eu.int.

European Commission (2005): *Impact assessment guidelines* (with annexes). European Commission SEC(2005)791. http://ec.europa.eu/governance/impact/docs/SEC2005_791_IA_guidelines_main.pdf.

European Court of Justice (2003): *Case C-320/03 R, Commission of the European Communities v Republic of Austria*. http://eur-lex.europa.eu.

European Court of Justice (2005): *Case C-320/03 R, Commission of the European Communities v Republic of Austria*. http://eur-lex.europa.eu.

European Union (2006): *Directive 2006/38/EC of the European Parliament and of the Council of 17 May 2006 amending Directive 1999/62/EC on the charging of heavy goods vehicles for the use of certain infrastructures;* in: Official Journal of the European Union 49 (L157), 8-23.

Friedman, D. (2000): *Law's Order. What Economics Has to Do with Law and Why It Matters;* Princeton University Press, New Jersey.

Glicksman, R.L. et al. (2003): *Risk Regulation at Risk. Restoring a Pragmatic Approach;* Stanford University Press, Stanford.

HEATCO (2005): Developing Harmonised European Approaches for Transport Costing and Project Assessment; Deliverable 2, State-of-the-art in Project Assessment. http:// heatco.ier.uni-stuttgart.de.

Heiner, R.A. (1983): The *Origin of Predictable Behavior;* in: American Economic Review 73 (4), 560-595.

IFEU (2004): *Vergleichende Analyse von Energieverbrauch und CO_2-Emissionen des Straßengüterverkehrs und des kombinierten Verkehrs Straße/Schiene;* Heidelberg.

Just, R.E./Hueth, D.L./Schmitz, A. (2004): *The Welfare Economics of Public Policy*; Edward Elgar, Cheltenham, UK.

Kaplow, L. (2002): *Why measure inequality?;* NBER Working Paper, nº 9342, Cambridge, Mass.

Kaplow, L./Shavell, S. (2002): *Fairness versus Welfare;* Harvard University Press, Cambridge, Mass.

Laffont, J.J. (1987): *Externalities;* in: Eatwell, J./Milgate, M./Newman, P. (eds.): New Palgrave. A Dictionary of Economics 2, 263-265, Macmillan Press, London.

Lockwood, B. (1987): *Pareto efficiency;* in: Eatwell, J./Milgate, M./Newman, P. (eds.): New Palgrave. A Dictionary of Economics 3, 811–813, Macmillan Press, London.

McGarity, T.O. (1998): *A Cost-Benefit State;* in: Administrative Law Review 50 (1), 7-80.

Medema, S.G./Mercuro, M. (1997): *Economics and the Law. From Posner to Post-Modernism;* Princeton university Press, New Jersey.

Medema, S.G./Zerbe, R.O. (2000): *The Coase Theorem;* in: Bouckaert, B./De Geest, G. (eds.), Encyclopedia of Law and Economics 1. The History and Methodology of Law and Economics, Edward Elgar,
Cheltenham.

Miceli, T. (1997): *Economics of the Law. Torts, Contracts, Property, Litigation;* Oxford University Press, New York.

Milgrom, P./Roberts, J. (1992): *Economics, Organization and Management;* Prentice Hall, New Jersey.

Morgenstern, R.D./Landy, M.K. (1997): *Economic Analysis: Benefits, Costs, Implications*; in: Morgenstern, R.D. (1997): Economic Analyses at EPA. Assessing Regulatory Impact, 455, 461-462.

Mueller, D. (2003): *Public Choice III*; Cambridge University Press, Cambridge, Mass.

Nicita, A. et al. (2007): *Towards a Theory of Incomplete Property Rights;* Paper presented at the INSIE Conference.

Nordhaus, W.D. (2007): *A Review of the Stern Review on the Economics of Climate Change;* in: Journal of Economic Literature 45 (3), 686-702.

Ogus, A. (2006): *Costs and Cautionary Tales. Economic Insights for the Law;* Hart, Oxford.

Pearce, D.W. (1983): *Cost-Benefit Analysis;* Second Edition, Macmillan Press, London.

Perloff, J.M. (2004): *Microeconomics*; Third Edition, Addison-Wesley, Boston.

Pigou, A.C. (1952): *The Economics of Welfare;* Macmillan Press, London.

Polinsky, A. (1979): *Controlling Externalities and Protecting Entitlements. Property Right Liability Rule, and Tax-Subsidy Approaches*; in: Journal of Legal Studies 8 (1), 1 - 48.

Posner, R.A. (1981): *The Economics of Justice;* Harvard University Press, Cambridge.

Posner, R.A. (1986): *Economic Analysis of Law;* Third Edition, Little, Brown and Co, Boston.

Rose-Ackerman, S. (1989): *Dikes, dams and vicious hogs. Entitlement and efficiency in tort law;* in: Journal of Legal Studies 18 (1), 25-50.

Rose-Ackerman, S. (1992): *Rethinking the Progressive Agenda. The Reform of the American Regulatory State;* Free Press, New York.

Schelling, (2007): *Climate Change: The Uncertainties, the Certainties, and What They Imply About Action*; in: Economists' Voice. http://www.bepress.com.

Schmidtchen D. (2003): *Szenarien für eine europäische Verkehrspolitik*; in: Schüller A. (2003): Orientierungen für ordnungspolitische Reformen, Studien zur Ordnungsökonomik 29.

Shavell, S. (1987): *Economic Analysis of Accident Law;* Harvard University Press, Cambridge/Mass.

Simon, H. (1961): *Administrative Behaviour, 2nd ed.*; Macmillan, New York.

Sinn, H.-W. (2007): *The Welfare State and the Forces of Globalization;* CESifo Working Paper N° 1925, Munich.

Stavins, R. et al. (2007): *Too Good To Be True? An Examination of Three Economic Assessments of California Climate Change Policy;* Harvard University, Faculty Research Working Paper Series RWP07-16 and AEI Brookings Joint Center Paper n° 07-01.

Stephen, F.H. (1988): *The Economics of the Law;* Wheatsheaf Books, Brighton.

Stern, N. (2007): *The Economics of Climate Change*; Cambridge University Press, Cambridge, UK.

Symposium (1987): *Economists on the Bench;* in: Law and Contemporary Problems 50, 1-286.

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

Verwaltungsgericht Kassel, (2005): *Beschluss in dem Verwaltungsverfahren der Firma Junior Transporte Reinhard Gronwald gegen das Land Hessen,* Geschäftsnummer 2 G 1465/05.

Viscusi, W.K. (2002): *Fatal Tradeoffs: Public and Private Responsibilities for Risk*; Oxford University Press, New York.

Watson, J. (2002): *Strategy. An Introduction to Game Theory*; Norton, New York and London.

Weitzman, M.L. (2007): *A Review of the Stern Review on the Economics of Climate Change;* in: Journal of Economic Literature 45 (3), 703-724.

Williamson O.E. (1985): *The Economic Institutions of Capitalism*; Macmillan, New York.

Wittman, D. (2006): *Economic Foundations of Law and Organization;* Cambridge University Press, New York.

Wolf, C. (1979): *A Theory of nonmarket Failure. Framework for Implementation Analysis*; in: Journal of Law and Economics 22 (1), 107-139.

Appendix

To minimise the total costs of damage and damage prevention requires that both the residents and the transport industry should increase spending on damage prevention until the last Euro spent reduces damage by one Euro.

A little formalism can prove this assertion. Let x and y be the costs (or sacrificed rents) of the transport industry and the residents, respectively, and let D(x,y) be a strictly convex function of the parties' costs of care x and y. The expected damage costs D(x,y) can be broken down into two parts: the probability of the damage, denoted p, and the damage itself, denoted d. Thus we have $D(x,y) = p(x,y) \cdot d(x,y)$. The total social cost is the sum $D(x,y) + x + y$, minimised at x^* and y^*.

The more the transport industry spends on damage prevention, the lower the expected damage, i.e. $D_x < 0$. We assume that marginal productivity decreases, which means that additional amounts spent on damage prevention reduce expected damage at a decreasing rate, i.e. the second derivative of the damage function, D(x) with respect to x, is $D_{xx} > 0$.

Assume that the impact of care by the residents on damage reduction reveals the same pattern as shown for the transport industry. That means that $D_y < 0$ and $D_{yy} > 0$.

Assume further that an extra Euro spent on damage reduction by the residents increases the marginal productivity of an extra Euro spent by the transport industry, i.e. $D_{xy} > 0$, and precautions is substitutable:[213] if the precaution level of one party falls, the other party should increase its level of precaution.

The socially optimal care level for both the transport industry and the residents is the solution to the problem:

$\mathrm{Min}_{x,y}\ [x+y+D(x,y)]$ (1)

The first order conditions for a minimum are

$1 + D_x(x, y) = 0$ or $- D_x (x,y) = 1$ (2)

and

$1 + D_y(x, y) = 0$ or $- D_y (x,y) = 1.$ (3)

Equation (2) defines the point $x^*(y)$ which represent the optimal care level of the transport industry for any precaution level y of the residents. Similarly, equation (3) defines the optimal levels of precaution y*(x) for the residents.

[213] Dxy < 0 means joint care, Dxy > 0 implies alternative care and Dxy = 0 implies independence of the parties' care.

Obviously, the social optimum is realised at the intersection of these loci such that $x^* \equiv x^*(y^*)$ and $y^* \equiv y^*(x^*)$.[214]

If $x < x^*$, then a marginal increase in x would decrease damage by more than one euro, and if $x > x^*$, then a decrease in x would result in an increase in expected damage by less than one euro (similar reasoning can be applied for y).

A numerical example may help understand the intuition of the model.[215] For simplicity we assume $D_{xy} = 0$, i.e. independence of the parties' levels of precaution. Next, assume the damage function $D(x,y) = 100-4x^{½}-16y^{½}$.

The partial derivatives of damage D(x,y), with respect to x and y, indicating the effect of an increase in x and y on D(x,y), respectively, are $-D_x = -2x^{-1/2} <0$ and $-D_y = -8y^{-1/2} <0$ (for all values of x and y). The marginal productivity of costs spent on damage prevention is assumed to be decreasing, i.e. $D_{xx} > 0$ and $D_{yy} > 0$. In the numerical example: $D_{xx} = x^{-3/2} > 0$ and $D_{yy} = 4x^{-3/2} > 0$. The social costs are $C = D(x,y)+x+y$, and numerically $C = 100-4x^{½}-16y^{½}+x+y$.

The first - order conditions for expected cost minimisation are

$C_x = 1+D_x(x,y) = 1-2x^{-1/2} = 0$

$C_y = 1+D_y(x,y) = 1-8y^{-1/2} = 0$.

This implies $x^* = 4$ and $y^* = 64$.[216]

With $x^* = 4$ and $y^* = 64$, total social costs are

$C = 100-4*4^{½} -16*64^{½}+4+64$

$C=32$.

This damage maximises society's welfare. With $x = 0$ and $y = 0$, social costs would be 100. It is thus preferable to invest $x > 0$ and $y > 0$. With $x = 4$ and $y = 64$, social costs are minimised.

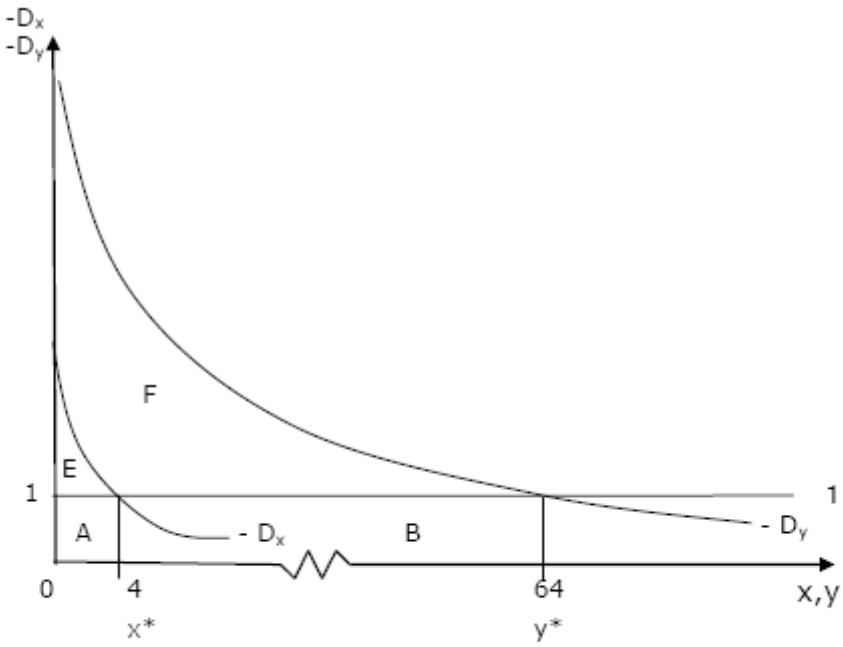
The following figure represents the bilateral prevention model:

[214] See Miceli 1997, p 18.

[215] The example is borrowed from Wittman 2006, p 133-134. .

[216] The second order conditions ensure that we have a minimum rather than a maximum or an inflection point (see Wittman 2006, p 134).

Figure A-1 Efficient avoidance of an externality



The following table lists some prominent points on these curves:

Table A-1 Example curves

x, y	- D_x	- D_y
1	2	8
4	1	4
9	2/3	8/3
16	2/4	8/4
25	2/5	8/5
36	2/6	8/6
49	2/7	8/7
64	2/8	8/8

We interpret the model as represented by figure 2-9:

Both parties contribute to the prevention of the damage: Under the assumptions made, the transport industry is not as productive in damage prevention as the residents. Efficiency thus requires that the residents should bear the greater burden.[217] The cost borne by the transport industry

[217] Of course, in the case in which the figures for y belong to the transport industry and those of x to the residents, the opposite would hold.

amounts to area A = 4, the cost borne by the residents amounts to areas A + B = 64. Thus total cost is 2A+B=68.

The transport industry prevents damage by area E+A. The residents prevent damage by area E+A+B+F. Welfare increases by the difference of the two areas: [2(E+A)+B+F]-[2A+B]=[2E+F].

Who is the cheapest cost avoider? The answer is, both are at the margin. They share the "job". Outside the optimum (x*,y*), for example with x < x* and y > y*, the transport industry is the cheapest cost avoider: it should increase x to x* and the residents should decrease y to y*.

With damage of, for example, D(x,y) = D(0, 0) = 67, i.e. a damage function $D(x,y) = 67-4x^{1/2}-16y^{1/2}$, neither the transport industry nor the residents should take care. The costs in the optimum would amount to 68, but the benefit would be a reduction of the damage from 67 down to zero – clearly a benefit that is not worth the cost of realising it. This conclusion holds for all values D(x,y) = D(0,0) < 68.

Note that the model can also be applied to cases in which the prices of both damage reduction inputs to are ≠ 1. Let a and b denote the price of the input respectively used by the transport industry and the residents to reduce damage. Here, the optimum requires $-D_x = a$ and $-D_y = b$.

The benefit from the double investment is depicted by the sum of the areas under the $D_{x,y}$ curves. The measures should be taken if the transaction costs that they imply to not exceed the benefit thus calculated.

Modifying the assumptions that the model and the figure are based on can lead to a solution in which only one side should take action reducing the damage.
Assume that the position of curve $-D_x$ is such that there is no intersection with the horizontal line 1. Now, the transport industry should do nothing, since the net benefit to society would be negative (marginal costs of 1 are always greater than the marginal benefit $-D_x$). The residents are the cheapest cost avoiders. The optimum (x* = 0, y* = 64) is a corner solution.[218] The remaining damage is 36. This case is generally called the unilateral case, because only one of the parties should take measures to reduce the externality. Note however that unilateral causation is a *result* of bilateral cost minimisation, and not a different type of causation.

The case of unilateral causation, i.e. the case where one of the D_i curves (with i = x,y) does not cut the marginal benefit curve is exactly the case described in the "but for" test. For example, a patient could have prevented suffering form a doctor's professional error by refusing an operation. However, a lawyer using the "but for" test would argue that it is not reasonable to refuse a life saving operation because of the risk of a mistake. We arrive at the same conclusion with the cheapest cost avoider method: the marginal cost of the prevention measure for the inhabitant would be so

[218] Note that this is different from the unilateral care case in which $D_x = 0$ or $D_y = 0$.

high that it can never be equal the marginal benefit of avoiding the damage. $-D_y$ would move far to the left and result in $y^* = 0$. Thus, prevention measures should be taken only by the airline. This implies that the property right should be given to the inhabitant. Only now that these two steps have been taken can we apply the polluter pays principle in accordance with the cheapest cost avoider principle: if the airline infringes the inhabitant's right, it should be liable.