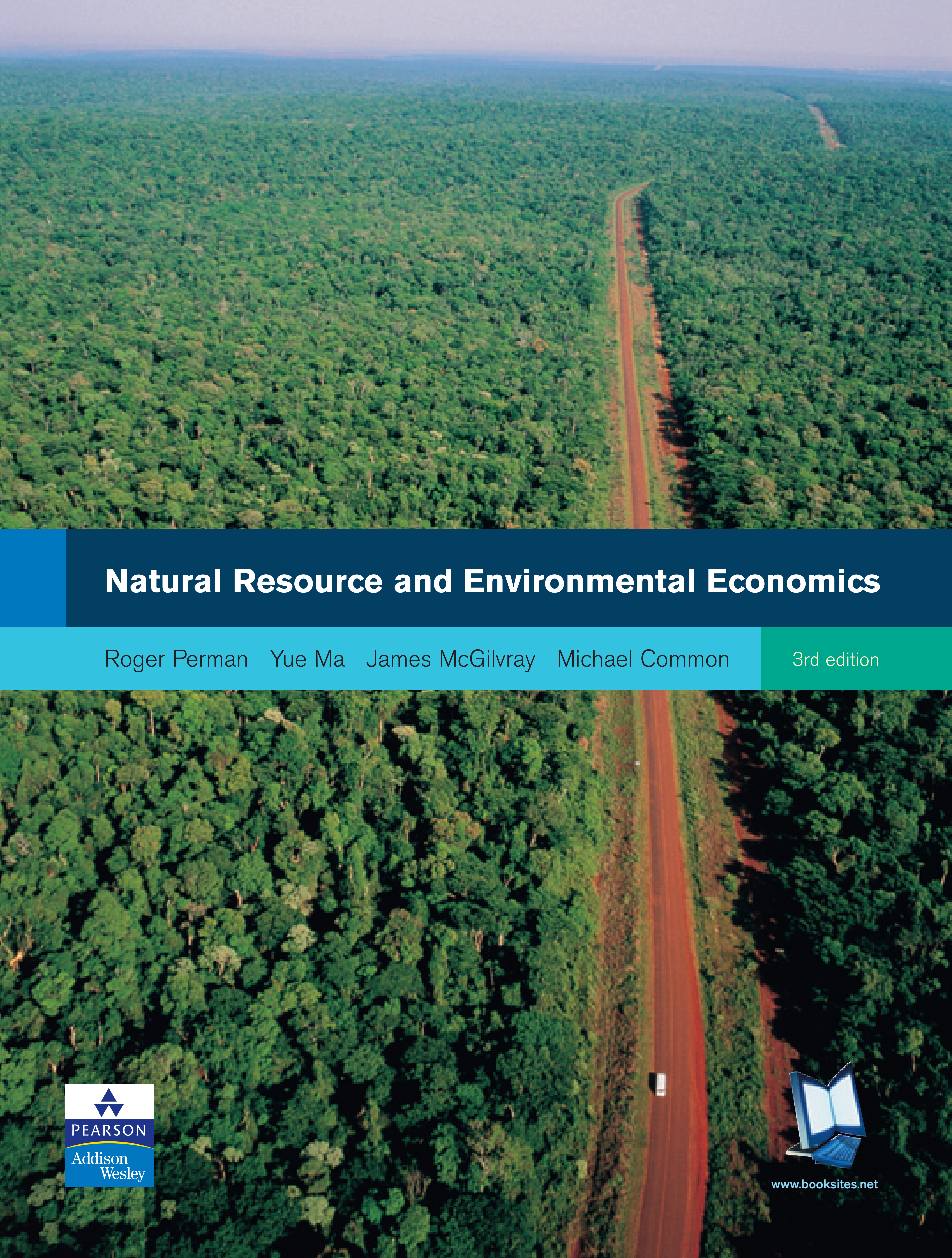

# Natural Resource and Environmental Economics

Roger Perman Yue Ma James McGilvray Michael Common

3rd edition

PEARSON
Addison Wesley

www.booksites.net

# Natural Resource and Environmental Economics

Third Edition

Roger Perman
Yue Ma
James McGilvray
Michael Common

PEARSON
Addison Wesley

Harlow, England • London • New York • Boston • San Francisco • Toronto
Sydney • Tokyo • Singapore • Hong Kong • Seoul • Taipei • New Delhi
Cape Town • Madrid • Mexico City • Amsterdam • Munich • Paris • Milan

**Pearson Education Limited**
Edinburgh Gate
Harlow
Essex CM20 2JE

and Associated Companies throughout the world

*Visit us on the World Wide Web at:*
www.pearsoneduc.com

First published 1996 Longman Group Limited
Second edition 1999 Addison Wesley Longman Limited
**Third edition 2003 Pearson Education Limited**



ISBN 0273655590

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

**Library of Congress Cataloging-in-Publication Data**
Natural resource and environmental economics / Roger Perman . . . [et al.].—3rd ed.
p. cm.
Rev. ed. of: Natural resource and environmental economics / Roger Perman, Yue Ma, James McGilvray. 1996.
Includes bibliographical references and index.
ISBN 0-273-65559-0 (pbk.)
1. Environmental economics. 2. Natural resources—Management. 3. Sustainable development. I. Perman, Roger, 1949– Natural resource and environmental economics.

HC79.E5 P446 2003
333.7—dc21 2002042567

10 9 8 7 6 5 4 3 2 1
06 05 04 03

Typeset in 9.75/12pt Times by 35
Printed and bound by Ashford Colour Press Ltd., Gosport

# Contents

# PART II Environmental pollution

# CHAPTER 6 Pollution control: targets

The use of coal was prohibited in London in 1273, and at least one person was put to death for this offense around 1300. Why did it take economists so long to recognize and analyze the problem?

**Fisher (1981), p. 164**

**Learning objectives**

*At the end of this chapter, the reader should be able to*

- understand the concept of a pollution target
- appreciate that many different criteria can be used to determine pollution targets
- understand that alternative policy objectives usually imply different pollution targets
- understand how in principle targets may be constructed using an economic efficiency criterion
- understand the difference between flow and stock pollutants
- analyse efficient levels of flow pollutants and stock pollutants
- appreciate the importance of the degree of mixing of a pollutant stock
- recognise and understand the role of spatial differentiation for emissions targets

## Introduction

In thinking about pollution policy, the economist is interested in two major questions. How much pollution should there be? And, given that some target level has been chosen, what is the best method of achieving that level? In this chapter we deal with the first of these questions; the second is addressed in the next chapter.

How much pollution there should be depends on the objective that is being sought. Many economists regard economic optimality as the ideal objective. This requires that resources should be allocated so as to maximise social welfare. Associated with that allocation will be the optimal level of pollution. However, the information required to establish the optimal pollution level is likely to be unobtainable, and so that criterion is not feasible in practice.[1] As a result, the weaker yardstick of economic efficiency is often proposed as a way of setting pollution targets.[2]

[1] In Chapter 5 we showed that identification of an optimal allocation requires, among other things, knowledge of an appropriate social welfare function, and of production technologies and individual preferences throughout the whole economy. Moreover, even if such an allocation could be identified, attaining it might involve substantial redistributions of wealth.

[2] If you are unclear about the difference between optimality and efficiency it might be sensible to look again at Chapter 5. It is worth recalling that the efficiency criterion has an ethical underpinning that not all would subscribe to, as it implicitly accepts the prevailing distribution of wealth. We established in Chapter 5 that efficient outcomes are not necessarily optimal ones. Moreover, moving from an inefficient to an efficient outcome does not necessarily lead to an improvement in social well-being.

The use of efficiency as a way of thinking about how much pollution there should be dates back to the work of Pigou, and arose from his development of the concept of externalities (Pigou, 1920). Subsequently, after the theory of externalities had been extended and developed, it became the main organising principle used by economists when analysing pollution problems.

In practice, much of the work done by economists within an externalities framework has used a partial equilibrium perspective, looking at a single activity (and its associated pollution) in isolation from the rest of the system in which the activity is embedded. There is, of course, no reason why externalities cannot be viewed in a general equilibrium framework, and some of the seminal works in environmental economics have done so. (See, for example, Baumol and Oates, 1988, and Cornes and Sandler, 1996.)

This raises the question of what we mean by the 'system' in which pollution-generating activities are embedded. The development of environmental economics and of ecological economics as distinct disciplines led some writers to take a comprehensive view of that system. This involved bringing the material and biological subsystems into the picture, and taking account of the constraints on economy–environment interactions.

One step in this direction came with incorporating natural resources into economic growth models. Then pollution can be associated with resource extraction and use, and best levels of pollution emerge in the solution to the optimal growth problem. Pollution problems are thereby given a firm material grounding and policies concerning pollution levels and natural resource uses are linked. Much of the work done in this area has been abstract, at a high level of aggregation, and is technically difficult. Nevertheless, we feel it is of sufficient importance to warrant study, and have devoted Chapter 16 to it.[3]

There have been more ambitious attempts to use the material balance principle (which was explained in Chapter 2) as a vehicle for investigating pollution problems. These try to systematically model interactions between the economy and the environment. Production and consumption activities draw upon materials and energy from the environment. Residuals from economic processes are returned to various environmental receptors (air, soils, biota and water systems). There may be significant delays in the timing of residual flows from and to the environment. In a growing economy, a significant part of the materials taken from the environment is assembled in long-lasting structures, such as roads, buildings and machines. Thus flows back to the natural environment may be substantially less than extraction from it over some interval of time. However, in the long run the materials balance principle points to equality between outflows and inflows. If we defined the environment broadly (to include human-made structures as well as the natural environment) the equality would hold perfectly at all times. While the masses of flows to and from the environment are identical, the return flows are in different physical forms and to different places from those of the original, extracted materials. A full development of this approach goes beyond what we are able to cover in this book, and so we do not discuss it further (beyond pointing you to some additional reading).

Economic efficiency is one way of thinking about pollution targets, but it is certainly not the only way. For example, we might adopt sustainability as the policy objective, or as a constraint that must be satisfied in pursuing other objectives. Then pollution levels (or trajectories of those through time) would be assessed in terms of whether they are compatible with sustainable development. Optimal growth models with natural resources, and the materials balance approach just outlined, lend themselves well to developing pollution targets using a sustainability criterion. We will show later (in Chapters 14, 16 and 19) that efficiency and sustainability criteria do not usually lead to similar recommendations about pollution targets.

Pollution targets may be, and in practice often are, determined on grounds other than economic efficiency or sustainability. They may be based on what risk to health is deemed reasonable, or on what is acceptable to public opinion. They may be based on what is politically feasible. In outlining the political economy of regulation in Chapter 8, we demonstrate

[3] Our reason for placing this material so late in the text is pedagogical. The treatment is technically difficult, and is best dealt with after first developing the relevant tools in Chapters 14 and 15.

that policy is influenced, sometimes very strongly, by the interplay of pressure groups and sectional interests. Moreover, in a world in which the perceived importance of international or global pollution problems is increasing, policy makers find themselves setting targets within a network of obligations and pressures from various national governments and coalitions. Pollution policy making within this international milieu is the subject of Chapter 10.

In the final analysis, pollution targets are rarely, if ever, set entirely on purely economic grounds. Standards setting is usually a matter of trying to attain multiple objectives within a complex institutional environment. Nevertheless, the principal objective of this chapter is to explain what economics has to say about determining pollution targets.

## 6.1 Modelling pollution mechanisms

Before going further, it will be instructive to develop a framework for thinking about how pollution emissions and stocks are linked, and how these relate to any induced damage. An example is used to help fix ideas. Box 6.1 outlines the stages, and some characteristics, of the oil fuel cycle. It illustrates the material and energy flows associated with the extraction and transportation of oil, its refining and burning for energy generation, and the subsequent transportation and chemical changes of the residuals in this process.

The contents of Box 6.1 lead one to consider several important ideas that will be developed in this

**Box 6.1 The oil-to-electricity fuel cycle**

Figure 6.1 describes the process steps of the oil-to-electricity fuel cycle. At each of these steps, some material transformation occurs, with potential for environmental, health and other damage.

The task given to the ExternE research team was, among other things, to estimate the external effects of power generation in Europe. A standard methodology framework – called the Impact Pathway Methodology – was devised for this task. The stages of the impact pathway are shown on the left-hand side of Figure 6.2. Each form of pollutant emission associated with each fuel cycle was investigated in this standard framework. One example of this, for one pollutant and one kind of impact of that pollutant, is shown on the right-hand side of Figure 6.2; coal use results in sulphur dioxide emissions, which contribute to acidification of air, ground and water systems.

An indication of the pervasiveness of impacts and forms of damage is shown in Table 6.1, which lists the major categories of damages arising from the oil-to-electricity fuel cycle. In fact, ExternE identified 82 sub-categories of the items listed in Table 6.1. It attempted to measure each of these 82 impacts for typical oil-fired power stations in Europe, and place a monetary value on each sub-category.

ExternE (1995) compiled a detailed summary of its estimates of the annual total damage impacts of one example of an oil fuel cycle

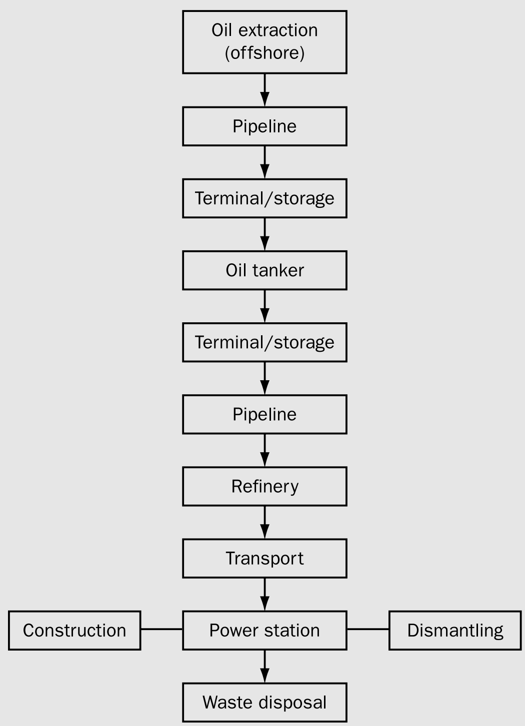


*Figure 6.1* Process steps of the oil-to-electricity fuel cycle
*Source*: ExternE (1995), figure 3.1, p. 30

**Box 6.1 *continued***

| Impact pathway stage | Example |
|---|---|
| Resource use | Coal fired power generation |
| Emission | $SO_2$ |
| Pollutant transport | Charge in atmospheric $SO_2$ concentration |
| Process | Fresh water acidification |
| Impact | Loss of fish stock |
| Damage | Loss of economic net benefit |

*Figure 6.2* The impact pathways methodology and one example
*Source*: Adapted from ExternE (1995), figure 1, p. iii

*Table 6.1* Major categories of damage arising from the oil-to-electricity fuel cycle

| Damage category |
|---|
| Oil spills on marine ecosystems |
| Public health: |
| Acute mortality |
| Acute morbidity |
| Ozone |
| Chronic morbidity |
| Occupational health |
| Agriculture |
| Forests |
| Materials |
| Noise |
| Global warming |

*Source*: Adapted from ExternE (1995)

*Table 6.2* ExternE estimates of the damage impacts of two power stations

| Category | Total |
|---|---|
| All Other | 0.7826 |
| Death | 18.4362 |
| Other human health | 4.30331 |
| Grand Total | 23.5221 |
| **Category** | **Total** |
| All Other | 3.33% |
| Death | 78.38% |
| Other human health | 18.29% |
| Grand Total | 100.00% |

*Source*: ExternE (1995), as compiled in the Excel workbook *ExternE.xls*. Full definitions of units and variables are given there

(the Lauffen power plant, Germany, employing a peak-load gas turbine plant operated with light fuel-oil and a base load combined cycle plant using heavy fuel-oil). Given its size – about 100 individual categories of impact are identified – we have chosen to present these findings separately, in the Excel workbook *ExternE.xls* in the *Additional Materials* for Chapter 6. For convenience, the Excel table also contains damage estimates for one example of a natural gas fuel cycle (the West Burton power station, a 652 MW Combined Cycle Gas Turbine Plant in the East Midlands of the UK). Data is shown in currency units of mecu (milli-ecu, or 0.001 ecu; at 1992 exchange rates \$US 1.25 ≈ 1 ecu).

It is useful to study this material for two reasons. First, it shows the huge breadth of types of pollution impact, and the great attention to detail given in well-funded research studies. Second, as Table 6.2 demonstrates, estimates of pollution damages are often dominated by values attributed to human mortality impacts. The data in Table 6.2 shows the sums of annual combined impacts of the two example power stations (expressed in units of mecu/kWh) for three very broad impact categories, and then in terms of percentages of total impact. Impacts on human mortality constitute over 78% of the identified and quantified impacts. It should be pointed out that the figures shown were arrived at when the ExternE analysis was incomplete; in particular, little attention had been given to greenhouse warming impacts of $CO_2$ emissions. Nevertheless, the figures here illustrate one property that is common to many impact studies: human health impacts account for a large proportion of the total damage values. Given that valuation of human life is by no means straightforward (as we shall indicate in Chapter 12), estimates produced by valuation studies can often be highly contentious.

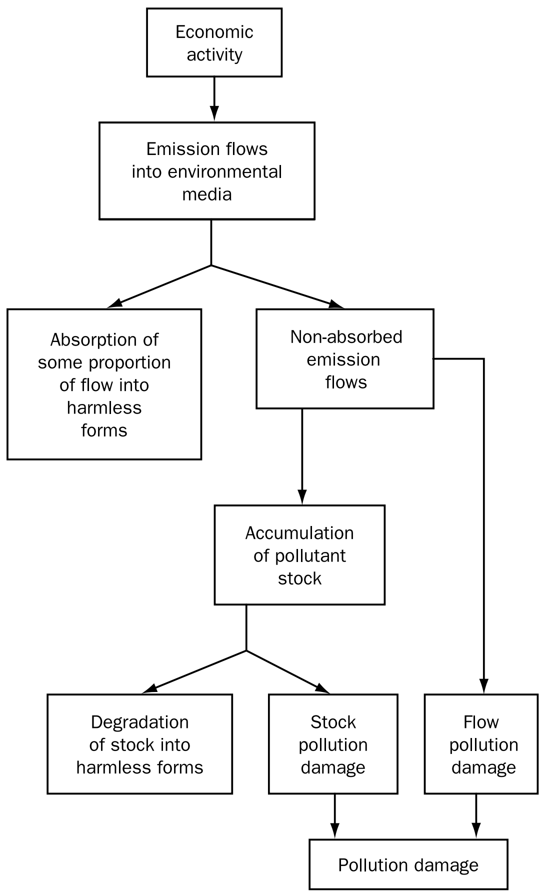


*Figure 6.3* Economic activity, residual flows and environmental damage

chapter. In particular, residual flows impose loads upon environmental systems. The extent to which these waste loads generate impacts that are associated with subsequent damage depends upon several things, including:

- the assimilative (or absorptive) capacity of the receptor environmental media;
- the existing loads on the receptor environmental media;
- the location of the environmental receptor media, and so the number of people living there and the characteristics of the affected ecosystems;
- tastes and preferences of affected people.

Figure 6.3 illustrates some of these ideas schematically for pollution problems in general. Some proportion of the emission flows from economic activity is quickly absorbed and transformed by environmental media into harmless forms. The assimilative capacity of the environment will in many circumstances be sufficient to absorb and transform into harmless forms some amount of wastes. However, carrying capacities will often be insufficient to deal with all wastes in this way, and in extreme cases carrying capacities will become zero when burdens become excessive. Furthermore, physical and chemical processes take time to operate. Some greenhouse gases, for example, require decades to be fully absorbed in water systems or chemically changed into non-warming substances (see Table 6.3).

This implies that some proportion of wastes will, in any time interval, remain unabsorbed or untransformed. These may cause damage at the time of their emission, and may also, by accumulating as pollutant stocks, cause additional future damage. Stocks of pollutants will usually decay into harmless forms but the rate of decay is often very slow. The half-lives of some radioactive substances are thousands of years, and for some highly persistent pollutants, such as the heavy metals, the rate of decay is approximately zero.

## 6.2 Pollution flows, pollution stocks and pollution damage

Pollution can be classified in terms of its damage mechanism. This has important implications for how pollution targets are set and for the way in which pollution is most appropriately controlled. The distinction here concerns whether damage arises from the flow of the pollutant (that is, the rate of emissions) or from the stock (or concentration rate) of pollution in the relevant environmental medium. We define the following two classes of pollution: flow-damage pollution and stock-damage pollution (but recognise that there may also be mixed cases).

*Flow-damage pollution* occurs when damage results only from the flow of residuals: that is, the rate at which they are being discharged into the environmental system. This corresponds to the right-hand side branch in Figure 6.3. By definition, for pure cases of flow-damage pollution, the damage will instantaneously drop to zero if the emissions flow becomes zero. This can only be exactly true when the pollutant exists in an energy form such as noise or light so that when the energy emission is terminated no residuals remain in existence. However, this characterisation of damages may be approximately true in a wider variety of cases, particularly when the residuals have very short lifespans before being transformed into benign forms.

*Stock-damage pollution* describes the case in which damages depend only on the stock of the pollutant in the relevant environmental system at any point in time. This corresponds to the central branch in Figure 6.3. For a stock of the pollutant to accumulate, it is necessary that the residuals have a positive lifespan and that emissions are being produced at a rate which exceeds the assimilative capacity of the environment. An extreme case is that in which the assimilative capacity is zero, as seems to be approximately the case for some synthetic chemicals and a number of heavy metals. (The left-hand branch in Figure 6.3 does not then exist.) Metals such as mercury or lead accumulate in soils, aquifers and biological stocks, and subsequently in the human body, causing major damage to human health. Persistent synthetic chemicals, such as PCBs (polychlorinated biphenyls), DDT and dioxins, have similar cycles and effects. Rubbish which cannot biodegrade is another case. So are, for all practical purposes, strongly radioactive elements such as plutonium with extremely long radiation half-lives.

Most important pollution problems have the attribute of a stock-damage pollution effect being present. The most prominent are those which affect human health and life expectancy. But the phenomenon is more pervasive than this. Pollution stocks are harmful to built structures (buildings, works of art and so on) and they may adversely affect production potential, particularly in agriculture. Stock pollution levels influence plant and timber growth, and the size of marine animal populations. Less direct effects operate through damages to environmental resources and ecological systems. There is another way in which stock effects operate. The assimilative capacity of the environment often depends on the emissions load to which relevant environmental media are exposed. This is particularly true when the natural cleaning mechanism operates biologically. In water systems, for example, bacterial decomposition of pollutants is the principal cleaning agency. But where critical loads are exceeded, this biological conversion process breaks down, and the water system can effectively become dead. Its assimilative capacity has fallen to zero.

Mixed cases, where pollution damage arises from both flow and stock effects, also exist. Waste emissions into water systems are sometimes modelled as mixed stock-flow pollutants. So too are damages arising from the emissions of compounds of carbon, sulphur and nitrogen. However, in these mixed cases, it may often be preferable to view the problem as one of a pure stock pollutant.

Using $M$ to denote the pollution flow, $A$ to denote the pollution stock and $D$ to denote pollution damage, we therefore have two variants of damage function:

$$\text{Flow-damage pollution:}\quad D = D(M) \tag{6.1a}$$

$$\text{Stock-damage pollution:}\quad D = D(A) \tag{6.1b}$$

For simplicity of notation, we shall from this point on call these ‘flow pollution’ and ‘stock pollution’.

## 6.3 The efficient level of pollution

We now investigate how pollution targets can be set using an efficiency criterion. Given that pollution is harmful, some would argue that only a zero level of pollution is desirable. But, as we shall see, pollution can also be beneficial. Therefore, zero pollution is not economically efficient except in particular special circumstances. In what sense is pollution beneficial? One answer comes from the fact that producing some goods and services that we do find useful may not be possible without generating some pollution, even if only a small amount. More generally, goods might only be producible in non-polluting ways at large additional expense. Thus, relaxing a pollution abatement constraint allows the production of goods that could not otherwise have been made, or to produce those goods at less direct cost. This is the sense in which pollution could be described as beneficial.

With both benefits and costs, economic decisions about the appropriate level of pollution involve the evaluation of a trade-off. Thinking about pollution as an externality arising from production or consumption activities makes this trade-off clear. The efficient level of an externality is not, in general, zero as the marginal costs of reducing the external effect will, beyond a certain point, exceed its marginal benefits.

The discussion of efficient pollution targets which follows is divided into several parts. In the first two (Sections 6.4 and 6.5) a static modelling framework is used to study efficient emissions of a flow pollutant. This explains the key principles involved in dealing with the trade-off. We next, in Section 6.6, investigate the more common – and important – case of stock-damage pollution. Two variants of stock damage are considered. Sections 6.7 and 6.8 deal with those stock pollutants for which the location of the emission source matters as far as the pollutant stock, and so the extent of damages, is concerned. Our emphasis here will be on the *spatial* dimension of pollution problems. Section 6.9 focuses on the *time* dimension of pollution problems. It studies long-lived pollutants, such as greenhouse gases, which can accumulate over time. At this stage, our treatment of persistent stock pollutants will be relatively simple. Later, in Chapter 16, a richer dynamic modelling framework will be used to identify emission targets where pollution is modelled as arising from the depletion of natural resources.

## 6.4 A static model of efficient flow pollution

A simple static model – one in which time plays no role – can be used to identify the efficient level of a flow pollutant. In this model, emissions have both benefits and costs. In common with much of the pollution literature, the costs of emissions are called damages. Using a concept introduced in Chapter 5, these damages can be thought of as a negative (adverse) externality. Production entails joint products: the intended good or service, *and* the associated pollutant emissions. In an unregulated economic environment, the costs associated with production of the intended good or service are paid by the producer, and so are internalised. But the costs of pollution damage are not met by the firm, are not taken into account in its decisions, and so are externalities. Moreover, in many cases of interest to us, it is also the case that the externality in question is what Chapter 5 called a public bad (as opposed to a private bad), in that once it has been generated, no one can be excluded from suffering its adverse effects.

For simplicity, we suppose that damage is independent of the time or source of the emissions and that emissions have no effect outside the economy being studied. We shall relax these two assumptions later, the first in Section 6.6 and in Chapter 7, and the second in Chapter 10.

An efficient level of emissions is one that maximises the net benefits from pollution, where net benefits are defined as pollution benefits minus pollution costs (or damages). The level of emissions at which net benefits are maximised is equivalent to the outcome that would prevail if the pollution externality were fully internalised. Therefore, the identification of the efficient level of an adverse externality in Figure 5.14, and the discussion surrounding it, is apposite in this case with an appropriate change of context.

In the case of flow pollution, damage ($D$) is dependent only on the magnitude of the emissions flow ($M$), so the damage function can be specified as

$$D = D(M) \tag{6.2}$$

Matters are a little less obvious with regard to the benefits of pollution. Let us expand a little on the earlier remarks we made about interpreting these benefits. Suppose for the sake of argument that firms were required to produce their intended final output without generating any pollution. This would, in general, be extremely costly (and perhaps even impossible in that limiting case). Now consider what will happen if that requirement is gradually relaxed. As the amount of allowable emissions rises, firms can increasingly avoid the pollution abatement costs that would otherwise be incurred. Therefore, firms make cost savings (and so profit increases) if they are allowed to generate emissions in producing their goods. The larger is the amount of emissions generated (for any given level of goods output), the greater will be those cost savings.

A sharper, but equivalent, interpretation of the benefits function runs as follows. Consider a representative firm. For any particular level of output it chooses to make, there will be an unconstrained emissions level that would arise from the cost-minimising method of production. If it were required to reduce emissions below that unconstrained level, and did so in the profit-maximising way, the total of production and control costs would exceed the total production costs in the unconstrained situation. So there are additional costs associated with emissions reduction. Equivalently, there are savings (or benefits) associated with emissions increases. It is these cost savings that we regard as the benefits of pollution.

Symbolically, we can represent this relationship by the function

$$B = B(M) \tag{6.3}$$

in which $B$ denotes the benefits from emissions.[4] The social net benefits (NB) from a given level of emissions are defined by

$$\text{NB} = B(M) - D(M) \tag{6.4}$$

It will be convenient to work with marginal, rather than total, functions. Thus $\mathrm{d}B/\mathrm{d}M$ (or $B'(M)$ in an alternative notation) is the marginal benefit of pollution and $\mathrm{d}D/\mathrm{d}M$ (or $D'(M)$) is the marginal damage of pollution. Economists often assume that the total and marginal damage and benefit functions have the general forms shown in Figure 6.4. Total damage is thought to rise at an increasing rate with the size of the pollution flow, and so the marginal damage will be increasing in $M$. In contrast, total benefits will rise at a decreasing rate as emissions increase (because per-unit pollution abatement costs will be more expensive at greater levels of emissions *reduction*). Therefore, the marginal benefit of pollution would fall as pollution flows increase.

It is important to understand that damage or benefit functions (or both) will not necessarily have these general shapes. For some kinds of pollutants, in particular circumstances, the functions can have very different properties, as our discussions in Section 6.11 will illustrate. There is also an issue about whether the benefit function correctly describes the social benefits of emissions. Under some circumstances, emissions abatement can generate a so-called double dividend. If it does, the marginal benefit function as defined in this chapter will overstate the true value of emissions benefits. For some explanation of the double dividend idea, see Box 6.3. Nevertheless, except where it is stated otherwise, our presentation will assume that the general shapes shown in Figure 6.4 are valid.

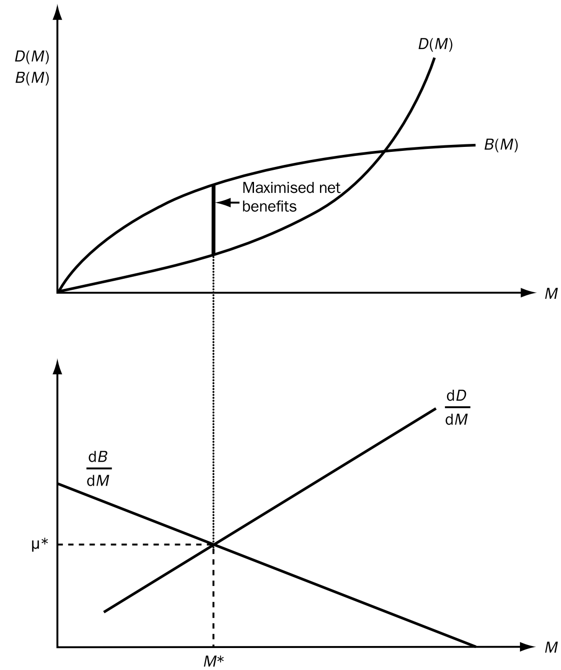


*Figure 6.4* Total and marginal damage and benefit functions, and the efficient level of flow pollution emissions

To maximise the net benefits of economic activity, we require that the pollution flow, $M$, be chosen so that

$$\frac{\mathrm{dNB}(M)}{\mathrm{d}M} = \frac{\mathrm{d}B(M)}{\mathrm{d}M} - \frac{\mathrm{d}D(M)}{\mathrm{d}M} = 0 \tag{6.5a}$$

or, equivalently, that

[4] Given our interpretation of the emissions benefit function (which involves optimised emissions abatement costs at any level of emissions below the unconstrained level), it will not be an easy matter to quantify this relationship numerically. However, there are various ways in which emissions abatement cost functions can be estimated, as you will see in Section 6.12. And with a suitable change of label (again, as we shall see later) abatement cost functions are identical to the benefit function we are referring to here.

$$\frac{\mathrm{d}B(M)}{\mathrm{d}M} = \frac{\mathrm{d}D(M)}{\mathrm{d}M} \qquad (6.5b)$$

which states that the net benefits of pollution can be maximised only where the marginal benefits of pollution equal the marginal damage of pollution.[5] This is a special case of the efficiency condition for an externality stated in Chapter 5.

The efficient level of pollution is $M^*$ (see Figure 6.4 again). If pollution is less than $M^*$ the marginal benefits of pollution are greater than the marginal damage from pollution, so higher pollution will yield additional net benefits. Conversely, if pollution is greater than $M^*$, the marginal benefits of pollution are less than the marginal damage from pollution, so less pollution will yield more net benefits.

The value of marginal damage and marginal benefit functions at their intersection is labelled $\mu^*$ in Figure 6.4. We can think of this as the equilibrium 'price' of pollution. This price has a particular significance in terms of an efficient rate of emissions tax or subsidy, as we shall discover in the following chapter. However, as there is no market for pollution, $\mu^*$ is a hypothetical or shadow price rather than one which is actually revealed in market transactions. More specifically, a shadow price emerges as part of the solution to an optimisation problem (in this case the problem of choosing $M$ to maximise net benefits). We could also describe $\mu^*$ as the shadow price of the pollution externality. If a market were, somehow or other, to exist for the pollutant itself (thereby internalising the externality) so that firms had to purchase rights to emit units of the pollutant, $\mu^*$ would be the efficient market price. Indeed, Chapter 7 will demonstrate that $\mu^*$ is the equilibrium price of tradable permits if an amount $M^*$ of such permits were to be issued.

Another interpretation of the emissions efficiency condition (equation 6.5b) is obtained by inspection of Figure 6.5. The efficient level of pollution is

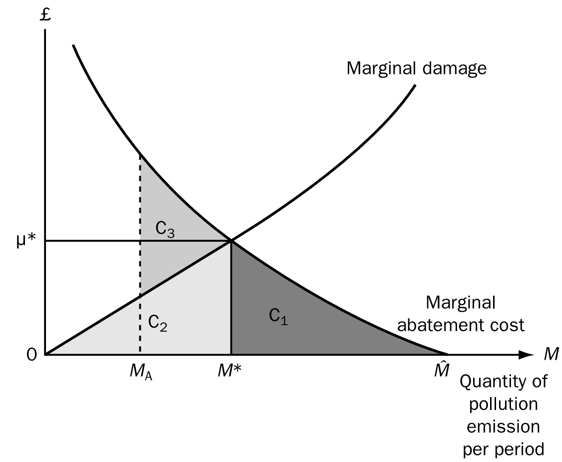


*Figure 6.5* The economically efficient level of pollution minimises the sum of abatement and damage costs

the one that minimises the sum of total abatement costs plus total damage costs. Notice that in the diagram we have relabelled the curve previously called marginal benefit as marginal abatement cost. The logic here should be clear given our earlier discussion about the derivation of the benefits of pollution function.[6]

To confirm this cost-minimising result, note that at the efficient pollution level, $M^*$, the sum of total damage costs (the area $C_2$) and total abatement costs (the area $C_1$) is $C_2 + C_1$. Any other level of emissions yields higher total costs. If too little pollution is produced (or too much abatement is undertaken) with a pollution flow restricted to $M_A$, it can be deduced that total costs rise to $C_1 + C_2 + C_3$, so $C_3$ is the efficiency loss arising from the excessive abatement. If you cannot see how this conclusion is reached, look now at Problem 2 at the end of this chapter. You should also convince yourself that too much pollution (too little abatement) results in higher costs than $C_1 + C_2$.

[5] This marginal equality applies when the optimum is at an interior point (does not fall at either extreme of the domain of the function). A sufficient second-order condition for this solution to be a net benefit maximum is that $\mathrm{d}^2NB/\mathrm{d}M^2 = \mathrm{d}^2B/\mathrm{d}M^2 - \mathrm{d}^2D/\mathrm{d}M^2 < 0$. Both an interior solution and the second-order condition are satisfied given the slopes and relative positions of the functions assumed in the text and shown in Figure 6.4 (see Chiang, 1984).

[6] The reinterpretation follows from the fact that reducing emissions incurs abatement costs. By construction, these (marginal) abatement costs are equal to the marginal benefits that will be lost if emissions fall. So, in Figure 6.5, if we start at the unconstrained emissions level, denoted as $\hat{M}$ in the diagram, then moving leftwards towards the origin corresponds to rising amounts of pollution abatement. Marginal abatement costs are low at small levels of abatement, and rise at an increasing rate as the abatement level becomes larger.

**Box 6.2 Efficient solution for a flow pollutant: a numerical example**

Suppose that the total damage and total benefits functions have the following particular forms:

$$D = M^2 \text{ for } M \geq 0$$

$$B = \begin{cases} 96M - 0.2M^2 & \text{for} \quad 0 \leq M \leq 240 \\ 11\,520 & \text{for} \quad M > 240 \end{cases}$$

**What is $M^*$?**

If $M$ is less than or equal to 240, then we have $B = 96M - 0.2M^2$ and so $\mathrm{d}B/\mathrm{d}M = 96 - 0.4M$. For any positive value of $M$ we also have $D = M^2$ which implies that $\mathrm{d}D/\mathrm{d}M = 2M$, Now setting $\mathrm{d}B/\mathrm{d}M = \mathrm{d}D/\mathrm{d}M$ we obtain $96 - 0.4M = 2M$, implying that $M^* = 40$.

Substituting $M^* = 40$ into the benefit and damage functions gives us the result that $B^* = 3520$ and $D^* = 1600$, and so maximised total net benefits (NB*) are 1920. Note also that at $M^*$ marginal benefit and marginal damage are equalised at 80 and so the shadow price $\mu^*$ – the value of value of marginal pollution damage at the efficient outcome – is 80.

You should now verify that $M^* = 40$ is a global optimum. This can be done by sketching the respective marginal functions and showing that net benefits are necessarily lower than 1920 for any (positive) level of $M$ other than 40.

**Additional materials**

It can be useful to write a spreadsheet to do the kind of calculations we have just gone through. Moreover, if the spreadsheet is constructed appropriately, it can also serve as a template by means of which similar calculations can be quickly implemented as required. Alternatively, we could use such a spreadsheet to carry out comparative statics; that is, to see how the solution changes as parameter values are altered.

We have provided an Excel workbook *Targets examples.xls* that can be used in these ways in the *Additional Materials* available on the textbook's web pages. That spreadsheet also shows how one of Excel's tools – 'Solver' – can be used to obtain the efficient level of $M$ directly, by finding the level of $M$ which maximises the net benefit function $\mathrm{NB} = B - D = (96M - 0.2M^2) - (M^2)$.

It can also be deduced from Figures 6.4 and 6.5 that the efficient level of pollution will not, in general, be zero. (By implication, the efficient level of pollution abatement will not, in general, correspond to complete elimination of pollution.) Problem 1 examines this matter.

We round off this section with a simple numerical example, given in Box 6.2. Functional forms used in the example are consistent with the general forms of marginal benefit and marginal damage functions shown in Figure 6.4. We solve for the values of $M^*$, $B^*$, $D^*$ and $\mu^*$ for one set of parameter values. Also provided, in the *Additional Materials* that are linked to this text, is an Excel spreadsheet (*Targets examples.xls*) that reproduces these calculations. The Excel workbook is set up so that comparative statics analysis can be done easily by the reader. That is, the effects on $M^*$, $B^*$, $D^*$ and $\mu^*$ of changes in parameter values from those used in Box 6.2 can be obtained.

## 6.5 Modified efficiency targets

Our notion of efficiency to this point has been a comprehensive one; it involves maximising the difference between *all* the benefits of pollution and *all* the costs of pollution. But, sometimes, one particular kind of pollution cost (or damage) is regarded as being of such importance that pollution costs should be defined in terms of that cost alone. In this case we can imagine a revised or modified efficiency criterion in which the goal is to maximise the difference between all the benefits of pollution and this particular kind of pollution damage.

Policy makers sometimes appear to treat risks to human health in this way. So let us assume policy makers operate by making risks to human health the only damage that counts (in setting targets). How would this affect pollution targets? The answer depends on the relationship between emissions and

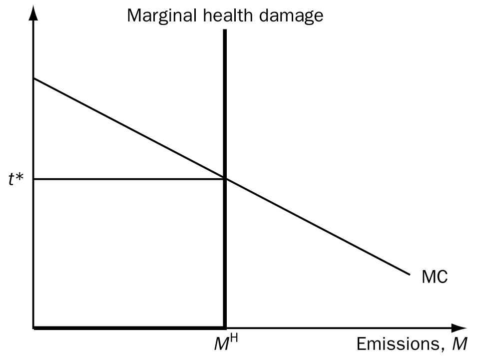


*Figure 6.6* Setting targets according to an absolute health criterion

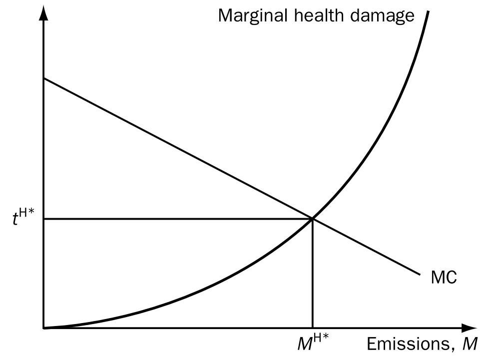


*Figure 6.7* A 'modified efficiency-based' health standard

health risks. One possible relationship is that illustrated by the ⌋-shaped relationship in Figure 6.6. Total (and marginal) health risks are zero below the threshold, but at the threshold itself risks to human health become intolerably large. It is easy to see that the value of marginal benefits is irrelevant here. A modified efficiency criterion would, in effect, lead to the emissions target being set by the damage threshold alone. Target setting is simple in this case because of the strong discontinuity we have assumed about human health risks. It is easy to see why an absolute maximum emission standard is appropriate.

But now suppose that marginal health damage is a rising and continuous function of emissions, as in Figure 6.7. A trade-off now exists in which lower health risks can be obtained at the cost of some loss of pollution benefits (or, if you prefer lower health risks involve higher emission abatement costs). It is now clear that with such a trade-off, both benefits and costs matter. A 'modified efficiency target' would correspond to emissions level $M^{H*}$.

**Box 6.3 No regrets and a double dividend from environmental control?**

It is sometimes possible to achieve environmental objectives at no cost or, better still, at 'negative' cost. Not surprisingly, ways of doing things that have such effects are known as 'no regrets' policies. There are several reasons why these may arise:

- double dividends;
- elimination of technical and economic inefficiencies in the energy-using or energy-producing sectors;
- induced technical change;
- achievement of additional ancillary benefits, such as improved health or visual amenity.

We will explain these ideas in the context of one potential example: reducing the emissions of carbon dioxide to reduce global climate change. First, the 'double dividend' hypothesis is explained.

**The double dividend hypothesis**

The double dividend idea arises from the possibility that the revenues from an emissions tax (or a system of permits sold by auction) could be earmarked to reduce marginal rates of other taxes in the economy. If those other taxes have distortionary (i.e. inefficiency-generating) effects, then reducing their rate will create efficiency gains. Thus an environmental tax with revenues ring-fenced for reducing distortionary taxes has a double benefit (dividend); the environment is improved *and* efficiency gains accrue to the economy as whole.

There are other reasons why 'no regret' options may be available. The existence of market imperfections can cause firms to be producing away from the frontier of what is technically and/or economically possible. Firms may be unaware of new techniques, or poorly informed

**Box 6.3 *continued***

about waste recycling mechanisms. Companies may have old, technologically obsolete capital, but are unable because of credit market imperfections to update even when that would generate positive net present value. An environmental programme that requires firms to use new, less polluting techniques, or which provides incentives to do so, can generate a different kind of double benefit. Pollution is reduced and productive efficiency gains are made.

One special case of this is dynamic efficiency gains, arising through induced technical change. It has long been recognised (see, for example, Porter, 1991) that some forms of regulatory constraint may induce firms to be more innovative. If a pollution control mechanism can be devised that accelerates the rate of technical change, then the mechanism may more than pay for itself over the long run. One area where this may be very important is in policy towards the greenhouse effect. Grubb (2000) argues persuasively that the provisions of the Kyoto Protocol will have beneficial induced effects on technical change. He writes:

> general economic processes of international investment and the dissemination of technologies and ideas – accelerated by the provisions on technology transfer and other processes under the Convention and the Protocol – could contribute to global dissemination of cleaner technologies and practices. In doing so, they will also yield multiplicative returns upon industrialised country actions.
>
> Grubb (2000), p. 124

More generally, there is a large set of possible ancillary benefits to environmental reforms. Perhaps the most important type is health benefits. Reductions of greenhouse gases tend to go hand in hand with reductions in emissions of secondary pollutants (such as particulates, sulphur dioxide, nitrogen dioxide and carbon monoxide), which can have important health impacts.

Some writers distinguish between a 'weak form' and a 'strong form' of the double dividend hypothesis. For a revenue-neutral environmental reform, the weak form refers to the case where total real resource costs are lower for a scheme where revenues are used to reduce marginal rates of distortionary taxes than where the revenues are used to finance lump-sum payments to households or firms. There is almost universal agreement that this hypothesis is valid. The strong form asserts that the real resource costs of a revenue-neutral environmental tax reform are zero or negative. Not surprisingly, this hypothesis is far more contentious.

For a more thorough examination of the double dividend hypothesis, and some empirical results, see the Word file *Double Dividend* in *Additional Materials*, Chapter 6.

**Box 6.4 Measures of stocks and flows for a variety of pollutants**

Pollutant emissions are measured (like all flows) in rates of output per period of time. For example, it is estimated that worldwide anthropogenic emissions of carbon dioxide, the most important greenhouse gas, were 6.9 gigatonnes of carbon equivalent per year (6.9 GtC/yr) as of 1990.[7] These flows accumulate through time as pollutant stocks, measured either in quantities in existence at some point in time, or in terms of some measure of concentration in an environmental medium of interest to us. Carbon dioxide atmospheric concentrations have risen from about 280 ppmv (parts per million by volume) in 1750 (the start of the industrial era) to 367 ppmv in 1999 (an increase of 31%). The current rate of change of the $CO_2$ concentration rate is estimated to be 1.5 ppmv per year (a growth rate of 0.4% per year). IPCC scenarios suggest that by 2100, concentrations will be in the range 549 to 970 ppm (90 to 250% above pre-industrial levels).

*Sources*: Technical Summary of the Working Group 1 Report (IPCC(1), 2001), particularly Figure 8, p. 36

[7] A metric tonne is equal to 1000 kilograms (kg). Commonly used units for large masses are (i) a gigatonne (Gt) which is $10^9$ tonnes, (ii) a megatonne (Mt) which is $10^6$ tonnes, and (iii) a petragram (Pg) which is equal to 1 Gt. Finally, 1 GtC = 3.7 Gt carbon dioxide.

*Table 6.3* Expected lifetimes for several pollutants

| | Pre-industrial concentration | Concentration in 1998 | Rate of concentration change | Atmospheric lifetime |
|---|---|---|---|---|
| $CO_2$ (carbon dioxide) | about 280 ppm | 365 ppm | 1.5 ppm/yr | 5 to 200 yr[1] |
| $CH_4$ (methane) | about 700 ppb | 1745 ppb | 7.0 ppb/yr | 12 yr |
| $N_2O$ (nitrous oxide) | about 270 ppb | 314 ppb | 0.8 ppb/yr | 114 yr |
| CFC-11 (chlorofluorocarbon-11) | zero | 268 ppt | −1.4 ppt/yr | 45 yr |
| HFC-23 (hydrofluorocarbon-23) | zero | 14 ppt | 0.55 ppt/yr | 260 yr |
| $CF_4$ (perfluoromethane) | 40 ppt | 80 ppt | 1 ppt/yr | >50 000 yr |
| Sulphur | Spatially variable | Spatially variable | Spatially variable | 0.01 to 7 days |
| $NO_X$ | Spatially variable | Spatially variable | Spatially variable | 2 to 8 days |

Note:
1. No single lifetime can be defined for $CO_2$ because of the different rates of uptake by different removal processes
*Sources*: Technical Summary of the IPCC Working Group 1 Report, IPCC(1) (2001), Table 1, p. 38

## 6.6 Efficient levels of emission of stock pollutants

The analysis of pollution in Section 6.4 dealt with the case of flow pollution, in which pollution damage depends directly on the level of emissions. In doing so, there were two reasons why it was unnecessary to distinguish between flows and stocks of the pollutant. First, both benefits and damages depended on emissions alone, so as far as the objective of net benefit maximisation was concerned, stocks – even if they existed – were irrelevant. But we also argued that, strictly speaking, stocks do not exist for pure flow pollutants (such as noise or light).

How do we need to change the analysis in the case of stock pollutants where damage depends on the stock level of the pollutant? It turns out to be the case – as we shall see below – that the flow pollution model also provides correct answers in the special (but highly unlikely) case where the pollutant stock in question degrades into a harmless form more-or-less instantaneously. In that case, the stock dimension is distinguishable from the flow only by some constant of proportionality, and so we can work just as before entirely in flow units. But in all other cases of stock pollutants, the flow pollution model is invalid.

The majority of important pollution problems are associated with stock pollutants. Pollution stocks derive from the accumulation of emissions that have a finite life (or residence time). The distinction between flows and stocks now becomes crucial for two reasons. First, without it understanding of the science lying behind the pollution problem is impossible. Second, the distinction is important for policy purposes. While the damage is associated with the pollution stock, that stock is outside the direct control of policy makers. Environmental protection agencies may, however, be able to control the rate of emission flows. Even where they cannot control such flows directly, the regulator may find it more convenient to target emissions rather than stocks. Given that what we seek to achieve depends on stocks but what is controlled or regulated are typically flows, it is necessary to understand the linkage between the two.

As we shall now demonstrate, the analysis of stock pollution necessitates taking account of space and time. For clarity of presentation it will be convenient to deal with these two dimensions separately. To do so, we draw a distinction between pollutants with a relatively short residence time (of the order of a day or so) and those with considerably longer lifetimes (years rather than days, let us say). Table 6.3 provides some idea of the active life expectancy of a range of pollutants under normal conditions.

## 6.7 Pollution control where damages depend on location of the emissions

In this section and the next we deal with stock pollutants which have relatively short residence times in the environmental media into which they are dumped. To help fix ideas, consider the graphic in Figure 6.8 which represents two polluting ‘sources’,

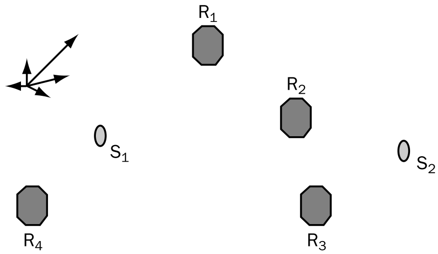


*Figure 6.8* A spatially differentiated airshed

$S_1$ and $S_2$, that are located near four urban areas, $R_1$, $R_2$, $R_3$ and $R_4$. These areas contain populations whose health is adversely affected by local ambient concentrations of the pollutant. Our interest lies in the amount of pollution these areas – called 'receptors' – receive from the emission sources. We assume that emissions from the two sources persist for at most a few days; atmospheric processes break up and degrade concentrations rather quickly, so that on any one day pollutant concentrations are determined purely by emissions of the last few days. There is no long-term accumulation effect taking place.

Now consider the extent of pollutant dispersion and mixing. Mixing of a pollutant refers to the extent to which physical processes cause the pollutant to be dispersed or spread out. One possibility is that emissions are 'uniformly mixing' (UM). A pollutant is uniformly mixing if physical processes operate so that the pollutant quickly becomes dispersed to the point where its spatial distribution is uniform. That is, the measured concentration rate of the pollutant does not vary from place to place. This property is satisfied, for example, by most greenhouse gases.

By definition, the location of the emission source of a UM pollutant is irrelevant as far as the spatial distribution of pollutant concentrations is concerned. Irrespective of the source location, pollutant stocks become evenly distributed across the whole spatial area of interest – in our picture over the whole rectangle depicted. All that matters, as far as concentration rates at any receptor are concerned, is the total amount of those emissions.

What can be said about the efficient level of emissions with the twin properties of short residence time (whose accumulation is therefore negligible) and uniform mixing? Intuition suggests that the simple flow pollution model developed in Section 6.4 can be used with only minor modification. To see why, note that there will be a one-to-one relationship between the level of emissions of the pollutant ($M$) and the pollutant stock size ($A$). Specifically, $M$ and $k$ are related by a fixed coefficient relationship of the form $A = kM$, with $k$ fixed for any particular kind of pollution. Therefore, while damage is a function of the stock, and benefit is a function of flow, the damage function can be translated into an equivalent flow function using the $A = kM$ relationship, permitting use of the flow pollution model. A simple numerical example is given in Box 6.5. This has been reproduced as an Excel spreadsheet in Sheet 2 of the workbook *Targets examples.xls*. As was the case for the numerical example in Box 6.2, the Excel workbook has been set up to allow comparative static analysis to be carried out, and shows the use of Solver to obtain a direct solution to the optimisation problem.

As we will now see, the flow pollution model cannot be used where the pollutant is not uniformly mixing nor where it has a relatively long lifespan. (Can you explain why?) Most air, water and ground pollutants are not uniformly mixing. Look at Figure 6.8 again. Suppose that the principal determinants of the spatial distribution of the pollutant are wind direction and velocity. In the diagram, the length and direction of the arrow vectors in the multiple arrow symbol represent the relative frequency of these two components. Clearly, emissions from $S_1$ are going to matter much more for the four receptor areas than emissions from $S_2$. Furthermore, looking at emissions from S1 alone, these are likely to raise pollutant concentration levels to a greater amount in R1 than in the other three receptors. R4 is likely to suffer the least from emissions by either source.

Other factors will, of course, come into play too. For example, suppose R1 is at high elevation, whereas R2 is situated in a depression surrounded by a ring of hills. Then R2 may experience the highest concentrations, both on average and at peak times. All of this amounts to saying that where pollutants are not uniformly mixing, location

**Box 6.5 Efficient solution for a uniformly mixed and short-lived stock pollutant: a numerical example**

As in Box 6.2 we suppose that total benefits function is given by:

$$B = \begin{cases} 96M - 0.2M^2 & \text{for} \quad 0 \le M \le 240 \\ 11\,520 & \text{for} \quad M > 240 \end{cases}$$

Our total damage, however, now needs to be specified appropriately for a stock pollutant and is taken to be:

$$D = 0.2A^2 \text{ for } A \ge 0$$

and in steady state we assume that $A = 2M$

**What are $M^*$ and $A^*$?**

We first consider the case in which there is an interior solution with $M$ positive but less than 240. The relevant first derivatives are:

$$\mathrm{d}B/\mathrm{d}M = 96 - 0.4M$$

$$\mathrm{d}D/\mathrm{d}M = 1.6M$$

(as $D = 0.2A^2$ implies $D = 0.2 \times (2M)^2 = 0.8M^2$ which implies $\mathrm{d}D/\mathrm{d}M = 1.6M$).

Now setting $\mathrm{d}B/\mathrm{d}M = \mathrm{d}D/\mathrm{d}M$ we obtain:

$$96 - 0.4M = 1.6M \rightarrow M^* = 48 \text{ and so } A^* = 96$$

**Additional materials**

As we remarked at the end of Box 6.2, a spreadsheet can be used for obtaining solutions to problems of this kind, or for carrying out comparative statics. Sheet 2 of the Excel workbook *Targets examples.xls* sets up a template for simple stock pollution models of this form. The interested reader may find it helpful to explore that sheet.

matters. There will not be a single relationship between emissions and concentration over all space. A given total value of $M$ will in general lead to differentiated values of $A$ across receptors. Moreover, if $M$ remained constant but its source distribution changed then the spatial configuration of $A$ would also change.

Non-uniform mixing is of great importance as many types of pollution fall into this category. Examples include ozone accumulation in the lower atmosphere, oxides of nitrogen and sulphur in urban airsheds, particulate pollutants from diesel engines and trace metal emissions. Many water and ground pollutants also do not uniformly mix. An environmental protection agency (EPA) may attempt to handle these spatial issues by controlling *ex ante* the location of pollution creators and victims. This approach, implemented primarily by zoning and other forms of planning control, forms a substantial part of the longer-term way of dealing with spatial aspects of pollution. However, in the next section we focus on the situation in which the location of polluters and people is already determined, and moving either is not a feasible option. Our interest must then lie in how targets for emissions from the various sources can be calculated (and, in the next chapter, on what instruments can be used).

## 6.8 Ambient pollution standards

It will be convenient to use a little elementary matrix algebra for the exposition of the arguments that follow. For the reader unfamiliar with matrix algebra, or who needs a quick refresher, a brief appendix is provided at the end of this chapter (Appendix 6.1) explaining the notation used in matrix algebra and stating some simple results. It would be sensible to read that now.

Some additional notation is now required. Using earlier terminology, we regard the environment as a series of spatially distinct pollution 'reception' areas (or receptors). Suppose that there are $J$ distinct receptors, each being indexed by the subscript $j$ (so $j = 1, 2, \ldots, J$) and $N$ distinct pollution sources, each being indexed by the subscript $i$ (so $i = 1, 2, \ldots, N$). Various physical and chemical processes determine the impact on pollutant concentration in any particular receptor from any particular source. For simplicity, we assume that the relationships are linear. In that case, a set of constant 'transfer coefficients' can be defined. The transfer coefficient $d_{ji}$ describes the impact on pollutant concentration at receptor $j$ attributable to source $i$.[8] The total level, or concentration rate, of pollution at location $j$, $A_j$, will be the sum of the contributions to pollution at that location from all $N$ emission sources. This can be written as

$$A_j = \sum_{i=1}^{N} d_{ji} M_i \tag{6.6}$$

where $M_i$ denotes the total emissions from source $i$.

A numerical example will help. In the case shown in Figure 6.8, we have $N = 2$ sources and $J = 4$ receptors. Then we have four equations corresponding to equation 6.6. These are

$$A_1 = d_{11}M_1 + d_{12}M_2 \tag{6.7a}$$

$$A_2 = d_{21}M_1 + d_{22}M_2 \tag{6.7b}$$

$$A_3 = d_{31}M_1 + d_{32}M_2 \tag{6.7c}$$

$$A_4 = d_{41}M_1 + d_{42}M_2 \tag{6.7d}$$

We can collect all eight $d_{ji}$ coefficients into a $J \times N$ matrix, **D**. Denoting the vector of emissions from the two sources as **M** and the vector of ambient pollution levels in the four receptors as **A** we have

$$\mathbf{A} = \mathbf{DM} \tag{6.8}$$

or

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \\ d_{41} & d_{42} \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \tag{6.9}$$

Knowledge of the **M** vector and the **D** matrix allows us to calculate ambient pollution levels at each receptor. If, for example, **D** and **M** are

$$\mathbf{D} = \begin{bmatrix} 0.7 & 0.1 \\ 0.9 & 0.2 \\ 0.3 & 0.2 \\ 0.1 & 0.0 \end{bmatrix} \text{ and } \mathbf{M} = \begin{bmatrix} 10 \\ 20 \end{bmatrix}$$

then $A_1 = 9$, $A_2 = 13$, $A_3 = 7$ and $A_4 = 1$. The Excel workbook *Matrix.xls* and Word file *Matrix.doc* in *Additional Materials*, Chapter 6, illustrate how this – and other similar – matrix calculations can be done using a spreadsheet program.

Armed with this terminology, we now answer the following question in a general way: what is the socially efficient level of emissions from each source? As in all previous cases in this chapter, it will be the set of emission levels that maximises net benefits. To see how this works here, note that there are $N$ emission sources, and so our solution will consist of $N$ values of $M_i$, one for each source. Benefits consist of the sum over all $N$ sources of each firm's pollution benefits. So we have

$$B = \sum_{i=1}^{N} B_i(M_i)$$

Damages consist of the sum over all $J$ receptor areas of the damage incurred in that area. That is,

$$D = \sum_{j=1}^{J} D_j(A_j)$$

Hence the net benefits function to be maximised (by appropriate choice of $M_i$, $i = 1, \ldots, N$) is

$$\text{NB} = \sum_{i=1}^{N} B_i(M_i) - \sum_{j=1}^{J} D_j(A_j) \tag{6.10}$$

By substitution of equation 6.6 into 6.10, the latter can be written as

$$\text{NB} = \sum_{i=1}^{N} B_i(M_i) - \sum_{j=1}^{J} D_j\left(\sum_{i=1}^{N} d_{ji} M_i\right) \tag{6.11}$$

A necessary condition for a maximum is that

$$\begin{aligned} \frac{\partial \text{NB}}{\partial M_i} &= B_i'(M_i) - \sum_{j=1}^{J} D_j'(A_j)\frac{\text{d}A_j}{\text{d}M_i} = 0 \quad \text{for } i = 1, \ldots, N \\ &= B_i'(M_i) - \sum_{j=1}^{J} D_j'(A_j) d_{ji} = 0 \text{ for } i = 1, \ldots, N \end{aligned} \tag{6.12}$$

[8] The linearity assumption is a very good approximation for most pollutants of interest. (Low-level ozone accumulation is one significant exception.) Each coefficient $d_{ji}$ will, in practice, vary over time, depending on such things as climate and wind conditions. However, if we measure average values of these coefficients over some period of time, they can be regarded as constant coefficients for the purposes of our analysis.

which, after rearranging, yields the set of $N$ marginal conditions

$$B_i'(M_i) = \sum_{j=1}^{J} D_j'(A_j)d_{ji} \text{ for } i = 1, \ldots, N$$

Where

$$D_j'(A_j) = \frac{\partial D_j}{\partial A_j} \tag{6.13}$$

The intuition behind this result is straightforward. The emissions target (or standard) for each firm should be set so that the private marginal benefit of its emissions (the left-hand side of the equation) is equal to the marginal damage of its emissions (the right-hand side of the equation). Note that because the $i$th firm's emissions are transferred to some or all of the receptors, the marginal damage attributable to the $i$th firm is obtained by summing its contribution to damage over each of the $J$ receptors.

An interesting property of the solution to equation set 6.13 is that not only will the efficient emission level differ from firm to firm, but also the efficient ambient pollution level will differ among receptors. It is easy to see why efficient emission levels should vary. Firms located at different sources have different pollution impacts: other things being equal, those sources with the highest pollution impact should emit the least. But what lies behind the result that efficient levels of pollution will vary from place to place? Receptors at different spatial locations will experience different pollution levels: other things being equal, those receptors which would (in an unconstrained world) experience the highest pollution-stock level should have the highest efficient ambient pollution level. Of course, these two considerations have to be met jointly; NB = $B - D$ is being maximised, and so we are searching for the best trade-off between the benefits reduction and damages reduction. Appendix 6.2 provides a worked numerical example of efficient emissions that illustrates this point.

In practice, environmental regulators might deem that it is unethical for $A$ to vary from place to place. So, they might impose an additional constraint on the problem to reflect this ethical position. One form of constraint is that the pollution level in no area should exceed some maximum level $A^*$ (that is $A_j^* \leq A^*$ for all $j$). Another, stricter, version would be the requirement that $A$ should be the same over all areas (that is $A_j^* = A^*$ for all $j$). In the latter case, the net benefit function to be maximised is

$$L = \sum_{i=1}^{N} B_i(M_i) - \sum_{j=1}^{J} D_j(A^*) \tag{6.14}$$

By imposing additional constraints, maximised net benefit is lower in equation 6.14 than in equation 6.10. An efficiency loss has been made in return for achieving an equity goal.

## 6.9 Intertemporal analysis of stock pollution

We now consider the case of stock pollutants that have a relatively long active (i.e. damaging) lifespan but which are uniformly mixing. Doing so has two implications. First, the uniformly mixing assumption implies that pollutant concentrations will not differ from place to place, and so the spatial dimension of emissions control is no longer of direct relevance. Second, persistence of pollution stocks over time means that the temporal dimension is of central importance. As we shall see, an efficient pollution control programme will need to take account of the trajectory of emissions over time, rather than just at a single point in time.

The model we use to examine pollution targets is the simplest possible one that can deal with the intertemporal choices involved. Damage at time $t$ is determined by the contemporaneous stock size or concentration rate of the pollutant in a relevant environmental medium. Gross benefits depend on the flow of emissions. Hence our damage and (gross) benefit functions have the general forms

$$D_t = D(A_t) \tag{6.15}$$

$$B_t = B(M_t) \tag{6.16}$$

The variables $A$ and $M$ in equations 6.15 and 6.16 are, of course, not independent of one another. With relatively long-lived pollutants, emissions add to existing stocks and those stocks accumulate over time. However, except in the special case where pollutants are infinitely long-lived, part of the existing stock will decay or degrade into a harmless form over time, thereby having a negative impact on stock accumulation. A convenient way of representing this stock–flow relationship is by assuming that the rate of change of the pollutant stock over time is governed by the differential equation

$$\dot{A}_t = M_t - \alpha A_t \tag{6.17}$$

where a dot over a variable indicates its derivative with respect to time, so that $\dot{A}_t = \mathrm{d}A/\mathrm{d}t$. To interpret this equation, it will be helpful to have an example in mind. Consider atmospheric carbon dioxide ($CO_2$), one source of which is emissions from the combustion of fossil fuels. Current emissions ($M_t$) add to $CO_2$ stocks, and so the concentration level rises; that is, $\dot{A}_t$ is positive. However, offsetting factors are at work too. Some of the existing $CO_2$ stock will be transformed into harmless substances by physical or chemical processes, or will be absorbed into oceans or other sinks where it has no damaging effect. In other words, part of the pollution stock decays. The amount of pollution decay is captured by the term $-\alpha A_t$.

The net effect on $A$ (and so whether $\dot{A}_t$ is positive or negative overall) depends on the magnitudes of the two terms on the right-hand side of equation 6.17.[9] The parameter $\alpha$ is a proportion that must lie in the interval zero to one. A pollutant for which $\alpha = 0$ exhibits no decay, and so the second term on the right-hand side of equation 6.17 is zero. This is known as a perfectly persistent pollutant. In this special case, integration of equation 6.17 shows that the stock at any time is the sum of all previous emissions. Notice that the absence of decay means that damages arising from current emissions will last indefinitely. This is approximately true for some synthetic chemicals, such as heavy metal residuals, and toxins such as DDT and dioxin. Moreover, the pollution stock and pollution damages will increase without bounds through time as long as $M$ is positive.

More generally, we expect to find $0 < \alpha < 1$, and denote this as an imperfectly persistent pollutant. Here, the pollutant stock decays gradually over time, being converted into relatively harmless elements or compounds. Greenhouse gases provide one example, but (as we show in Chapter 10) with slow or very slow rates of decay. The second limiting case, where $\alpha = 1$, implies instantaneous decay, and so the pollutant can be regarded as a flow rather than a stock pollutant. We need deal with this special case no further here.

The specification given in equation 6.17 imposes the restriction that the parameter $\alpha$ is constant; a constant *proportion* of the pollution stock decays over any given interval of time. This may be invalid in practice. If the restriction is approximately true equation 6.17 might still be used for reasons of convenience and simplicity. But if it is grossly inaccurate, and the decay rate (or assimilation rate as it is often called) changes substantially over time, or varies with changes in either $A$ or $M$, then it is not an appropriate basis for modelling. We will return to this matter later.

We mentioned earlier that, unlike in the previous cases investigated in this chapter, the relationship between $M$ and $A$ is not independent of time. By integrating equation 6.17 over time we obtain

$$A_t = \int_{\tau=t_0}^{\tau=t} (M_t - \alpha A_t)\,\mathrm{d}\tau$$

where $t_0$ denotes the first point in time at which the pollutant in question was emitted. Thus the pollution stock level at any time $t$, $A_t$, depends on the entire history of emissions up to that point in time. Even if emissions had been at a constant level in the past and were to remain so in the future, $A$ would not be constant throughout time. Put another way, as emissions

[9] In this chapter, we are working principally with economic models specified in continuous time terms. However, sometimes it is convenient to work in a discrete time framework. Doing this requires defining the meaning to be attached to time subscripts for stock variables. A convention that we follow throughout this text is that for any stock variable the subscript $t$ denotes the *end* of period $t$. Then the discrete time counterpart of equation 6.17 would be:

$$A_t - A_{t-1} = M_t - \alpha A_{t-1}$$

Notice that the last term on the right-hand side now has the time subscript $t-1$, as compared with $t$ in equation 6.17. Given our convention, $A_{t-1}$ refers to the pollution stock at the end of period $t-1$ (or, equivalently, start of period $t$). The discrete time counterpart of equation 6.17 would then say that the inflow (new emissions) is taking place contemporaneously with the outflow (stock decay), and that it is the difference between inflow and outflow during period $t$ that determines whether stock will rise, fall or remain constant between the end of period $t-1$ and the end of period $t$. This is intuitively sensible.

at time $t$ add to pollution stocks at that time *and* in future time periods, there is no one-to-one relationship between $A$ and $M$. It is because time matters here in a fundamental way that the variables in equations 6.15 and 6.16 are time-dated.[10]

As time periods are linked together through a stock–flow relationship, efficient pollution targets and policies must be derived from an intertemporal analysis. We proceed by assuming that the policy maker aims to maximise discounted net benefits over some suitable time horizon. For simplicity, the horizon is taken to be of infinite span. Using $t = 0$ to denote the current period of time, and defining the net benefits of pollution as gross benefits minus damages (specified respectively by equations 6.15 and 6.16) the policy maker's objective is to select $M_t$ for $t = 0$ to $t = \infty$ to maximise

$$\int_{t=0}^{t=\infty} (B(M_t) - D(A_t))e^{-rt}dt \tag{6.18}$$

where $r$ is the social (consumption) discount rate.

A complete description of efficient stock pollution will, therefore, consist not of a single number for, but a *trajectory* (or time path) of, emission levels through time. In general, this optimal trajectory will be one in which emission levels vary throughout time. However, in many circumstances, the trajectory will consist of two phases. One of these phases is a so-called *steady state* in which emissions (and concentration levels) remain constant indefinitely at some level. The other is an adjustment phase; the trajectory describes a path by which emissions (and concentrations) move from current levels to their efficient, steady-state levels. This adjustment process may be quick, or it may take place over a long period of time.

Even with complete information, obtaining such a trajectory is technically difficult, involving the calculus of optimal control. We will explain this technique in Chapter 14, and apply it to the pollution model being examined here in Chapter 16. In this chapter, we consider only the second of the two phases described above: the efficient *steady-state* pollution level.[11] In a steady state, by definition, the pollution flow and the pollution stock are each at a constant, unchanging level.[12] Hence the time subscripts we have attached to variables become redundant and can be dropped. Moreover, with an unchanging stock $\dot{A}_t = 0$ and so equation 6.17 simplifies to $M = \alpha A$. The intuition that lies behind this is straightforward: for a pollutant that accumulates over time, the pollution stock can only be constant if emission inflows to the stock ($M$) are equal to the amount of stock which decays each period ($\alpha A$). It then follows that in a steady state, the stock–flow relationship between $A$ and $M$ can be written as

$$A = \frac{M}{\alpha} \tag{6.19}$$

This shows that, in a steady state, the smaller is the value of $\alpha$ the larger will be the pollution stock for any given level of emissions.

The full derivation of the steady-state solution to this problem is presented in Chapter 16. You may wish to return to, and reread, this section after studying that later chapter. Here, we just state one major result from that solution, interpret it intuitively, and discuss some of its characteristics. If you are prepared to take this result on trust, little will be lost by not going through its derivation.

The key result we draw upon from Chapter 16 is that an efficient steady-state level of pollution emissions requires that the following condition be satisfied:

$$\frac{dB}{dM} = \frac{dD}{dA}\left(\frac{1}{r+\alpha}\right) \tag{6.20}$$

Equation 6.20 is a variant of the familiar marginal condition for efficiency. The marginal benefit and

[10] In the last section, the relationship between stocks and flows of the pollutant was complicated because *space* mattered; the effect of $M$ on $A$ depended on the respective locations of the pollution source and recipient. There we used $i$ and $j$ terminology to denote that dependence on location. Here the relationship is complicated by the fact that *time* matters, hence the use of $t$ terminology.

[11] Doing this assumes that the problem is one in which a steady-state solution exists, which is not always true. Chapter 16 will briefly examine the adjustment process to a steady state, and whether such a state exists.

[12] There is a second sense in which the term *steady state* is sometimes used: as a state in which all variables of interest in some system are growing at a constant rate. We do not use this alternative meaning in this text.

the marginal cost of the chosen emissions level should be equal. More precisely, it can be read as an equality between the present value of the gross benefit of a marginal unit of pollution (the left-hand side of 6.20) and the present value of the damage that arises from the marginal unit of pollution (the right-hand side of 6.20). Note that a marginal emission today has benefits only today, and so the present value of that marginal emission is identical to its current marginal benefit. In contrast, the damage arising from the marginal emission takes place today and in future periods. The 'discount factor' $1/(r + \alpha)$ has the effect of transforming the single period damage into its present-value equivalent. (A fuller explanation of this interpretation is given in Chapter 16.) At the level of $M$ that satisfies equation 6.20, the value taken by the expression on each side of the equation is known as the shadow price of a unit of emission. It is labelled as $\mu$ in several of the diagrams in this chapter and will figure prominently in our discussions in the next chapter.[13]

Examination of equation 6.20 shows two very important results:

1. Other things being equal, the faster is the decay rate, the higher will be the efficient level of steady-state emissions. Reasoning: For any given value of $\mathrm{d}D/\mathrm{d}A$, a rise in $\alpha$ implies that the value of $\mathrm{d}B/\mathrm{d}M$ would have to fall to satisfy the marginal equality. A lower value of $\mathrm{d}B/\mathrm{d}M$ implies higher emissions. Intuition: The greater is the rate of decay the larger is the 'effective' discount rate applied to the marginal stock damage term and so the smaller is its present value. A higher discount rate means we attach less weight to damages in the future, and so the emission level can be raised accordingly.
2. Other things being equal, the larger is the consumption discount rate, the higher will be the efficient level of steady-state emissions. Reasoning: For any given value of $\mathrm{d}D/\mathrm{d}A$, a rise in $r$ implies that the value of $\mathrm{d}B/\mathrm{d}M$ would have to fall to satisfy the marginal equality. A lower value of $\mathrm{d}B/\mathrm{d}M$ implies higher emissions. Intuition: The greater is the consumption discount rate $r$, the larger is the discount rate applied to the stock damage term and so the smaller is its present value. A higher discount rate means we attach less weight to damages in the future, and so the emission level can be raised accordingly.

*Table 6.4* Special cases of equation 6.21

| | Imperfectly persistent pollutant $\alpha > 0$ | Perfectly persistent pollutant $\alpha = 0$ |
|---|---|---|
| $r = 0$ | A | D |
| $r > 0$ | B | C |

Problem 4 at the end of this chapter asks the reader to explore these and other results from the stock pollution model. The model is simulated in the Excel workbook *Stock1.xls*.

For the purpose of looking at some special cases of equation 6.20, it will be convenient to rearrange that expression as follows (the full derivation is given in Chapter 16):

$$\frac{\mathrm{d}D}{\mathrm{d}M} = \frac{\mathrm{d}B}{\mathrm{d}M}\left(1 + \frac{r}{\alpha}\right) \tag{6.21}$$

Four special cases of equation 6.21 can be obtained, depending on whether $r = 0$ or $r > 0$, and on whether $\alpha = 0$ or $\alpha > 0$. We portray these combinations in Table 6.4.

### Case A: $r = 0$, $\alpha > 0$

In this case the pollutant is imperfectly persistent and so eventually decays to a harmless form. With $r = 0$, no discounting of costs and benefits is being undertaken. Equation 6.21 collapses to:[14]

[13] In some of the economics literature, the shadow price of emissions is constructed to be a negative quantity (and would correspond here to the negative of $\mu$). This arises because some authors choose to attach a different interpretation to the shadow price. Whenever a different interpretation is being used in our text, that will be made clear to the reader explicitly.

[14] Notice that equation 6.23 appears to be identical to the efficiency condition for a flow pollutant. But it is necessary to be careful here, as 6.23 holds only in a steady state, and is not valid outside those states for a stock pollutant.

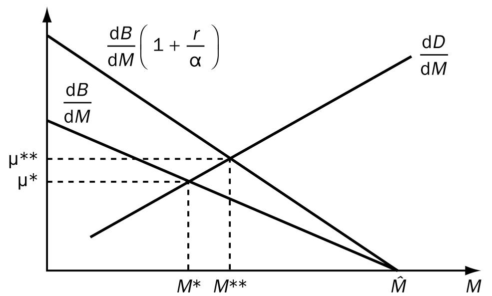


*Figure 6.9* Efficient steady-state emission level for an imperfectly persistent stock pollutant. Two cases: $\{r = 0 \text{ and } \alpha > 0\}$ and $\{r > 0 \text{ and } \alpha > 0\}$

$$\frac{dD}{dM} = \frac{dB}{dM} \tag{6.22}$$

This has a straightforward interpretation. An efficient steady-state rate of emissions for a stock pollutant requires that the contribution to benefits from a marginal unit of pollution flow be equal to the contribution to damage from a marginal unit of pollution flow. The steady-state equilibrium is shown in Figure 6.9 (by the intersection of the functions $dD/dM$ and $dB/dM$). Net benefits are maximised at the steady-state pollution flow $M^*$. In the steady state, $A^*$ will be at the level at which $\alpha A^* = M^*$, and both the pollution stock and emissions track along through time at constant levels. You may find it useful to look at Box 6.6 at this point; this goes through a simple numerical example to illustrate the nature of the equilibrium.

### Case B: $r > 0$, $\alpha > 0$

With $r$ and $\alpha$ being positive numbers, the equilibrium condition is given by equation 6.21 in unchanged form. The marginal equality in this case incorporates the additional term $1/(r + \alpha)$ to reflect the presence of discounting at a positive rate. This is shown diagrammatically in Figure 6.9, with $M^{**}$ denoting the equilibrium emission level. It is instructive to compare this equilibrium with that obtained in Case A. As $r$ increases above zero, the marginal benefits function rotates clockwise about the point $\hat{M}$. Discounting, therefore, increases the steady-state level of emissions. Moreover, the larger is the discount rate, the larger is the amount by which efficient steady-state emissions rise. Intuitively, a larger value of $r$ reduces the present value of the future damages that are associated with the pollutant stock. In effect, higher weighting is given to present benefits relative to future costs the larger is $r$. However, notice that the shadow price of one unit of the pollutant emissions becomes larger as $r$ increases.

**Box 6.6 Steady-state efficient solution for a stock pollutant: a numerical example**

**No discounting, $r = 0$ (Case A: $r = 0$, $\alpha > 0$)**

Let $\alpha = 0.5$, $D = A^2$, $B = 96M - 2M^2$.
What are $M^*$ and $A^*$?

$B = 96M - 2M^2 \rightarrow dB/dM = 96 - 4M$

$D = A^2 = (M/\alpha)^2 = (1/0.5)^2M^2$
$= 4M^2 \rightarrow dD/dM = 8M$

Now setting $dB/dM = dD/dM$ we obtain:

$96 - 4M = 8M \rightarrow M^* = 8$

Therefore $A = (M/\alpha) \rightarrow A^* = 16$

This result is obtained by inspection and by use of Solver in Sheet 1, and shown graphically in Chart 1, of Excel workbook *Stock1.xls* in the *Additional Materials* for Chapter 6.

**Positive discounting, $r > 0$ (Case B: $r > 0$, $\alpha > 0$)**

Let $\alpha = 0.5$, $r = 0.1$, $D = A^2$, $B = 96M - 2M^2$.
What are $M^*$ and $A^*$?

$B = 96M - 2M^2 \Rightarrow dB/dM = 96 - 4M$

$D = A^2 = (M/\alpha)^2 = (1/0.5)^2M^2$
$= 4M^2 \rightarrow dD/dM = 8M$

Now setting $\frac{dD}{dM} = \frac{dB}{dM}\left(1+\frac{r}{\alpha}\right)$ we obtain:

$8M = (96 - 4M)(1 + \{0.1/0.5\}) \rightarrow M^* = 9$

Therefore $A = (M/\alpha) \rightarrow A^* = 18$

This result is obtained by inspection and by use of Solver in Sheet 2, and shown graphically in Chart 2, of Excel workbook *Stock1.xls*. Note that we use Solver there to find the value of $M$ that sets marginal net benefits (expressed in terms of emissions) equal to zero.

### Cases C ($r > 0$, $\alpha = 0$) and D ($r = 0$, $\alpha = 0$)

In both Cases C and D the pollutant is perfectly persistent, and so never decays to a harmless form. One might guess that something unusual is happening here by noting that equation 6.21 is undefined when $\alpha = 0$; division by zero is not a legitimate mathematical operation. The intuition that lies behind this is straightforward. No steady state exists except for the case in which $M$ is zero. A steady state cannot exist for any positive value of $M$ as $A$ would rise without bound. But then pollution damage would also rise to infinity.

It follows that, at some point in time, the environmental protection agency would have to require that emissions be permanently set to zero to avoid the prospect of intolerable damage. The pollution stock level would remain at whatever level $A$ had risen to by that time. Pollution damage would also continue indefinitely at some constant level, but no additional damage would be generated. The zero-emissions steady-state solution turns out to be perfectly in accord with good sense.

One caveat to this conclusion should be noted. Although a perfectly persistent pollutant has a zero natural decay rate, policy makers may be able to find some technique by which the pollutant stock may be artificially reduced. This is known as clean-up expenditure. If such a method can be found, and can be implemented at reasonable cost, it allows the possibility of some perpetual level of emissions. We examine this possibility further in Chapter 16.

Of course, even if the EPA accepted that emissions would have to be set to zero at some date (and remain zero thereafter), the question remains of which date the switch to zero should be made. Steady-state analysis is unable to answer this question. To obtain that answer, another technique (or another criterion than economic efficiency) is required. Chapter 16 shows how optimal control can be used to find both the efficient steady-state solution and the optimal adjustment path to it.

## 6.10 Variable decay

The stock pollution models used in this chapter have assumed that the proportionate rate of natural decay of the stock, $\alpha$, is constant. This assumption is commonly employed in environmental economics analysis, but will not always be valid. In many situations,

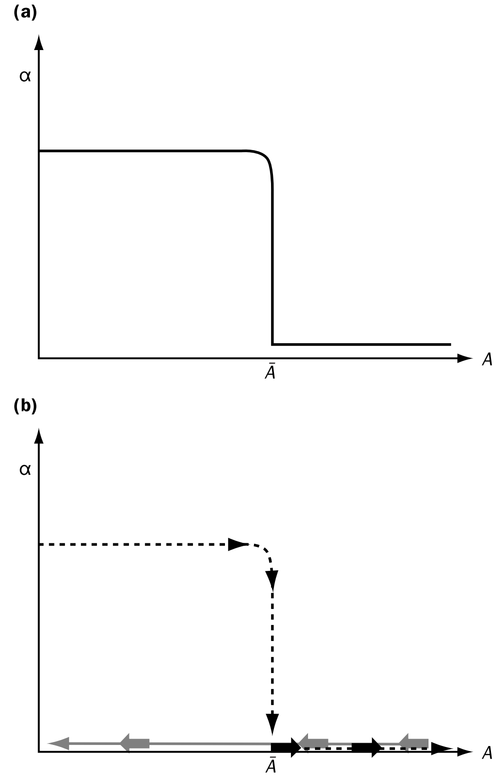


*Figure 6.10* Threshold effects and irreversibilities

one would expect that the rate of decay depend on the size of the pollution stock, or on some other associated variable. For example, it is thought that the decay rate of greenhouse gases alters (in quite complex ways) as mean temperature levels change. Of particular importance are the existence of threshold effects (where the decay rate changes in a sudden, discontinuous way) and irreversibilities (where the nature of a relationship changes depending on the direction in which a variable is moving). One example of a threshold effect is illustrated in the top panel of Figure 6.10. Here the decay rate of a waterborne pollutant collapses towards zero as some threshold level of biological oxygen demand (BOD) on a river is reached. This critical level of BOD is reached when the pollution stock is at $\bar{A}$. The lower panel illustrates a threshold effect combined with an irreversibility. The arrows denote the direction in which $A$ is changing. As the pollution stock rises from a low level, $\alpha$ collapses suddenly at the threshold $\bar{A}$ and remains close to zero as $A$ continues to rise. At high levels of pollution, the biological ability of the river to break down harmful pollutants might be largely destroyed. So if the change is reversed, and $A$ falls from a high value, the value of $\alpha$ would remain very low (as shown by the path with left-pointing arrows). This path dependence is also known as hysteresis; in this example, the history of pollutant flows matters, and reversing pollution pressures does not bring one back to the *status quo ex ante*.

Another way of thinking about this issue is in terms of carrying capacities (or assimilative capacities, as they are sometimes called) of environmental media. In the case of water pollution, for example, we can think of a water system as having some capacity to transform pollutants into harmless forms. The stock pollution model of Section 6.8 has in effect assumed unlimited carrying capacities: no matter how large the load on the system, more can always be carried and transformed at a constant proportionate rate.

Whether this assumption is plausible is, in the last resort, an empirical question. Where it is not, modelling procedures need to reflect limits to carrying capacity. The suggestions for further reading point you to some literature that explores models with variable pollution decay rates.

## 6.11 Convexity and non-convexity in damage and abatement cost functions

When benefit and damage functions were first presented in Section 6.4, a number of assumptions were made about their shapes. Those assumptions relate to the concept of convexity of a function. After explaining what is meant by a convex function, this section gives some examples of why the relevant functions may not be convex, and then shows some consequences of non-convexity.

Consider a function, $f(x)$, of a single variable $x$. The function is *strictly convex* if the line segment connecting any two distinct points on the function

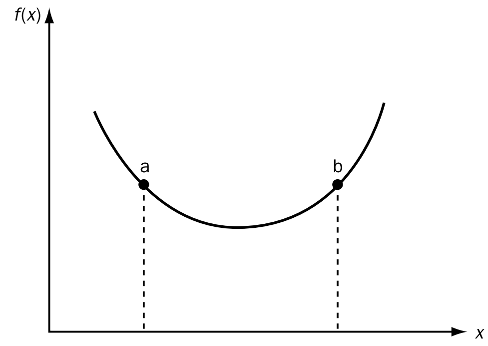


*Figure 6.11* A strictly convex function

lies everywhere above the function $f(x)$, except at the two points themselves. A function is convex (as opposed to strictly convex) if the line segment lies everywhere above or on the function $f(x)$, but not below it. As an example, the function graphed in Figure 6.11 is strictly convex.

Looking back at Figure 6.4, it is clear that the damage function $D(M)$ is convex.[15] This is not true for the benefits function $B(M)$ as that is drawn in Figure 6.4. However, suppose that we reinterpret benefits as avoided abatement costs, as suggested earlier. Now construct the horizontal image of $B(M)$, so that moving to the right on this mirror image corresponds to more pollution abatement. Then the abatement cost function will be convex.

Actually, this terminological contortion is not really necessary. What really matters, as we shall see, is whether the functions describing the problem being investigated are smooth, continuous, and lead to unique marginal efficient conditions. All of these properties are satisfied by the benefit and damage functions used in Figure 6.4. It is clear from the lower panel of Figure 6.4 that there is just one level of pollution at which the marginal efficiency condition is satisfied: the marginal benefit of pollution (or equivalently marginal cost of abatement) is equal to the marginal damage of pollution. This implies that marginal analysis alone is sufficient for identifying the efficient level of pollution.[16]

[15] In fact, as drawn it is strictly convex. But what matters is whether the weaker property of convexity is satisfied. So we shall use the word 'convex' from now on to cover strict as well as (weak) convexity.

[16] Mathematically, the efficient pollution level is obtained from the first-order conditions for optimisation; second-order conditions will automatically be satisfied (and so do not need to be checked).

### 6.11.1 Non-convexity of the damage function and its implications

There are many reasons why the damage function and the abatement cost function may be non-convex. Here we restrict attention to the more commonly discussed case of non-convex damages. So what might cause a pollution damage function to not be of the smooth, continuously increasing form that we have assumed so far? One example was given implicitly in Section 6.10 where we introduced the ideas of threshold effects and irreversibility. A closely related example to that is acidic pollution of rivers and lakes. Here, pollution may reach a threshold point at which the lake become biologically dead, unable to support animal or plant life. No further damage is done as pollution levels rise beyond that point. The total and marginal damages function in this case will be of the form shown in Figure 6.12.

Another example, discussed in Fisher (1981), is non-convexity of damages arising from averting behaviour by an individual. Suppose a factory emits particulate emissions that create external damages for an individual living in the neighbourhood of the factory. The marginal damage initially rises with the amount of the pollution. However, at some critical level of pollution flow, the affected individual can no longer tolerate living in the neighbourhood of the factory, and moves to live somewhere else where the damage to him or her becomes zero. As far as this particular individual is concerned, their marginal damage function is also of the form shown in Figure 6.12. However, if there are many individuals living in the neighbourhood, with varying tolerance levels, many of whom are prepared to move at some level of pollution, the aggregate marginal pollution damage function will be the sum of a set of individual functions, each of which has the same general form but with differing pollution tolerance levels. The aggregate damage function will be of the non-convex form shown in the top panel of Figure 6.13, with its marginal counterpart being shown by the curve labelled MD in the central panel.

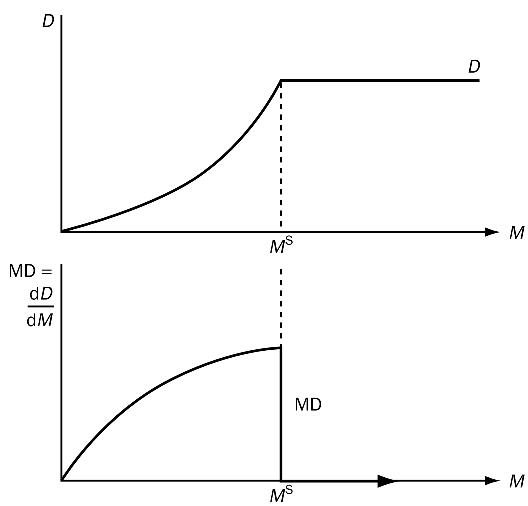


*Figure 6.12* A non-convex damage function arising from pollution reaching a saturation point

Now combine the marginal damage function for the averting behaviour example with a marginal benefit function of conventional shape. This is shown in the central panel of Figure 6.13. Marginal damage and benefits are equalised here at three emission levels. To ascertain which of these, if any, is the efficient level of pollution, it is necessary to inspect the level of total net benefits at these three points, *and* at all other levels of emission (as net benefits will not necessarily even correspond to a marginal equality when one or more function is not convex). The two points labelled A and B are 'local optima', as they satisfy the second-order conditions for a local maximum of net benefits, as shown in the lower panel of Figure 6.13. In this case it can be seen by inspection of the NB curve that $M_3$ is a 'global' net benefits-maximising pollution level. Note that in moving from $M_1$ to $M_3$, net benefits at first fall (by the area labelled **a**) and then rise (by the area labelled **b**).

Why does non-convexity matter? There are two major reasons why this is a matter of concern. The first could be described as a 'practical' matter: calculating the efficient level of emissions (or pollution stock) is likely to be more complicated than where all functions are convex. This is partly a matter of computational difficulty. But more importantly, it is to do with the fact that the information required to identify the (non-convex) functions may be immense and very costly to obtain. Obtaining

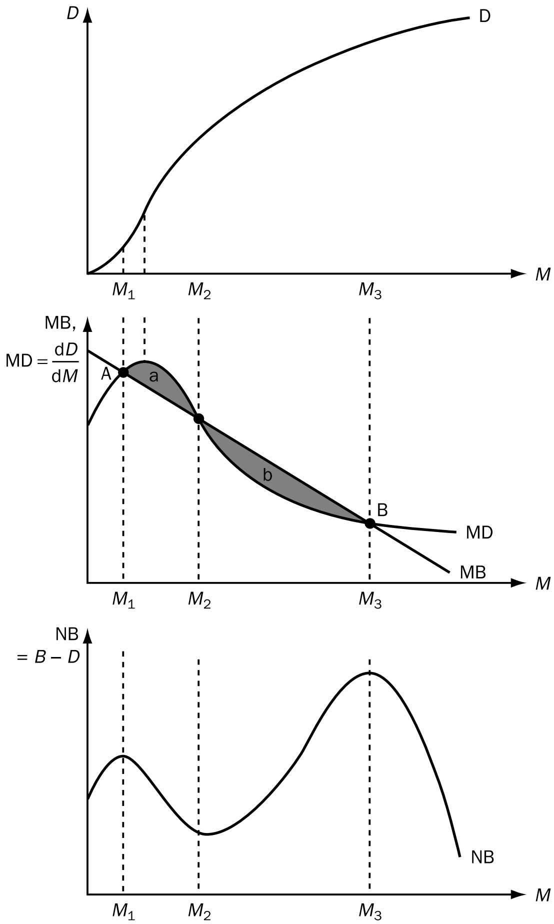


*Figure 6.13* Multiple marginal equalities arising from a non-convex damage function: the case of behavioural adjustments of individuals

reliable estimates of functions will be particularly difficult where information is limited or uncertain.

The second reason for concern is more fundamental. Non-convexity may be important because it exists but we do not recognise that it exists. In that case, some commonly advocated tools could give seriously misleading results. For example, a failure to recognise the existence of threshold effects or irreversibilities could render project appraisal using cost–benefit analysis completely wrong. (One example is explored in Problem 5 at the end of this chapter.)

One reason why policy makers may fail to recognise non-convexity is to do with the way information is acquired. We often find out about things by exploring a relevant 'local neighbourhood'. For example, cross-section sampling techniques may generate data on emissions and damage that are relatively closely clustered around current levels, and tell us little or nothing about properties of the function outside the current sample range. Inspection of that data may suggest convexity when in fact the function is only convex over part of its range. This becomes important – and potentially dangerous – if the policy maker falsely projects the apparently convex function outside this current range.

## 6.12 Estimating the costs of abating pollution[17]

There are many ways in which estimates can be made of the costs of pollution abatement. Two broad classes can be identified:

- engineering models;
- economic models.

In practice, most studies have used linked engineering–economic models, but the relative attention paid to each component varies widely.

### 6.12.1 Engineering models

These typically use what is called a 'bottom-up' approach. An emissions abatement objective is defined. Then all the techniques by which this target could be achieved are listed. For each technique, the researcher calculates the expected expenditures by firms on pollution abatement equipment and other investments, fuel, operation, maintenance and other labour costs. The costs incurred by each firm are

[17] For a more extensive version of the material in this section, see *Additional Materials*: Chapter 6 '*Abatement costs*'.

then added up to arrive at the total economy-wide abatement cost. Hence the name ‘bottom-up’. For a complete accounting of control costs, expenditures incurred by regulatory agencies should be added in. Best achievable abatement costs are those which are the minimum among those techniques studied. A more modest variant of this approach would involve the researcher obtaining cost estimates of one technique rather than all available. This requires making assumptions about the form of responses of firms to the controls they face.

There are some desirable properties in estimating abatement costs in this way. They are simple to understand, and simple (at least in principle) to undertake. Engineering models are typically highly disaggregated. They consider technology options in a rich, detailed way, providing large amounts of information at the micro-production level. This technology-rich property means that engineering models are very well suited to costing specific projects, such as using wind power to generate 25% of a country’s electricity. They are also capable of dealing in a careful way with some kinds of ‘no-regret’ or ‘free-lunch’ possibilities arising from technical and economic inefficiencies in existing method of production. (See Box 6.3 for more details.)

But this approach also has some serious limitations. Each technology is assessed independently via an accounting of its costs and savings, but possible interdependencies (or linkages and feedback) between the elements being studied and the economy as a whole are not taken into account. This leads to biased estimates of the true costs of abatement. Some examples of important linkages that matter – but which are typically ignored by engineering models – are:

- productivity changes induced by regulatory control;
- changes in unemployment;
- change in overall industrial structure of the economy.

The most fundamental problem is that engineering models ignore changes in relative prices, and the associated impacts on factor substitution and the behaviour of firms and individuals. Results can be seriously misleading because of this, particularly when long-term effects are being investigated.

### 6.12.2 Economic models

These are typically ‘top-down’ models.[18] They are constructed around a set of aggregate economic variables, the relationships among which are determined by (micro or macro) economic theory and equilibrium principles. These relationships are estimated econometrically, using time-series data. Alternatively, relationships are calibrated to match with data for one chosen base year. To obtain cost estimates, some project of interest such as the introduction of a carbon tax is taken as an exogenous shock. The model is solved for equilibrium before and after the shock. By comparing the values of relevant variables in the baseline and shocked case, cost estimates are obtained.

The top-down nature of these models means that they tend to be highly aggregated, and that they do not have the richness of detail (particularly about energy technology options) that can be captured in engineering models. The strength of economic models lies in their ability to deal with supply and demand relationships, and to capture behavioural changes and substitution effects that are important for making inferences about long-term consequences. In addition, they are good for the analysis of distributional effects, and for simulating the use of economic instruments.

But economic models alone treat the energy sector as a relatively undifferentiated whole, and so are of limited use for answering questions that involve changes within the energy sector. Aggregate output–energy use relationships tend to be relatively inflexible, and so economic models are not well suited to examining possible decoupling effects. One major practical limitation of economic models is their assumption that resource allocation in the baseline case is already fully efficient. As a result, they can say nothing about negative cost potential from removing existing inefficiencies.

[18] See IPCC(3) (2001) for further analysis of bottom-up and top-down models.

Economic models typically yield higher abatement cost estimates than engineering models. This arises because (a) they do not consider existing inefficiencies and (b) they take account of losses of consumer surplus arising from price increases as regulated firms attempt to pass additional costs on to consumers.

### 6.12.3 Linked or integrated engineering–economic models

Ideally, one would like to base cost estimates on models that combine the advantages of economic and engineering models. This might be done by linking the two, or by more systematically developing an integrated modelling approach. Among the many attempts that have been made to do this, we find the following types.

#### 6.12.3.1 Input–output (IO) models

IO models (see Chapter 9 for a more developed account) partition the economy into a number of sectors, and then represent the economy mathematically by a set of simultaneous linear equations. These equations embody the input–output relationships between those sectors. IO models, therefore, capture sectoral interdependences and spillovers. So, for example, if the use of coal were to be reduced, IO models could explore the ramifications of this for the economy as a whole, and so give some idea about the likely costs. However, the fixed coefficients in the IO equations preclude modelling of behavioural changes and factor substitution effects as relative prices change. Hence, they will tend to overestimate abatement costs. IO models are useful for short-run modelling where disaggregated sectoral detail is required.

#### 6.12.3.2 Macroeconomic models

Macroeconomic models give a key role to changes in effective demand and investigate the resulting quantity changes. More sophisticated models also incorporate overall wage and price-level changes, and describe the dynamics of, and adjustment to, new equilibria as a result of shocks. When these models are linked with others that deal more richly with the energy sector, they can be useful for investigating the short-run and medium-term cost implications of environmental policy changes.

#### 6.12.3.3 Computable general equilibrium (CGE) models

CGE models (see Chapter 9 for more details) simulate the behaviour of agents based on optimising microeconomic theory. The models are solved for sets of prices and wages that generate general equilibrium. CGE models are typically static models, and do not analyse adjustment processes from one equilibrium state to another. They are widely used to simulate the consequences of emissions taxes.

#### 6.12.3.4 Dynamic energy optimisation models

These are 'bottom-up', technology rich, partial equilibrium energy-sector models. They are used to minimise cost of the energy sector over a long-term horizon, yielding a partial equilibrium for energy markets. Sophisticated versions allow energy demand to respond to price, and examine the dynamics of changes in the energy sector (and so can trace out the evolution through time of changes in the size and type of capital stock used in the energy sector. Energy optimisation models are often linked with macro models.

#### 6.12.3.5 Purpose-built integrated energy–economic system simulation (E–E) models

E–E models are usually purpose-built to estimate abatement costs in one particular context (such as the costs of abatement required to attain Kyoto Protocol targets for greenhouse gases). They are bottom-up representations of energy demand and supply technologies, and as such typically have a very rich specification of technologies at a highly disaggregated level. A purpose-built economic component is constructed that is consistent with the energy structure of the model. E–E models are often used to simulate the consequences (and costs) of various scenarios.

In practice, most E–E models are hybrids, with problems of inconsistency between components. For

**Box 6.7 IPCC estimates of the costs of $CO_2$ abatement to reach Kyoto Protocol targets**

The *gross* abatement costs to attain Kyoto targets for carbon dioxide reduction depend on several factors:

1. The magnitude of emissions reduction required to meet the target. Assumptions made about marginal sources of supply (cost and availability of carbon-based and carbon-free technologies)
2. Short- and long-run price elasticities
3. Whether or not there is emissions trading (and how extensive this is)

Point 1 implies that the emissions 'baseline' is critical to the magnitude of total abatement costs. The larger emissions growth would be in the absence of control, the higher will be total abatement costs required to attain the Kyoto target. Emissions baseline growth rate of $CO_2$ depends on GDP growth, the rate of decline of energy per unit output, and the rate of decline of $CO_2$ emissions per unit energy.

The *net* costs depend on the gross costs and also on

1. Availability of no-regrets efficiency gains (e.g. can revenues be used to reduce marginal rates on other distortionary taxes – such as income, sales, or employment taxes – or reduce other technical/economic inefficiencies?)
2. Whether abatement will generate other ancillary benefits
3. The magnitude of any induced technical progress. Of importance here, in terms of the timing of costs, is whether the innovation route is via R&D or learning-by-doing.

Working Group III of the Intergovernmental Panel on Climate Change (IPCC) of the United Nations commissioned a number of independent modelling groups to simulate emissions reductions achieved by carbon taxes. Each of these groups employed some variant of energy–economy model. Tax revenues were recycled via lump-sum payments to the whole economy. The value of the tax rate required to achieve an emissions target indicates the marginal abatement cost in that model. With each team using different assumptions about baseline emissions and different model structures and/or parameter values, the exercise allows multi-model comparisons to be made, and the sensitivity of findings to variations in assumptions can be explored.

The estimated marginal abatement costs from these various models for attaining Kyoto Protocol targets by 2010 are shown in Table 6.5. Figures are given for three scenarios. The first scenario is one in which no trading of allowances is allowed

*Table 6.5* Marginal abatement costs (1990 US$/tC) for attainment of Kyoto target by 2010

| Model | No trading | | | | Annex 1 trading | Global trading |
|---|---|---|---|---|---|---|
| | US | OECD-Europe | Japan | CANZ | | |
| ABARE-GTEM | 322 | 665 | 645 | 425 | 106 | 23 |
| AIM | 153 | 198 | 234 | 147 | 65 | 38 |
| CETA | 168 | | | | 46 | 26 |
| Fund | | | | | 14 | 10 |
| G-Cubed | 76 | 227 | 97 | 157 | 53 | 20 |
| GRAPE | | 204 | 304 | | 70 | 44 |
| MERGE3 | 264 | 218 | 500 | 250 | 135 | 86 |
| MIT-EPPA | 193 | 276 | 501 | 247 | 76 | |
| MS-MRT | 236 | 179 | 402 | 213 | 77 | 27 |
| RICE | 132 | 159 | 251 | 145 | 62 | 18 |
| SGM | 188 | 407 | 357 | 201 | 84 | 22 |
| WorldScan | 85 | 20 | 122 | 46 | 20 | 5 |
| Administration | 154 | | | | 43 | 18 |
| EIA | 251 | | | | 110 | 57 |
| POLES | 135.8 | 135.3 | 194.6 | 131.4 | 52.9 | 18.4 |

*Source*: IPCC(III) 2001, Table TS.4, p. 56
One set of results (Oxford) has been omitted from this table, as it had not been fully reviewed at the time of writing, and relied on early 1980s data for initial parameterisation.
Models do not take account of induced technical progress, Clean Development Mechanism, sinks, negative cost options, targeted recycling of revenues, ancillary benefits, inclusion of non-$CO_2$ gases, or inefficiencies in implementation. Models here are typically general equilibrium rather than bottom-up technology-rich models.

**Box 6.7 *continued***

between countries – each country must independently achieve the emission target for it specified in the Protocol (see Chapter 10 for details of these targets). In this case, marginal abatement costs are shown for four 'blocs' of countries. It is evident that the marginal abatement costs vary considerably over countries, implying that the total global emissions reduction is not being achieved at least cost.

A second scenario allows trading of allowances (permits) among the Annex 1 countries (roughly speaking, the industrialised economies). Notice how partial trading dramatically reduces marginal (and so total) abatement costs. This is even more evident in results for the third scenario in which trading can take place between any countries. The efficiency gains that this generates mean that marginal costs are reduced by around an order of magnitude (a tenfold reduction) in some cases.

example, one may have as its basis a sophisticated engineering model that can be used to calculate direct technical costs. Linked to this might be a module which uses observed market behaviour to estimate technology adaptations. Further components estimate welfare losses due to demand reductions, and the revenue gains and losses due to trade changes.

## 6.13 Choosing pollution targets on grounds other than economic efficiency

This chapter has been largely concerned with pollution targets set in terms of an economic efficiency criterion. But there are (at least) two reasons why this focus is unduly restrictive. First, in the context of limited or imperfect information, there may be immense difficulties in identifying economically efficient targets.[19] In that case, efficiency-based targets may be of theoretical interest only and have little practical significance. We examine this issue at some length in Chapter 8.

Second, policy makers are likely to have multiple objectives. Efficiency matters, but it is not the only thing that matters. It is not surprising, therefore, that targets (or 'environmental standards' as they are sometimes called) are often chosen in practice on the basis of a mix of objectives. The mix may include health or safety considerations, equity, and perceptions of what is technically feasible (usually subject to some 'reasonable cost' qualification). In recent years, sustainability has taken its place as another stated goal of policy. As we show in Chapter 8, sustainability in conjunction with imperfect information and uncertainty may also point to some form of precautionary principle being incorporated in the set of objectives pursued by policy makers.

National and international policy is also determined in the context of a network of pressures and influences. It is not surprising, therefore, that political feasibility plays a significant role. This has been particularly important in the area of international environmental agreements over such things as ozone depletion, acid rain and the greenhouse effect, as we show in Chapter 10.

Tables 6.6 and 6.7 list some existing environmental standards and the criteria that appear to have been used in their selection. In the next chapter we investigate which instruments are available to an environmental protection agency for attaining a given pollution target, however that target may have been determined.

[19] Many of the problems posed by imperfect information also apply to targets set on the basis of sustainability, health, or indeed any other criterion. Nevertheless, they apply particularly strongly to efficiency-based targets. However, as we shall see in the following chapter, several of the alternative criteria can be interpreted as appropriate for target setting precisely when information is imperfect. They should then be thought of as responses to uncertainty rather than as being weakened or limited by it.

*Table 6.6* Environmental targets

| Pollutant | Target | Relevant criterion |
|---|---|---|
| **United Kingdom** | | |
| Grains emitted in cement production | 0.1–0.2 grains per cubic foot | Best practicable means |
| Sewage concentration | Max. 30 mg/litre suspended solids<br>Max. BOD 20 mg/litre | 1976 National Water Council: precautionary principle, perceived health risks |
| Cadmium/lead | Discharges into North Sea to fall by 70% between 1985 and 1995 | Health criterion |
| PCBs | Phase out by 1999 | Strict precautionary principle – health risks |
| Waste recycling | 50% domestic waste to be recycled | Political target? |
| **United States** | | |
| Criteria air pollutants | See Table 6.7 | Health risks |
| **International** | | |
| CFCs | CFC production to fall to 80% and 50% of 1986 levels by 1994 and 1999 respectively | Political feasibility, with final targets set in terms of critical load |

Key: BOD = biochemical oxygen demand. The concepts of 'Best practicable means', 'Critical load', and 'Precautionary principle' are explained elsewhere in the chapter

*Table 6.7* Primary NAAQS for Criteria Air Pollutants, 1997

| | | Concentration level | |
|---|---|---|---|
| Pollutant | Averaging time | ppm | μg/m$^3$ |
| Particulate matter (PM10) | Annual | – | 50 |
| | 24-hour | – | 150 |
| Particulate matter (PM2.5) | Annual | – | 15 |
| | 24-hour | – | 65 |
| Sulphur dioxide | Annual | 0.030 | 80 |
| | 24-hour | 0.140 | 365 |
| Carbon monoxide | 8-hour | 9.000 | 10 |
| | 1-hour | 35.000 | 40 |
| Nitrogen oxide | Annual | 0.053 | 100 |
| Ozone | 8-hour | 0.008 | – |
| Lead | Max. quarterly | – | 1.5 |

## Summary

- We do not expect pure market economies to deliver efficient outcomes in terms of pollution. Pollution tends to be an externality to the market process and as a result is not adequately reflected in private market decisions. Put another way, while firms would meet the costs of controlling or abating pollution, the benefits of abatement would not be received by firms (although they would by society). Hence, in considering pollution abatement, the control level that maximises net benefits to firms is different from the level that maximises social net benefits.
- Economists often recommend that pollution targets should be set using an economic efficiency criterion. This can be thought of as selecting pollution targets so as to maximise social net benefits
- Economic efficiency is not the only relevant criterion for pollution target setting. Several others were discussed in the chapter. Which criteria are important to policy makers will tend to reflect their policy objectives and the constraints under which they operate.

- There are important differences between flow pollutants and stock pollutants in terms of the mechanisms by which damage is generated. This distinction has implications for the way in which targets are derived using an economic efficiency criterion. For stock pollutants, persistence implies that attention must be given to the accumulation (and decay) of pollutants over time, and so an intertemporal analysis is required. This is not necessary for the analysis of flow pollutants.
- For long-lived stock pollutants, pollution targets are best thought of in terms of emissions paths over time. Efficient pollution paths will not in general imply the same level of control at all points in time. However, it is often useful to think of steady-state outcomes and to investigate what (constant) level of pollution control would be efficient in an equilibrium steady state.
- Where a stock pollutant is not uniformly mixing, the spatial distribution of emissions sources becomes relevant. If targets are set in terms of pollutant concentrations, then the allowable emissions of any particular source will depend on its location.

## Further reading

Excellent and extensive presentations of the economics of pollution are to be found in Fisher (1981, chapters 5 and 6), Anderson (1985, 1991), Hartwick and Olewiler (1986, 1998) and Kolstad (2000). Baumol and Oates (1988) is a classic source in this area, although the analysis is formal and quite difficult. Cornes and Sandler (1996) provides a powerful theoretical underpinning in terms of the theory of public goods.

Tietenberg (1992) gives very extensive, descriptive coverage of several specific types of pollution. Other useful treatments which complement the discussion in this chapter are Dasgupta (1982, chapter 8), and two survey articles by Fisher and Peterson (1976) and Cropper and Oates (1992). Smith (1972) gives a mathematical presentation of the theory of waste accumulation. Several excellent articles can be found in the edited volume by Bromley (1995).

In this chapter we have taken a 'normative' approach to the setting of pollution targets, analysing what such targets should be in terms of some criterion of the public interest. An alternative literature considers targets in 'positive' terms, dealing with how targets are actually set. This approach focuses on the behaviour of interest groups, attempting to gain rents by manipulating government policy to their advantage. Good introductory accounts of this 'political economy' of regulation can be found in Goodstein (1995, 1999) and Kolstad (2000), particularly chapter 8. More advanced references are Laffont and Tirole (1993) which discusses theories of regulation, and Stigler (1971) and Peltzman (1976); these last two references are seminal works on the interest group theory of regulation.

Grubb (1998) provides a very interesting account of greenhouse gas policy, focusing on technological responses to the Kyoto Protocol. Ulph (1997) considers the relationship between environmental policy and innovation. Porter (1991) articulates the argument that strict environmental policy may be a factor which stimulates the rate of technological innovation. The double dividend hypothesis is discussed by Bovenberg (1997). The collection of readings edited by Carraro and Siniscalco (1997) focuses on the application of game theory to environmental problems. This is a particularly useful tool in the analysis of international pollution problems, as we shall see in Chapter 10, but has interesting applications too for domestic pollution policy. One of the first studies about the difficulties in designing optimal taxes (and still an excellent read) is Rose-Ackerman (1973).

Some journals provide regular applications of the economic theory of pollution. Of particular interest are the *Journal of Environmental Economics and Management*, *Ambio*, *Environmental and Resource Economics*, *Land Economics*, *Ecological Modelling*, *Marine Pollution Bulletin*, *Ecological Economics* and *Natural Resources Journal*.

## Discussion questions

1. 'Only the highest standards of environmental purity will do.' Discuss.
2. 'A clean environment is a public good whose benefits cannot be privately appropriated. Therefore private industry which is run for private gain will always be the enemy of a clean environment.' Examine this proposition.
3. Discuss the relevance and application of the concept of externalities in environmental economics.

## Problems

1. Under which circumstances will the economically optimal level of pollution be zero? Under which circumstances will it be optimal to undertake zero pollution abatement?
2. We have seen that the efficient level of pollution is the one that minimises the sum of total abatement costs plus total damage costs. Refer now to Figure 6.5. Show that if pollution abatement takes place to the extent $\hat{M} - M_A$ the sum of total damage costs and total abatement costs is $C_1 + C_2 + C_3$. Prove that 'too little' abatement (relative to the optimal quantity) results in higher costs than $C_1 + C_2$.
3. Explain the concept of the 'efficient level of pollution'. What information is required in order to identify such an efficient quantity?
4. Using equation 6.20 or 6.21, deduce the effect of (i) a decrease in $\alpha$ and (ii) an increase in $r$ (*ceteris paribus*) on:
   (a) $M^*$
   (b) $A^*$
   (c) $\mu^*$

   Note that you could answer this question analytically. Alternatively, you could explore the issue numerically using the Excel file *Stock1.xls* (found in the *Additional Materials* for Chapter 6).
5. This problem illustrates how marginal analysis might give misleading results in the presence of non-convexity. It is based on an example from Goodstein (1995). Nitrogen oxides ($NO_x$), in combination with some volatile organic compounds and sunlight, can produce damaging lower-atmosphere ozone smog. Initially, the damage rises at an increasing rate with $NO_x$ emissions. However, high levels of $NO_x$ act as ozone inhibitors, and so beyond some critical level of emissions, higher levels of $NO_x$ reduce ozone damage.
   (i) Sketch a marginal damage (MD) function that is consistent with these properties.
   (ii) Add to your diagram a conventionally shaped marginal benefits function (or marginal abatement cost function) that intersects the MD function in more than one place.
   (iii) By an appropriate choice of some initial level of emissions, demonstrate that the following rule may give misleading results. Rule: emissions should be increased (decreased) if a small increase in emissions increases (decreases) net benefits.

## Appendix 6.1 Matrix algebra

### A6.1.1 Introduction

In this chapter, and in a few of the later ones (particularly Chapter 9 and the appendix to Chapter 14), some use is made of matrix algebra notation and elementary matrix operations. This appendix provides, for the reader who is unfamiliar with matrix algebra, a brief explanation of the notation and an exposition of a few of its fundamental operations. We deal here only with those parts of matrix algebra that are necessary to understand the use made of it in this text. The reader who would like a more extensive account should go to any good first-year university-level mathematics text. For example, chapter 4 of Chiang (1984) provides a relatively full account of introductory-level matrix algebra in an accessible form.

### A6.1.2 Matrices and vectors

A *matrix* is a set of elements laid out in the form of an array occupying a number of rows and columns. Consider an example where the elements are numbers. Thus, the array of numbers

| | |
|---|---|
| 0.7 | 0.1 |
| 0.9 | 0.2 |
| 0.3 | 0.2 |
| 0.1 | 0.0 |

can be called a matrix. In such an array, the relative positions of the elements *do* matter. Two matrices are identical if the elements are not only the same but also occupy the same positions in each matrix. If the positions of two or more elements were interchanged, then a different matrix would result (unless the interchanged elements were themselves identical).

It is conventional, for presentational purposes, to place such an array within square brackets, and to label the matrix by a single bold letter (usually upper-case).[20] So in the following expression, **A** is the name we have given to this particular matrix of eight numbers.

$$\mathbf{A} = \begin{bmatrix} 0.7 & 0.1 \\ 0.9 & 0.2 \\ 0.3 & 0.2 \\ 0.1 & 0.0 \end{bmatrix}$$

It is also conventional to define the *dimension* of a matrix by the notation $m \times n$ where $m$ is the number of rows occupied by the elements of the matrix and $n$ is the number of columns occupied by elements of the matrix. So, for our example, **A** is of dimension $4 \times 2$ as its elements span four rows and two columns.

Notice that because elements of matrices span rows and columns, they can be handled very conveniently within spreadsheet programs.

Sometimes we want to define a matrix in a more general way, such that its elements are numbers, but those numbers are as yet unspecified. To do this we could write **A** in the more general form

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix}$$

Notice the way in which each of the elements of this matrix has been labelled. Any one of them is $a_{ij}$ where $i$ denotes the row in which it is found and $j$ denotes its column. With this convention, the bottom right element of the matrix – here $a_{42}$ – will necessarily have a subscript identical to the dimension of the matrix, here $4 \times 2$.

It is convenient to have another shorthand notation for the matrix array. This is given by

$$\mathbf{A} = [a_{ij}] \quad \begin{matrix} i = 1, \ldots, m \\ j = 1, \ldots, n \end{matrix}$$

The bracketed term here lets the reader know that what is being referred to is a matrix with $m \times n$ elements $a_{ij}$.

#### A6.1.2.1 A special form of matrix: the identity matrix

A matrix is said to be *square* if its row and column dimensions are equal (it has the same number of rows and columns). Thus, the matrix

$$\mathbf{B} = \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$$

is a $2 \times 2$ square matrix. Furthermore, if the coefficients of a square matrix satisfy the restrictions that each element along the leading (top left to bottom right) diagonal is 1 and every other coefficient is zero, then that matrix is called an *identity matrix*. Thus the matrix

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is a $2 \times 2$ identity matrix. An identity matrix is often denoted by the symbol **I**, or sometimes by $\mathbf{I}_n$

[20] The use of square brackets is not universal; some authors prefer round brackets or braces.

where the $n$ serves to indicate the row (and column) dimension of the identity matrix in question. In our example, it would be $\mathbf{I}_2$.

#### A6.1.2.2 Vectors

A vector is a special case of a matrix in which all elements are located in a single row (in which case it is known as a row vector) or in a single column (known as a column vector). Looking at the various rows and columns in the $4 \times 2$ matrix $\mathbf{A}$ above, it is evident that we could make up six such vectors from that matrix. We could construct four row vectors from the elements in each of the four rows of the matrix. And we could make up two column vectors from the elements in each of the two columns.[21] The four row vectors constructed in this way are

$$[a_{11} \quad a_{12}] \quad [a_{21} \quad a_{22}] \quad [a_{31} \quad a_{32}] \quad \text{and} \quad [a_{41} \quad a_{42}]$$

each of which is of dimension $1 \times 2$, while the two column vectors, each of dimension $4 \times 1$, are given by

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{bmatrix} \text{ and } \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ a_{42} \end{bmatrix}$$

#### A6.1.2.3 The transpose of a matrix or a vector

Various 'operations' can be performed on matrices.[22] One of the most important – and commonly used – is the operation of forming the 'transpose' of a matrix. The transpose of a matrix is obtained by interchanging its rows and columns, so that the first column of the original matrix becomes the first row of the transpose matrix, and so on. Doing this implies that if the original matrix $\mathbf{A}$ were of dimension $m \times n$, its transpose will be of dimension $n \times m$. The transpose of $\mathbf{A}$ is denoted as $\mathbf{A}'$, or sometimes as $\mathbf{A}^{\mathrm{T}}$.

Consider two examples. First, let $\mathbf{a}$ be the $4 \times 1$ column vector

$$\mathbf{a} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{bmatrix}$$

then its transpose, $\mathbf{a}'$ is given by the row vector $\mathbf{a}' = [a_{11}\ a_{21}\ a_{31}\ a_{41}]$.

As a second example, consider the first array that we introduced in this appendix. That matrix and its transpose are given by

$$\mathbf{A} = \begin{bmatrix} 0.7 & 0.1 \\ 0.9 & 0.2 \\ 0.3 & 0.2 \\ 0.1 & 0.0 \end{bmatrix} \qquad \mathbf{A}' = \begin{bmatrix} 0.7 & 0.9 & 0.3 & 0.1 \\ 0.1 & 0.2 & 0.2 & 0.0 \end{bmatrix}$$

#### A6.1.2.4 Bold notation for vectors and matrices

As we mentioned earlier, it is conventional to use the **bold** font to denote vectors or matrices, and to use an ordinary (non-bold) font to denote a scalar (single number) term. Hence, in the following expression, we can deduce from the context and the notation employed that each of $\mathbf{a_1}$ and $\mathbf{a_2}$ is a column vector consisting respectively of the first column of scalars and the second column of scalars. We know that the element $a_{21}$, for example, is a scalar because it is not written in bold font.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} = [\mathbf{a_1} \quad \mathbf{a_2}]$$

### A6.1.3 Other operations on matrices

As with algebra more generally, several operations such as addition and multiplication can, under some conditions, be performed on matrices.

#### A6.1.3.1 Addition and subtraction

Two matrices can be added (or subtracted) if they have the same dimension. Essentially, these operations

[21] One could also, of course, make up other vectors as mixtures of elements from different rows or columns.

[22] From this point on in this appendix, we shall use the term matrix to include both vectors and matrices, unless the context requires that we distinguish between the two.

involve adding (or subtracting) comparably positioned elements in the two individual matrices. Suppose that we wish to add the two $(m \times n)$ matrices $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$. The sum, $\mathbf{C} = [c_{ij}]$ is defined by

$$\mathbf{C} = [c_{ij}] = [a_{ij}] + [b_{ij}] \quad \text{where } c_{ij} = a_{ij} + b_{ij}$$

Example:

$$\begin{bmatrix} 7 & 1 \\ 9 & 2 \\ 3 & 2 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 9 & 1 \\ 0 & 4 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 7+3 & 1+0 \\ 9+9 & 2+1 \\ 3+0 & 2+4 \\ 1+2 & 0+3 \end{bmatrix} = \begin{bmatrix} 10 & 1 \\ 18 & 3 \\ 3 & 6 \\ 3 & 3 \end{bmatrix}$$

Matrix subtraction is equivalent, but with the addition operation replaced by the subtraction operation in the previous expression.

#### A6.1.3.2 Scalar multiplication

Scalar multiplication involves the multiplication of a matrix by a single number (a scalar). To implement this, one merely multiplies every element of the matrix by that scalar.

Example:

$$\text{If } \mathbf{A} = \begin{bmatrix} 0.7 & 0.1 \\ 0.9 & 0.2 \\ 0.3 & 0.2 \\ 0.1 & 0.0 \end{bmatrix} \text{ then } 2\mathbf{A} = \begin{bmatrix} 1.4 & 0.2 \\ 1.8 & 0.4 \\ 0.6 & 0.4 \\ 0.2 & 0.0 \end{bmatrix}$$

#### A6.1.3.3 Multiplication of matrices

Suppose that we have two matrices, **A** and **B**. Can these be multiplied by one another? The first thing to note is that here (unlike with ordinary algebra) the order of multiplication matters. Call **A** the lead matrix and **B** the lag matrix. For the matrix multiplication to be possible (or even meaningful) the following condition on the dimensions of the two matrices must be satisfied:

Number of columns in **A** = Number of rows in **B**

If this condition is satisfied, then the matrices are said to be 'conformable' and a new matrix **C** can be obtained which is the matrix product **AB**. The matrix **C** will have the same number of rows as **A** and the same number of columns as **B**.

How are the elements of **C** obtained? The following rule is used.

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \text{ for } i = 1 \text{ to } m \text{ and } j = 1 \text{ to } n$$

Example:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \\ 1 & 2 \end{bmatrix} \; \mathbf{B} = \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix} = \mathbf{AB} =$$

$$\begin{bmatrix} (2\times3)+(1\times4) & (2\times2)+(1\times1) \\ (0\times3)+(3\times4) & (0\times2)+(3\times1) \\ (1\times3)+(2\times4) & (1\times2)+(2\times1) \end{bmatrix} = \begin{bmatrix} 10 & 5 \\ 12 & 3 \\ 11 & 4 \end{bmatrix}$$

An intuitive way of thinking about this is as follows. Suppose we want to find element $c_{ij}$ of the product matrix **C** (the element in the cell corresponding to row $i$ and column $j$). To obtain this, we do the following:

- multiply the first element in row $i$ by the first element in column $j$
- multiply the second element in row $i$ by the second element in column $j$

  .

  .

  and so on up to
- multiply the final element in row $i$ by the final element in column $j$

The sum of all these multiplications gives us the number required for $c_{ij}$. (Note that this process requires the dimension condition that we stated earlier to be satisfied.) This process is then repeated for all combinations of $i$ and $j$.

Doing this kind of exercise by hand for even quite small matrices can be very time-consuming, and prone to error. It is better to use a spreadsheet for this purpose. To see how this is done – and to try it out for yourself with an Excel spreadsheet, *Matrix.xls* – read the file *Matrix.doc* in the *Additional Materials* for Chapter 6.

However, we suggest you calculate the products **AB** and **BA** of the following two 2 × 2 matrices **A** and **B** to convince yourself that **AB** does not equal **BA**.

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$$

#### A6.1.3.4 Division

Whereas obtaining the product of two matrices is a meaningful operation in matrix algebra, and can be done providing the two matrices are 'conformable', the same cannot be said of matrix division. Indeed, the division of one matrix by another is not a meaningful operation.

#### A6.1.3.5 The inverse matrix

However, a related concept – matrix inversion – does exist and is fundamental to much that is done in matrix algebra. To motivate this concept, think of ordinary algebra. If $a$ and $b$ are two numbers then the division of $a$ by $b$ (i.e. $a/b$) can be done, provided that $b$ is non-zero. But notice that $a/b$ can also be written as $ab^{-1}$, where $b^{-1}$ is the inverse (or reciprocal of $b$).

Where $\mathbf{B}$ is a matrix, we can under some conditions obtain its inverse matrix, $\mathbf{B}^{-1}$. And if we have a second matrix, say $\mathbf{A}$, which has the same number of rows as $\mathbf{B}^{-1}$ has columns, then the product $\mathbf{B}^{-1}\mathbf{A}$ can be obtained.

How is the inverse of B defined? The matrix inverse must satisfy the following equality:

$$\mathbf{BB}^{-1} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$$

That is, the product of a matrix and its inverse matrix is the identity matrix. Inspecting the dimension conditions implied by this definition shows that a matrix can only have an inverse if it is a square matrix.

Let us look at an example. The inverse of the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$$

is given by

$$\mathbf{A}^{-1} = \begin{bmatrix} 0 & 1 \\ 0.5 & -1.5 \end{bmatrix}$$

as

$$\begin{bmatrix} 0 & 1 \\ 0.5 & -1.5 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0.5 & -1.5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We will not give any methods here by which an inverse can be obtained. There are many such rules, all of which are tedious or difficult to implement once the matrix has more than 3 rows. Instead, we just report that a modern spreadsheet package can obtain inverse matrices by one simple operation, even for matrices of up to about 70 rows in size. There is clearly no need to bother about deriving an inverse by hand! And, of course, it is always possible to verify that the inverse is correct by checking that its product with the original matrix is $\mathbf{I}$.

Once again, to see how this is done, see *Matrix.doc* and *Matrix.xls*.

### A6.1.4 The uses of matrix algebra

The two main uses we make of matrix algebra in this text are

- to describe a system of linear equations in a compact way;
- to solve systems of equations or to carry out related computations.

Each of these is used in this chapter (in Section 6.8, where we discuss ambient pollution standards) and in Chapter 9. As an example of the first use, it is evident that the system of equations used in our ambient pollution example,

$$A_1 = d_{11}M_1 + d_{12}M_2$$

$$A_2 = d_{21}M_1 + d_{22}M_2$$

$$A_3 = d_{31}M_1 + d_{32}M_2$$

$$A_4 = d_{41}M_1 + d_{42}M_2$$

can be more compactly written as $\mathbf{A} = \mathbf{DM}$

where

$$\mathbf{D} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \\ d_{41} & d_{42} \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}$$

Check for yourself that, after the matrix multiplication $\mathbf{DM}$, this reproduces the original system of four equations.

The potential power of matrix algebra as a computational or solution device is illustrated in our analysis of input–output analysis in Chapter 9. We will leave you to follow the exposition there. As you will see, it is in this context that the inverse of a matrix is useful.

## Appendix 6.2 Spatially differentiated stock pollution: a numerical example

This appendix provides a numerical example of a spatially differentiated ambient pollution problem. We obtain the efficient level of $M$ for each source and $A$ for each receptor. Some of the material below is copied from the output of a Maple file *ambient.mws*. The interested reader can find the Maple file itself in the *Additional Materials* for Chapter 6.

The problem is one in which in the relevant spatial area ('airshed') there are two emissions sources, and two pollution receptors. The **D** matrix of transition coefficients is, therefore of the following form:

$$D_{ij} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}$$

for which we use below the specific values

$$\begin{bmatrix} 2 & 4 \\ 3 & 2 \end{bmatrix}$$

Assumptions used:

1. The marginal damage of pollution function is MD($A$) = $A$ (a very simple special case), and is identical everywhere.
2. The marginal benefit of emissions function, MB($M$), is identical for each firm, and is given by

$$\text{MB}(M_i) = a - bM_i$$

where we assume $a = 344$ and $b = 7$.

As shown in the text, an efficient solution requires that for each $i$, $i = 1,2$

$$\text{MB}(M_i) = \sum_{j=1}^{N}\left(\frac{\partial}{\partial A_j}D(A_j)\right)d_{ji}$$

which under Assumption (1) is

$$\text{MB}(M_i) = \sum_{j=1}^{N} A_j d_{ji}$$

This is here a two-equation linear system:

$$a - bM_1 = d_{11}A_1 + d_{21}A_2$$

$$a - bM_2 = d_{12}A_1 + d_{22}A_2$$

which gives:

$$a - bM_1 = d_{11}(d_{11}M_1 + d_{12}M_2) + d_{21}(d_{21}M_1 + d_{22}M_2)$$

$$a - bM_2 = d_{12}(d_{11}M_1 + d_{12}M_2) + d_{22}(d_{21}M_1 + d_{22}M_2)$$

We next define an expression (called sys1) that consists of these two equations:

$$sys1 := \{a - bM_2 = d_{12}(d_{11}M_1 + d_{12}M_2) + d_{22}(d_{21}M_1 + d_{22}M_2),$$
$$a - bM_1 = d_{11}(d_{11}M_1 + d_{12}M_2) + d_{21}(d_{21}M_1 + d_{22}M_2)\}$$

This can be solved (using the 'solve' command in Maple) to obtain solutions for $M_1$ and $M_2$ in terms of the parameters, $a$, $b$ and the components of the **D** matrix.

The solutions are given by

$$M_1 = \frac{(b - d_{11}d_{12} - d_{21}d_{22} + d_{12}^2 + d_{22}^2)a}{\left(\begin{array}{c} b^2 + bd_{12}^2 + bd_{22}^2 + d_{11}^2 b + d_{11}^2 d_{22}^2 + \\ d_{21}^2 b + d_{21}^2 d_{12}^2 - 2d_{11}d_{12}d_{21}d_{22} \end{array}\right)}$$

$$M_2 = \frac{a(b - d_{11}d_{12} + d_{21}^2 + d_{11}^2 - d_{21}d_{22})}{\left(\begin{array}{c} b^2 + bd_{12}^2 + bd_{22}^2 + d_{11}^2 b + d_{11}^2 d_{22}^2 + \\ d_{21}^2 b + d_{21}^2 d_{12}^2 - 2d_{11}d_{12}d_{21}d_{22} \end{array}\right)}$$

To obtain specific values for the solutions, we now substitute the particular values $a = 344$, $b = 7$, $d_{11} = 2$, $d_{12} = 4$, $d_{21} = 3$ and $d_{22} = 2$ for the parameters, giving the solution:

$$\{M_1 = 13, M_2 = 6\}$$

We next find the efficient ambient pollution levels in the two receptor areas. First define a new system of equations:

$$sys11 := \{A_1 = d_{11}M_1 + d_{12}M_2, A_2 = d_{21}M_1 + d_{22}M_2\}$$

This can be solved (using the 'solve' command) to obtain solutions for $A_1$ and $A_2$ in terms of the components of the **D** matrix and the emission levels, $M_1$ and $M_2$:

$$sols22 := \{A_1 = d_{11}M_1 + d_{12}M_2, A_2 = d_{21}M_1 + d_{22}M_2\}$$

To obtain specific values for the solutions, we now substitute our assumed particular values for the parameters, giving

$$\{A_1 = 50, A_2 = 51\}$$

# CHAPTER 7 Pollution control: instruments

Economists can only repeat, without quite understanding, what geologists, ecologists, public health experts, and others say about physical and physiological facts. Their craft is to perceive how economies and people in general will respond to those facts. **Dorfman (1985), p. 67**

**Learning objectives**

*After reading this chapter, the reader should understand*

- how bargaining processes might bring about efficient resource allocations (and so might lead to the attainment of efficient pollution outcomes without regulatory intervention)
- the conditions which limit the likelihood of bargaining solutions to pollution problems being achieved
- the instruments available to attain a pollution target
- the mechanisms by which pollution instruments operate in attaining targets
- the comparative merits of alternative instruments
- the significance, in instrument choice, of whether a pollutant is uniformly mixing

## Introduction

The previous chapter dealt with pollution targets. Here we consider how an environmental protection agency (EPA) could attain a predetermined pollution target by investigating the instruments that could be used.

In some circumstances no intervention would be required. Perhaps fortuitously, the prevailing level of pollution is not different from the target. Or intervention may be unnecessary because of the existence of voluntary bargaining. We show in Section 7.3 that bargaining between generators and victims of pollution could lead to an outcome in which the unregulated amount of pollution is equal to the pollution target. But we also show that such an outcome is unlikely for most important types of pollution problem. Where bargaining fails to reduce pollution to its targeted level, intervention of some form is called for.

This chapter is organised around three main themes. First, we describe the instruments that are available, and how each operates. Second, we provide a comparative assessment of those instruments. Finally, we consider whether there are particular circumstances – or particular types of pollution problems – which tend to favour the use of specific instruments. Of decisive importance is a matter raised in the previous chapter: whether or not the pollutant being targeted is uniformly mixing.

For the most part, our *analysis* will be quite general. That is, we will be thinking about instruments in the context of 'pollution problems' in general, rather than separately for air pollution, water pollution, soil contamination, and so on. However, the generality of the analysis will be limited in one important way. We will focus on pollution problems that are national (or sub-national) in scope, rather than on ones which are international. Control and regulation of *international* pollution problems will be addressed specifically in Chapter 10. The reason for segmenting the material in this way has nothing to do with the relative importance of different pollution problems. It is because dealing with international pollution issues brings another dimension into the picture: developing, coordinating and monitoring control across sovereign states. At this stage, we wish to keep this dimension out of our treatment.[1]

Although the analysis in this chapter is general in its scope, the examples and applications deal with specific contexts and case studies. Several applications not covered in this chapter – specifically instruments for conserving biological diversity, mobile source (transport) pollution, and agricultural pollution – are examined in the Word files *Biodiversity*, *Transport* and *Agriculture* in the *Additional Materials* for Chapter 7.

## 7.1 Criteria for choice of pollution control instruments

There are many instruments available to an EPA charged with attaining some pollution target. How should it choose from among these? If attaining the target were all that mattered, instrument choice would be relatively simple. The best instrument would be the one which meets the target with greatest reliability. But the EPA is unlikely to have only this objective. Government typically has multiple objectives, and the terms of reference that policy makers impose on their agents will tend to reflect that diversity of objectives. Even where these terms of reference are not explicit, the network of influences and pressures within which the EPA operates will lead it to adopt multiple goals *de facto*.

Instrument choice can be envisaged in the following way. Each available instrument can be characterised by a set of attributes, relating to such things as impacts on income and wealth distribution, the structure of incentives generated, and the costs imposed in abating pollution. A score can be given to each instrument, dependent on how well its attributes match with the set of objectives sought by the EPA. (A hypothetical example of this is explored in Problem 1 at the end of this chapter.) This perspective is useful as it draws attention to what kinds of attributes a ‘good’ instrument might have. Table 7.1 lays out a set of criteria in terms of which the relative merits of instruments can be assessed.

The brief descriptions in the right-hand column of the table should be sufficient to convey what the various criteria mean. Fuller definitions and explanations of the first five items will be given later in the chapter. The remaining four all relate, in some way or other, to decision making under conditions of limited information or uncertainty, and will be investigated in the next chapter. However, three observations about these criteria warrant mention now (and will be developed later).

*Table 7.1* Criteria for selection of pollution control instruments

| Criterion | Brief description |
|---|---|
| Cost-effectiveness | Does the instrument attain the target at least cost? |
| Long-run effects | Does the influence of the instrument strengthen, weaken or remain constant over time? |
| Dynamic efficiency | Does the instrument create continual incentives to improve products or production processes in pollution-reducing ways? |
| Ancillary benefits | Does the use of the instrument allow for a ‘double dividend’ to be achieved? |
| Equity | What implications does the use of an instrument have for the distribution of income or wealth? |
| Dependability | To what extent can the instrument be relied upon to achieve the target? |
| Flexibility | Is the instrument capable of being adapted quickly and cheaply as new information arises, as conditions change, or as targets are altered? |
| Costs of use under uncertainty | How large are the efficiency losses when the instrument is used with incorrect information? |
| Information requirements | How much information does the instrument require that the control authority possess, and what are the costs of acquiring it? |

[1] As you will see, our attempt to avoid dealing with the international dimension in this chapter will be compromised as soon as we get to grips with biodiversity. For that reason, it is taken up again in Chapter 10.

First, the use of any instrument is likely to involve conflicts or trade-offs between alternative criteria. Instrument choice will, therefore, depend on the relative weights attached to the criteria by the EPA. Second, it is likely that the weights (and so the choice of instrument) will vary over different types of pollution. For example, where a dangerous and persistent toxin is concerned, the EPA may regard cost efficiency as being of low importance relative to the long-run effect of the chosen instrument. Third, no single instrument is best for dealing with all types of pollution in all circumstances. We shall see in the next chapter that this is true *a fortiori* where instrument choice takes place under conditions of uncertainty. One particular criterion – cost efficiency – has received so much attention in the environmental economics literature that it warrants special attention now.

## 7.2 Cost efficiency and cost-effective pollution abatement instruments

Suppose a list is available of all instruments which are capable of achieving some predetermined pollution abatement target.[2] If one particular instrument can attain that target at lower real cost than any other can then that instrument is cost-effective.[3] Cost-effectiveness is clearly a desirable attribute of an instrument. Using a cost-effective instrument involves allocating the smallest amount of resources to pollution control, conditional on a given target being achieved. It has the minimum opportunity cost. Hence, the use of cost-effective instruments is a prerequisite for achieving an economically efficient allocation of resources.[4]

Let us explore some ramifications of the cost-effectiveness criterion. There will (usually) be many sources of an emission, and so many potential abaters. This raises the question of how the overall target should be shared among the sources. The principle of cost efficiency provides a very clear answer: a necessary condition for abatement at least cost is that the marginal cost of abatement be equalised over all abaters. This result is known as the least-cost theorem of pollution control. It is derived algebraically in the first part of Appendix 7.1. You will find it useful to read that now.

The intuition behind this result is easily found. Consider a situation in which marginal abatement costs were not equalised. For example, suppose that at present abatement levels two firms, A and B, have marginal abatement costs of 60 and 100 respectively. Clearly if B did one unit less abatement and A did one more (so that total abatement is unchanged) there would be a cost reduction of 40. Cost savings will accrue for further switches in abatement effort from B to A as long as it is more expensive for B to abate pollution at the margin than it is for A.

Let us examine these ideas a little further.[5] Suppose government wishes to reduce the total emission of a particular pollutant from the current, uncontrolled, level $\hat{M}$ (say, 90 units per period) to a target level $M^*$ (say, 50 units). This implies that the abatement target is 40 units of emission per period. Emissions arise from the activities of two firms, A and B. Firm A currently emits 40 units and B 50 units.

The following notation is used. The subscript $i$ indexes one firm (so here $i$ = A or B). $M_i$ is the actual level of the $i$th firm's emissions, which will depend on what control regime is in place. Two particular levels are of special interest. $\hat{M}_i$ is the profit-maximising level of emissions by firm $i$ in the absence of any controls set by government and in the absence of any pollution charges. $M_i^*$ is an emission ceiling (upper limit) set for the firm by the EPA. The quantity of pollution abatement by the $i$th firm is $Z_i$, given by $Z_i = \hat{M}_i - M_i^*$. Hence we assume that

[2] You will notice that we refer here to a pollution reduction (or abatement) target, rather than to a target level of pollution itself. This conforms to conventional usage in the literature on instruments. In this chapter, the context should make it clear whether the target being referred to relates to pollution or pollution abatement.

[3] Strictly speaking an instrument is cost-effective if its real resource cost is no greater than that of any other instrument available. This means that a cost-effective instrument may not be unique. For example, suppose that two instruments each incur costs of £10m to bring sulphur dioxide pollution down to some target level, while all others cost more than £10m. Then those two instruments are cost-effective.

[4] It is this which explains why the cost-effectiveness criterion has figured so prominently in the economics literature.

[5] The following problem is replicated in the Excel workbook *Leastcost.xls*, found in the *Additional Materials* for Chapter 7.

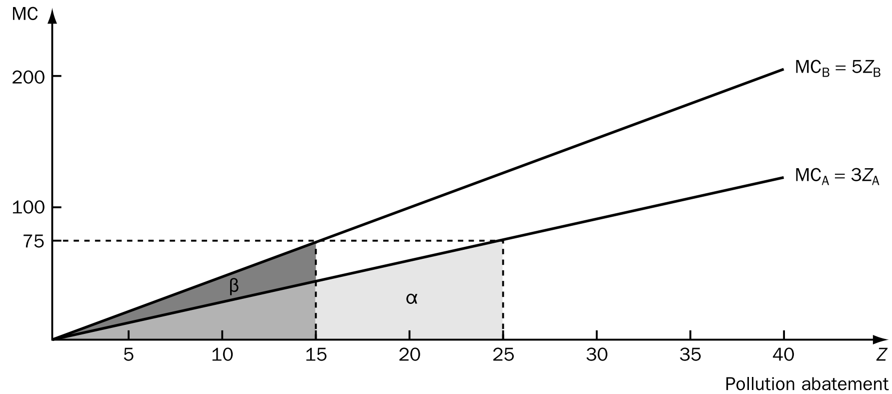


*Figure 7.1* Marginal abatement cost functions for the two firms

whenever an emissions regulation is in operation the amount of emissions the firm actually produces is that set by the EPA. $C_i$ is the total abatement cost of the $i$th firm.

Suppose that the total abatement cost functions of the two firms are $C_A = 100 + 1.5Z_A^2$ and $C_B = 100 + 2.5Z_B^2$. Therefore, the marginal abatement cost functions are $MC_A = 3Z_A$ and $MC_B = 5Z_B$. These are sketched in Figure 7.1. The least-cost solution is obtained by finding levels of $Z_A$ and $Z_B$ which add up to the overall abatement target $Z = 40$ *and* which satisfy the least-cost condition that $MC_A = MC_B$. This gives the answer $Z_A = 25$ and $Z_B = 15$. Figure 7.1 shows this least cost solution. At those respective abatement levels both firms have marginal abatement costs of 75. Minimised total abatement costs (1700) can be read from the diagram. The darker shaded area denoted β shows B's total abatement costs (662.5), while the lighter area denoted α represents A's total abatement costs (1037.5).

To verify this result, you could use the Lagrange multiplier technique, obtain the necessary first-order conditions, and solve these for the two firms' abatement levels. This was explained in the appendix to Chapter 4, where this problem – albeit with different numbers – was solved to show how the technique works. A convenient alternative, taking only a couple of minutes, is to use Excel's Solver routine to do this task for us. The mechanics of doing so are given in *Leastcost.xls* (in *Additional Materials*, *Chapter 7*) and you are recommended to study this Excel workbook now.

It is instructive to compare this solution with two others. First, one might think that as firm A has a lower marginal abatement cost schedule than B it should undertake all 40 units of abatement. It is easy to verify that this results in higher costs (2500) than those found in the least-cost solution (1700). Second, an equity argument might be invoked to justify sharing the abatement burden equally between the two firms. But it is easy to show (for example by looking at *Sheet1* of *Leastcost.xls*) that this also leads to higher costs (1800 in fact). If the regulator wanted such an equitable outcome, it would come at an additional real cost to the economy of 100 units (1800 – 1700). Note that the greater the difference in the firms' abatement cost functions, the greater would be the cost penalty from not pursuing the least-cost outcome. (See Problem 2.)

Some important conclusions emerge from this analysis:

- A least-cost control regime implies that the marginal cost of abatement is equalised over all firms undertaking pollution control.
- A least-cost solution will in general not involve equal abatement effort by all polluters.
- Where abatement costs differ, cost efficiency implies that relatively low-cost abaters will undertake most of the total abatement effort, but not all of it.

We shall use these results later in this chapter to establish whether particular kinds of pollution control instrument are cost-effective.

## 7.3 Instruments for achieving pollution abatement targets

In this section, we describe and explain the instruments available for pollution control. For convenience, the most common are listed in Table 7.2. Our emphasis is on the method of operation of each instrument and whether the instrument is cost-efficient. A more complete examination of the relative advantages of the instruments is left until later in the chapter.

### 7.3.1 Institutional approaches which facilitate internalisation of externalities

The various approaches to environmental policy that we consider in this section are best thought of not as pollution control instruments as such but rather as institutions which may avert the need to use pollution control instruments. Each shares the characteristic of potentially preventing the emergence of externalities, or internalising externalities which have arisen. In doing so, it is possible that decentralised behaviour by consumers and producers may generate efficient outcomes and so obviate the need for the regulatory intervention, at least if targets are set on efficiency grounds.

#### 7.3.1.1 Bargaining solutions and the limitations on bargaining solutions to environmental problems

The way in which bargaining can internalise externalities and so achieve efficient outcomes was explained in Chapter 5. There we considered an example of a musician disturbing a single neighbour, and how bargaining between those two parties could generate an efficient quantity of music playing. However, our discussion also demonstrated that efficient bargaining outcomes are often hard to obtain, and are sometimes impossible. These limitations are particularly likely for many kinds of environmental problem. We now briefly review why this should be so.

First, the likelihood of bargaining taking place is low unless enforceable property rights exist. For many environmental resources, well-defined and enforceable property rights do not exist. Second, bargaining is facilitated by the existence of a relatively small number of affected parties, and by all such parties being easily identifiable. Again, many environmental problems fail to satisfy either of those properties. Typically, environmental degradation affects many people and in many cases, as with vehicle pollution, is attributable to a large number of sources. It is often difficult to identify all affected parties, and the transactions costs associated with undertaking a bargaining exercise can be enormous. Hence where the number of affected individuals is large, the scope for efficient bargaining behaviour is very restricted.

Another pertinent issue relates to the possibility of intertemporal bargaining, including bargaining between current and future generations. Many environmental externalities cut across generations – our behaviour today imposes externalities on future persons. While bargaining between affected individuals at one point in time seems feasible, it is difficult to imagine that this could happen between representatives of the present generation and those not yet living. One would not, therefore, expect that bargaining between directly affected individuals and firms would offer much prospect of bringing about an efficient response to global climate change, involving as it does many generations.

Finally, bargaining solutions are extremely unlikely to be able to bring about socially efficient provision or conservation of public goods. Given that a substantial proportion of natural resources – or the services that they yield – have public good characteristics, this is a profound limitation.

What do these observations imply about the role for government? If, despite these limitations, bargaining does offer the prospect of substantial efficiency gains, then government should facilitate it wherever that is cost-effective. It could do so by clearly defining and explicitly allocating property rights wherever that is practicable (and ethically acceptable). Government might also seek to develop and sustain an institutional structure that maximises the scope for bargaining behaviour, as is sometimes done for employment disputes. Gains may derive from government's taking some responsibility for environmental monitoring so as to identify pollution producers and recipients, and disclosing information from this to affected parties. Finally, access to the judicial system should be easy and cheap. This will

*Table 7.2* Classification of pollution control instruments

| Instrument | Description | Examples |
|---|---|---|
| ***Institutional approaches to facilitate internalisation of externalities*** | | |
| Facilitation of bargaining | Cost of, or impediments to, bargaining are reduced | Polluter information placed in the public domain |
| Specification of liability | Codification of liability for environmental damage | Respiratory damage in Japan |
| Development of social responsibility | Education and socialisation programmes promoting citizenship | Energy-conservation media campaigns<br>Environmental labelling |
| ***Command and control instruments*** | | |
| Input controls over quantity and/or mix of inputs | Requirements to use particular inputs, or prohibitions/restrictions on use of others | Bans on use of toxic cleansing agents |
| Technology controls | Requirements to use particular methods or standards | Requirement to install catalytic converters in exhausts. BATNEEC |
| Output controls: | | |
| Output quotas or prohibitions | Non-transferable ceilings on product outputs | Ban on use of DDT<br>Singapore: vehicle quotas<br>Effluent discharge licences |
| Emissions licences | Non-transferable ceilings on emission quantities | |
| Location controls (zoning, planning controls, relocation) | Regulations relating to admissible location of activities | Heavy industry zoning regulations |
| ***Economic incentive (market-based) instruments*** | | |
| Emissions charges/taxes | Direct charges based on quantity and/or quality of a pollutant | Air pollution charges (e.g. $NO_X$ charges in France and Sweden; $SO_2$ charges in France and Japan)<br>Carbon/energy taxes<br>Water effluent charges (evidence of effectiveness in Germany, Netherlands and Malaysia)<br>Noise pollution charges (Belgium, France, Germany, Japan, Netherlands, Norway, Switzerland)<br>Fertiliser and pesticide taxes (Austria, Belgium, Scandinavian countries) |
| User charges/fees/natural resource taxes | Payment for cost of collective services (charges), or for use of a natural resource (fees or resource taxes) | User charges on municipal waste collection, treatment or disposal<br>Hazardous waste, wastewater user, and aircraft noise charges<br>Water extraction charges (thought to be effective in several Asian countries)<br>Congestion pricing (France, Norway, Singapore, USA) |
| Product charges/taxes | Applied to polluting products | Hungary: vehicle tyres<br>Finland: nuclear waste<br>Italy: plastic bags<br>Belgium: disposables tax |
| Emissions abatement and resource management subsidies | Financial payments designed to reduce damaging emissions or conserve scarce resources | Quebec: subsidy for energy generated from waste<br>Norway: grants to ecological farming |
| Marketable (transferable, marketable) emissions permits | Two systems: those based on emissions reduction credits (ERCs) or cap-and-trade | Denmark: $CO_2$ emissions from power plants |
| Deposit-refund systems | A fully or partially reimbursable payment incurred at purchase of a product | Austria: refillable plastic bottles<br>Quebec: one-way beer and soft-drink bottles<br>Also used in Korea, Greece, Norway and Sweden |
| Non-compliance fees | Payments made by polluters or resource users for non-compliance, usually proportional to damage or to profit gains | Greece: car emissions<br>Sweden: sea dumping of oil from ships |
| Performance bonds | A deposit paid, repayable on achieving compliance | Australia: mine sites<br>US: open pits |
| Liability payments | Payments in compensation for damage | Japan: waste – restoration of sites polluted by illegal dumping |

Notes to table:

1. Many of the examples in the table are drawn from OECD (1999) and EPA (1999). These references are available online, the first via the OECD web site www.oecd.org, the second at http://yosemite1.epa.gov/ee/epalib/incent.nsf. They provide extensive accounts of incentive-based environmental controls used in OECD countries.
2. Particular countries are mentioned purely as examples. Listings are not exhaustive.

also facilitate use of the liability principle that we shall discuss in the next section.

Nevertheless, the limitations to bargaining that we have described do appear to be very substantial, and it would be inappropriate to place too much reliance on such a mechanism. There is one important exception to this conclusion, however. When it comes to dealing with pollution, or other environmental, problems that spill over national boundaries, the absence of supra-national sovereign institutions means that there is often little or no alternative to bargaining solutions. These are unlikely, of course, to take place directly between affected individuals or firms. Rather, international policy coordination and cooperation is negotiated between representatives of affected national governments.

Discussions about greenhouse gas emissions or about the maintenance of biological diversity are two of the more well-known examples of such international bargaining processes, and have the potential to generate massive collective benefits. As international policy cooperation about environmental problems is the subject of a separate chapter (Chapter 10), we shall postpone further consideration of this matter until then.

#### 7.3.1.2 Liability

The role that may be played by the judicial system in helping to bring about efficient outcomes has been implicit in our discussion of bargaining. But that role can be taken a step further. Suppose that a general legal principle is established which makes any person or organisation liable for the adverse external effects of their actions. Then any polluter knows that there is some probability, say $p$, of being identified and successfully prosecuted, and so made to pay for that pollution. One variant of this scheme has the prosecuted polluter paying $p$ times the value of the damages, so that the expected value of the liability equals the value of pollution damage.[6]

The liability principle is related to property rights. Where pollution is a private good, the liability is equivalent to a statement of enforceable property rights vested in the victims, and enforcement would be done through civil law. But where the pollutant is a public good, this way of making the polluter pay is not usually feasible. In that case, the EPA acts as an agent of the public interest, enforcing the liability principle on behalf of affected parties. An interesting question is whether any damages obtained in this way should be returned to individuals as compensation. We explore this matter in Discussion Question 1.

Using the liability principle is not without its problems. One difficulty arises where damage only becomes apparent a long time after the relevant pollutants were discharged. Tracking down those who are liable may be a substantial undertaking, and those responsible – individuals or firms – may no longer exist. An interesting development is the process of establishing legal liability throughout the life cycle of a product, using the principle that producers are responsible for damage from 'cradle to grave'.

#### 7.3.1.3 Development of social responsibility

Pollution problems happen, in the final analysis, because of self-interested but uncoordinated, or sometimes thoughtless, behaviour. Encouraging people to behave as responsible citizens can help to attain environmental goals. Clearly, the government of the day has limited influence over the cultural context of human behaviour. But it would be wrong to ignore the opportunities that exist for using educational institutions and the mass communications media to help achieve specific targets and to promote ethical behaviour.

The evidence that individuals do not exclusively act in a narrowly utilitarian way suggests that this objective may be more than just wishful thinking. Among the very many examples that could be cited are support for green parties and the increasing importance being given to environmental issues by voters, the success of some ethical investment funds, our willingness to support charities. Perhaps the strongest evidence is to be found in our family and social lives, where much of what we think and do has a social – rather than purely self-interested – basis. Although we write little about 'cultural' instruments in this text, the authors recognise that they may be the most powerful ways of achieving general environmental goals.

[6] It is important to note, however, that damages may be assessed differently by a court from the way we have in mind, and so the liability principle may generate different outcomes from the 'efficient outcomes' achieved through bargaining.

**Box 7.1 Liability for environmental damage**

An important example of the liability for damage principle can be found in the regulations relating to hazardous waste disposal in the USA. Under the terms of the Resource Conservation and Recovery Act, a 'cradle-to-grave' tracking and liability principle has been adopted.

The *Superfund* concerns abandoned waste dumps. The fund is built up from various sources including damages settlements. The principle of 'strict, joint and several liability' establishes a special form of retrospective liability, in which parties that have dumped waste (legally or illegally) can be sued for the whole costs of clean-up, even though they were only partial contributors to the dump. The sued party may then attempt to identify others responsible to recover some of the damages. Moreover, liability lies with the generators of waste as well as those who subsequently reprocess or dispose along the waste cycle.

The use of liability payment schemes is now widespread, with examples to be found in Quebec, Denmark, Finland, Germany, Japan, Sweden and Turkey. Several countries have instituted general liability schemes (e.g. Denmark, Finland, Sweden and Turkey), in some cases requiring compulsory environmental damage insurance for large polluters (e.g. Finland). Other governments have specified liability schemes for particular categories of polluter (Quebec – tioxide (titanium dioxide) pollution; Germany – noise; USA – hazardous waste).

Since the 1970, Japanese courts have developed an extensive liability case law, relating primarily to waste, air and water pollution. Japanese businesses contribute to a compensation fund. Until 1988, persons with bronchial asthma and other respiratory diseases were entitled to compensation from the fund without judicial procedure. After 1988, new sufferers were no longer entitled to automatic compensation, as air pollution was no longer unequivocally accepted as the principal contributory factor to respiratory illnesses.

*Source*: OECD (1999)

One particular policy mechanism which could be said to be in the 'social responsibility' category is environmental labelling, used in virtually all industrialised economies and in many developing countries. This has been credited with reducing VOC (volatile organic compound) emissions in Germany, and with increasing paper recycling in Korea (EPA, 1999).

### 7.3.2 Command and control instruments

The dominant method of reducing pollution in most countries has been the use of direct controls over polluters. This set of controls is commonly known as *command and control* instruments. Figure 7.2 provides a schema by which these instruments can be classified. There we see that the regulations can be classified in terms of what is being targeted.

The first panel (Figure 7.2a) represents the various relationships that link production to pollution levels. Emissions are by-products in the production of intended final output. The amount (and type) of emissions will depend on which goods are being produced, and in what quantities. It will also depend on the production techniques being employed, and on the amount (and mix) of inputs being used. For uniformly mixing pollutants (UMPs), pollution levels will depend only on total emissions levels. In the case of non-uniformly-mixing pollutants (indicated in the diagram by the dotted lines in the branch to the right) the spatial distribution of ambient pollution levels will also depend on the location of emission sources.

Command and control instruments can be designed to intervene at any of these stages. So, as the second panel (Figure 7.2b) illustrates, regulations may apply to outputs of emissions themselves, to the quantity of final production, to production techniques used, or to the level and/or mix of productive inputs. For non-UMPs, controls may also apply to location of emission sources.

In general, there should be advantages in directing the controls at points closest (in this sequence of linkages) to what is ultimately being targeted: that is, ambient pollution levels. This allows polluters the most flexibility in how a pollution reduction is to be achieved. But it may not always be feasible – or desirable on other grounds – to set regulations in that way.

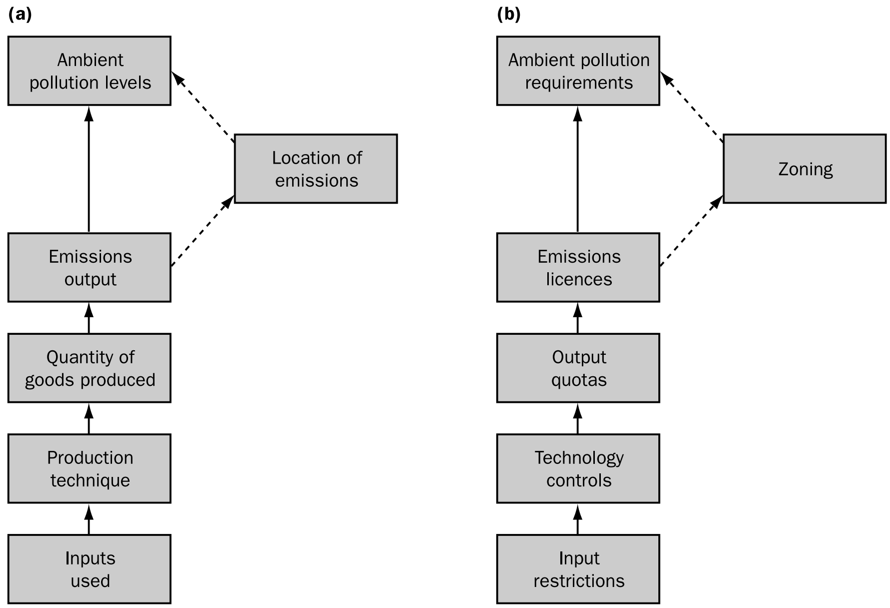


*Figure 7.2* A classification of command and control instruments

There is huge variation from place to place in the detail of regulatory systems. It would be of little use – and probably impossible – to list the plethora of command and control regulations. Our coverage of command and control instruments will be limited, therefore, to some general comments on the main categories shown in Figure 7.2b, together with some illustrative examples in boxes. For further detail, the reader should visit the text's Accompanying Website www.booksites.net/perman, which provides links to many sites that provide regularly updated accounts of regulatory regimes in various countries.

Some examples of the use of command and control in the USA are given in Box 7.2. The material

**Box 7.2 Environmental protection in the USA**

The United States system of environmental controls is one of the most comprehensive to be found. A set of Congressional statutes provides the legal framework for the regulatory system, and give responsibility to the United States Environmental Protection Agency (US EPA) for implementing and administering the system. A comprehensive, and well-indexed, account of US environmental policy can be found on the 'Laws & regulations' section of the US EPA web site (www.epa.gov/epahome/lawreg.htm). Here we focus on a small, but important, part of that system.

Table 7.3 outlines the regulatory framework in six particular areas: air and water pollution, hazardous waste disposal, agricultural chemicals, toxic substances, and species protection. It identifies the regulatory area in each case, and states the criteria that must be considered by US EPA in setting standards.

**Air quality**

The *Clean Air Act* defines ambient air quality standards for all parts of the USA for two types of pollutant: criteria (common) and hazardous air pollutants. *Criteria* air pollutants consist of particulates, $SO_2$, CO, $NO_2$, low-level ozone and lead. Each of these is given a *primary NAAQS* (National Ambient Air Quality Standard), set to protect human health. Some are also given a *secondary* NAAQS to protect wildlife, visibility and ecological systems. The levels of NAAQS for the criteria pollutants were listed in Table 6.7 in Chapter 6.

The system for *criteria* air pollutants is as follows. For *stationary sources* of air pollutants,

**Box 7.2 *continued***

the principal control instrument is technology-based regulation. This is supported by maximum allowable emissions rates in some cases. *Existing pollution sources* must satisfy 'reasonably available control technology' (RACT). *New pollution sources* must meet more restrictive 'new source performance standards' (NSPS), based on the criterion of commercially available 'best technological system of emissions reduction'. Where NAAQS have not been met, stricter criteria may be used, such as 'lowest achievable emissions rate' (LAER), or in Class 1 (unspoilt) areas 'best available control technology' (BACT). What counts as satisfying these requirements is often laid out in great detail by US EPA after thorough study of particular production processes. Firms may be required to use particular techniques to recover fumes or waste products, or they may be prohibited from using certain production processes. Not surprisingly, the interpretation of these different criteria and the particular requirements that US EPA mandates for firms, are contentious, and lead to significant amounts of judicial action.

For *mobile source* air pollution, control is largely directed at vehicle manufacturers, again in the form of required technology controls. Stricter controls are used in some non-attainment areas (such as mandated use of low-polluting fuels).

Although air pollution is mainly controlled by technology-based regulation, there are some exceptions. A flexible incentive-based system has been developed for acid-rain-inducing pollutants, and will be examined in Chapter 10. Individual states may also, if they wish pursue higher than national standards. Some states are experimenting with various market-based controls, such as those being used in the Los Angeles basin area.

In the cases of *hazardous air pollutants* (about 200 air toxins listed by US EPA other than the criteria pollutants), 'large' stationary sources must use 'maximum achievable control technology' (MACT). Additional control

*Table 7.3* Factors to be considered by the US EPA in setting standards and regulations

| Statute | Coverage | Factors to be considered in setting standards |
|---|---|---|
| Clean Air Act (CAA) (as amended 1990) www.epa.gov/oar/oaqps/peg_caa/pegcaain.html | Ambient air quality standards | Standards to be set on safety grounds (to achieve an 'adequate margin of safety')<br>US EPA must consider benefits of regulation but *not* costs |
| Clean Water Act (CWA) 1987 (in conjunction with Federal Water Pollution Control Act) www.epa.gov/region5/defs/html/cwa.htm | Effluent emissions, from stationary point sources and non-point sources | Standards to be set on safety grounds<br>Waters required to be at least 'swimmable and fishable'<br>US EPA must consider benefits and costs of regulation (but balancing is not required) |
| The Resource Conservation and Recovery Act (RCRA)<br>Comprehensive Environmental Response, Compensation, and Liability Act (Superfund)<br>The Superfund Amendments and Reauthorization Act (SARA)<br>www.epa.gov/superfund | Hazardous waste disposal on land, both current disposal (RCRA) and abandoned waste dumps (Superfund) | Standards to be set on safety grounds<br>US EPA must consider benefits of regulation but *not* costs |
| The Endangered Species Act (ESA) | | Ecological sustainability standard<br>Protection of species at any cost |
| Federal Insecticide, Fungicide and Rodenticide Act (FIFRA)<br>Toxic Substances Control Act (TSCA)<br>Emergency Planning and Right-to-Know Act | Restricting the use of dangerous substances<br>FIFRA: agricultural chemicals<br>TSCA: toxic substances | Targets to be set on efficiency grounds<br>Benefits and costs of regulation to be balanced in both cases |

**Box 7.2 *continued***

measures, and/or new listed pollutants, may be added by the US EPA if risk analysis suggests that this is warranted. With the passage of time, US EPA has gone some way along the process of defining acceptable risk in operational terms. For example, 'ample margin of safety' is now defined to mean that cancer risks to the most exposed population do not exceed 1 in 10 000. (The long-term target has been specified as 1 in 1 million for the population at large.)[7] Note from Table 7.3 that the Clean Air Act requires the US EPA to only take account of the benefits of control in setting regulations over toxic air emissions. However, in 1987 a Court of Appeals ruling found that US EPA has been (unlawfully) considering both benefits and costs in setting ambient standards. As a result, US EPA tightened its standards (so that control was extended to cover emissions for which it previously felt that the cost-to-benefit ratio was too high to justify control).

**Clean water**

Water standards are again typically based on technology controls. In the initial control phase, this required the use of 'best *practical* technology' (BPT). Later control phases mandated the more stringent 'best *available* technology' (BAT). In addition to BAT, dischargers must acquire (non-marketable) effluent emissions licences, often containing very detailed plans about how discharges are treated as well as the amounts that may be discharged. What counts as 'best' is defined by US EPA (although, again, not without much judicial challenge). Technology controls ('best-management practices' are also employed to reduce runoff from non-point sources (industrial and agricultural sites).

**Hazardous waste disposal**

Under the terms of the Resource Conservation and Recovery Act, the US EPA has developed a list of about 450 hazardous substances. Disposal is controlled through location restrictions, required staff training, groundwater monitoring by disposing firms, and the requirement to construct detailed plans for site closure and post-closure practice. Operators must also undertake sufficient insurance cover. These, and other, restrictions are supported by a licence system. An interesting innovation here is the adoption of a 'cradle-to-grave' tracking and liability principle (see Box 7.1). The Superfund has provided a mechanism for dealing with abandoned waste dumps. The fund is built up from general taxation and from taxes on the petroleum and chemical industries. The principle of 'strict, joint and several liability' (see Box 7.1) establishes strong incentives throughout the waste cycle to minimise the amount of waste produced.

**Toxic substances**

The TSCA requires US EPA to review all new chemicals, and gives it authority to restrict use of, or ban, existing chemicals. Unlike most areas of environmental regulation, the TSCA requires balancing of the costs of regulation (in money terms) and the benefits of regulation (in terms of cancer or other serious health impacts avoided). A study by Van Houtven and Cropper (1996) investigated US EPA bans on the use of asbestos in particular uses under the provisions of the TSCA. Of the 39 uses of asbestos it investigated, Van Houtven and Cropper found that US EPA was able to measure costs and benefits in 31 cases. Of these, 21 products were banned.

**Agricultural chemicals**

FIFRA imposes a duty of registration of all new pesticides. New ingredients in agricultural chemicals cannot be introduced until the US EPA is satisfied, after cost–benefit analysis, that the product will generate positive net benefits. As an input to this study, manufacturers must submit a detailed scientific study of the ingredient. US EPA may also carry out Special Reviews on existing pesticides. As with TSCA, FIFRA requires that the EPA 'balance' benefits against costs in arriving at its decisions about bans or other restrictions. The Van Houtven and Cropper study investigated 245 food crop applications of 19 pesticide active ingredients. Of these, 96 applications were banned after US EPA Special Reviews.

[7] Actual risks have often been very much higher. A US EPA study in the late 1980s revealed that risks were worse than 1 in 1000 in 205 communities around the country.

there shows that the administration of instruments is not entirely separate from the setting of targets (or 'standards' as they are also known). In the examples, the 'goal' passed on to the US Environmental Protection Agency (USEPA) is given in the form of a general principle regarding the criterion that should be used in setting standards, together with some direction about what information should be used in its deliberations. The USEPA is then required to translate that goal into specific targets and/or regulations and to administer their implementation.

In the sections that follow, we describe in a little more detail the three most commonly used types of command and control instrument, and then investigate instruments that use the price mechanism to create incentives for pollution abatement. In doing so, the likely cost-efficiency of each instrument will be discussed. A more complete appraisal of the relative merits of each instrument using the criteria listed in Table 7.1 will be left until Section 7.6.

#### 7.3.2.1 Non-transferable emissions licences

Suppose that the EPA is committed to attaining some overall emissions target for a particular kind of pollutant. It then creates licences (also called permits or quotas) for that total allowable quantity. After adopting some criterion for apportioning licences among the individual sources, the EPA distributes licences to emissions sources. We use the term *non-transferable licences* to refer to a system where the licences cannot be transferred (exchanged) between firms: each firm's initial allocation of pollution licences sets the maximum amount of emissions that it is allowed.[8] Successful operation of licence schemes is unlikely if polluters believe their actions are not observed, or if the penalties on polluters not meeting licence restrictions are low relative to the cost of abatement. Licence schemes will have to be supported, therefore, by pollution monitoring systems and by sufficiently harsh penalties for non-compliance.

Under special conditions, the use of such emissions licences will achieve an overall target at least cost (that is, be cost-efficient). But it is highly unlikely that these conditions would be satisfied. We know (see the first part of Appendix 7.1) that cost-efficiency requires the marginal cost of emissions abatement to be equal over all abaters. If the EPA knew each polluter's abatement cost function, it could calculate which level of emissions of each firm (and so which number of licences for each firm) would generate this equality *and* meet the overall target.

It is very unlikely that the EPA would possess, or could acquire, sufficient information to set standards for each polluter in this way. The costs of collecting that information could be prohibitive, and may outweigh the potential efficiency gains arising from intervention. Moreover, there is a problem of information asymmetries; those who possess the necessary information about abatement costs at the firm level (the polluters) do not have incentives to provide it in unbiased form to those who do not have it (the regulator).[9] We examine these incentives in a little more detail in Section 7.6. A system of long-term relationships between regulator and regulated may overcome these asymmetries to some extent, but might bring other problems (such as high administrative cost and regulatory capture – to be defined and explained in Chapter 8) in its wake. Given all this, it seems likely that arbitrary methods will be used to allocate licences, and so the controls will not be cost-efficient. Box 7.10 gives some indication of how great this cost-inefficiency is in practice.

#### 7.3.2.2 Instruments which impose minimum technology requirements

Another command and control approach involves specifying required characteristics of production processes or capital equipment used. In other words, minimum technology requirements are imposed upon potential polluters. Examples of this approach have been variously known as *best practicable means* (BPM), *best available technology* (BAT) and *best available technology not entailing excessive cost* (BATNEEC). Some further information on technology controls is given in Box 7.3.

[8] We use the term 'licence' to denote non-transferable emissions quotas. Later in the chapter, transferable quotas will be discussed. To avoid confusion, we call these 'permits'.

[9] Another possibility is that firms themselves may also be unaware of their abatement costs.

**Box 7.3 Required technology controls**

Regulations mandating the use of particular technologies are common forms of pollution control instrument in Europe, North America and the other OECD countries. In the UK, a criterion underlying required technology standards has been 'best practicable means'. The adjective *practicable* has never been given a precise legal definition, but the 1956 Clean Air Act stated that

> Practicable means reasonably practicable having regard, amongst other things, to local conditions and circumstances, to the financial implications and the current state of technology.

Despite an element of tautology in this statement, it can be interpreted as meaning that a practicable control technology should be technologically effective, subject to the constraint that it is not excessively costly. In recent years, the cost qualification has been given greater priority, and has been enshrined in the principle of BATNEEC: the best available technology not entailing excessive cost. This puts the instrument closer to the kind advocated by economists, as the 'excessive cost' condition implies a quasi-cost–benefit calculation in the administration of the instrument.

However, while the cost of control is often measured by the regulator in money terms (for example, the additional money cost of one technique over another), the benefits are not usually measured in money terms; instead, benefits are seen in terms of reduced probabilities of death or serious damage to health. In this sense, although some balancing of costs against benefits does often take place, the approach being used is not 'cost–benefit analysis' in the economics sense of that term. Rather than using the public's estimate of benefits (in terms of willingness to pay) the regulator has to come to a view as to what cost is reasonable to save a life or reduce a health risk. Some information on this is provided in Box 7.3a. Equivalent kinds of money-cost relative to health-benefit comparisons are also made in the US regulatory system.

The manner in which technology-based instruments have been implemented varies considerably between countries. In the UK, officials of the Inspectorate of Pollution negotiate controls with plant managers, but have the right, in the last instance, to require the adoption of certain control technologies. The United States Environmental Protection Agency administers a rather more uniform control programme: in 1990, Congress required the EPA to establish technology-based standards for about 200 specific pollutants.

**Box 7.3a The value of life, as revealed by actions of the United States Environmental Protection Agency**

A series of recent papers have attempted to deduce what value the United States Environmental Protection Agency (US EPA) places on saving lives. In one of these papers, Van Houtven and Cropper (1996) examined four particular areas over which the US EPA has sought to achieve regulation. We noted in Box 7.2 that the US EPA can issue bans on particular uses of asbestos. Van Houtven and Cropper investigated 39 applications for asbestos use. From data on the costs of regulation and the number of lives expected to be saved in each application, the authors were able to estimate the value of a statistical life that is implied by US EPA decisions. By definition, if an action results in the expected level of deaths falling by one person over some relevant time period, that action has saved one statistical life. Van Houtven and Cropper found that, on average, products were banned when the cost of saving one life was below $49 million (in 1989 US dollar prices).

Van Houtven and Cropper obtain a very similar implied value ($51.51) million for a fatal cancer avoided in their study of 245 pesticide applications (of which 96 were banned). Decisions here were taken under the auspices of FIFRA agricultural chemicals legislation, which also requires cost and benefit balancing to be used by the US EPA. Van Houtven and Cropper also investigated controls of toxic air pollutants – specifically benzene, arsenic, asbestos and mercury – under the provisions of the Clean Air Act. Prior to 1987, the implied value of a fatal

**Box 7.3a *continued***

cancer avoided was about $16 million. As we remarked earlier, a 1987 Court of Appeals ruling found that EPA has been unlawfully considering the costs of regulation in making its decisions. In so doing, some emissions had been allowed where the US EPA had estimated the cost-to-benefit ratio to be too high to justify control. The tighter standards the US EPA subsequently imposed (based only on the benefits of control) implied a value of a statistical life after 1987 of $194 million.

These values are considerably higher than the values which people seem to be willing to pay to reduce the risk of death. For example, Viscusi (1992, 1993) estimated the compensating wage differential required by workers to take on high-risk jobs. Observed wage differentials imply a value of a statistical life of $5 million, just one-tenth of that implied by US EPA regulations that required balancing.

In the previous edition of this text, we remarked that the US EPA appeared to be using the principle of cost-effectiveness in making decisions. For example, that would entail that, for any given sized health benefit, those products with the lower control costs are banned while those with higher costs are not banned. But other research suggests that this is questionable. For example, Viscusi (1996) examines a number of command and control regulations designed to save lives and protect health. Table 7.4 shows the costs of a statistical life saved for each category of regulation. Huge variability is evident, although some of this reflects differences in what the US EPA is required to consider in making decisions (that is: just benefits, benefits and costs but without balancing, or benefits and costs with balancing).

Another example of widely varying marginal costs is given in a study by Magat *et al.* (1986) of the marginal treatment cost of biological oxygen demand (BOD) from US rivers and lakes. The authors estimated that marginal costs of attaining regulatory standards varied from as little as $0.10 per kilogram of BOD removal to as much as $3.15.

In the case of both BOD removal and reduction of the risk of death, there appear to be very large efficiency gains possible from reallocating control (and so control expenditures) from high-cost to low-cost areas.

*Table 7.4* The statistical value of a life as revealed by US EPA command and control regulations

| Regulation | Initial annual risk | Expect annual lives saved | Cost per expected life saved ($US 1984) |
|---|---|---|---|
| Unvented space heaters | 2.7 in $10^5$ | 63.000 | 0.10 |
| Airplane cabin fire protection | 6.5 in $10^8$ | 15.000 | 0.20 |
| Auto passive restraints/belts | 9.1 in $10^5$ | 1850.000 | 0.30 |
| Underground construction | 1.6 in $10^3$ | 8.100 | 0.30 |
| Servicing wheel rims | 1.4 in $10^5$ | 2.300 | 0.50 |
| Aircraft seat cushion flammability | 1.6 in $10^7$ | 37.000 | 0.60 |
| Aircraft floor emergency lighting | 2.2 in $10^8$ | 5.000 | 0.70 |
| Crane suspended personnel platform | 1.8 in $10^3$ | 5.000 | 1.20 |
| Concrete and masonry construction | 1.4 in $10^5$ | 6.500 | 1.40 |
| Benzene/fugitive emissions | 2.1 in $10^5$ | 0.310 | 2.80 |
| Grain dust | 2.1 in $10^4$ | 4.000 | 5.30 |
| Radionuclides/uranium mines | 1.4 in $10^4$ | 1.100 | 6.90 |
| Benzene in workplace | 8.8 in $10^4$ | 3.800 | 17.10 |
| Ethylene oxide in workplace | 4.4 in $10^5$ | 2.800 | 25.60 |
| Arsenic/copper smelter | 9.0 in $10^4$ | 0.060 | 26.50 |
| Uranium mill tailings, active | 4.3 in $10^4$ | 2.100 | 53.00 |
| Asbestos in workplace | 6.7 in $10^5$ | 74.700 | 89.30 |
| Arsenic/glass manufacturing | 3.8 in $10^5$ | 0.250 | 142.00 |
| Radionuclides/DOE facilities | 4.3 in $10^6$ | 0.001 | 210.00 |
| Benzene/ethylbenzenol styrene | 2.0 in $10^6$ | 0.006 | 483.00 |
| Formaldehyde in workplace | 6.8 in $10^7$ | 0.010 | 72000.00 |

*Source*: Viscusi (1996), pp. 124–125

In some variants of this approach, specific techniques are mandated, such as requirements to use flue-gas desulphurisation equipment in power generation, designation of minimum stack heights, the installation of catalytic converters in vehicle exhaust systems, and maximum permitted lead content in engine fuels. In other variants, production must employ the (technically) best technique available (sometimes subject to a reasonable cost qualification). The specific technique adopted is sometimes negotiated between the EPA and the regulated parties on an individual basis.

Much the same comments about cost-effectiveness can be made for technology controls as for licences. They are usually not cost-efficient, because the instrument does not intrinsically focus abatement effort on polluters that can abate at least cost. Moreover, there is an additional inefficiency here that also involves information asymmetries, and which relates back to a point made earlier about Figure 7.2. Technology requirements restrict the choice set allowed to firms to reduce emissions. Decisions about emissions reduction are effectively being centralised (to the EPA) when they may be better left to the firms (who will choose this method of reducing emissions rather than any other only if it is least-cost for them to do so).

Required technology controls blur the pollution target/pollution instrument distinction we have been drawing in this and the previous chapter. The target actually achieved tends to emerge jointly with the administration of the instrument. We need to be a little careful here. Sometimes government sets a general target (such as the reduction of particulates from diesel engines by 25% over the next 5 years) and then pursues that target using a variety of instruments applied at varying rates of intensity over time. In this case, no single instrument need necessarily have a particular target quantity associated with it. Nevertheless, it does matter (as far as cost-efficiency is concerned) if the actual operation of any particular component of this programme does not involve any comparison of the benefits and costs of that component (because then the wrong mix of components will be used). There are many examples of technology control where it appears to be the case that emphasis is given almost exclusively to the costs of pollution reduction technologies, and in particular to what kind of cost premium is involved in using the technically best method as compared with its lower-ranked alternatives. (See Box 7.10, for example. And think about saving lives via safety regulations.)

Although technology-based instruments may be lacking in cost-effectiveness terms, they can be very powerful; they are sometimes capable of achieving large reductions in emissions quickly, particularly when technological 'fixes' are available but not widely adopted. Technology controls have almost certainly resulted in huge reductions in pollution levels compared with what would be expected in their absence.

#### 7.3.2.3 Location

Pollution control objectives, in so far as they are concerned only with reducing human exposure to pollutants, could be met by moving affected persons to areas away from pollution sources. This is only relevant where the pollutant is not uniformly mixing, so that its effects are spatially differentiated. Implementing this *ex ante*, by zoning or planning decision, is relatively common. *Ex post* relocation decisions are rarer because of their draconian nature. There have been examples of people being removed from heavily contaminated areas, including movements away from irradiated sites such as Chernobyl, Times Beach (Missouri) and Love Canal (New York). However, it has been far more common to move pollution sources away from areas where people will be affected, or to use planning regulations to ensure separation. Planning controls and other forms of direct regulation directed at location have a large role to play in the control of pollution with localised impacts and for mobile source pollution. They are also used to prevent harmful spatial clustering of emission sources.

Location decisions of this kind will not be appropriate in many circumstances. Moving people away from a pollution source cannot, for example, reduce impacts on ecosystems. Relocating (or planning the location of new) emission sources has wider applicability, but will be of no use in cases where pollution is uniformly mixing. In Section 7.5 we shall consider a number of incentive-based instruments that are designed, among other things, to influence the spatial location of emissions sources. These are not, however, examples of command and control instruments.

## 7.4 Economic incentive (quasi-market) instruments

Command and control instruments operate by imposing mandatory obligations or restrictions on the behaviour of firms and individuals. Incentive-based instruments work by creating incentives for individuals or firms to voluntarily change their behaviour. These instruments alter the structure of pay-offs that agents face.

Employing incentives to make behaviour less polluting can be thought about in terms of prices and markets. Taxes, subsidies and transferable permits create markets (or quasi-markets, something equivalent to markets) for the pollution externality.[10] In these markets, prices exist which generate opportunity costs that profit-maximising firms will take account of in their behaviour.

### 7.4.1 Emissions taxes and pollution abatement subsidies

In this section, we examine tax and subsidy instruments used to alter the rate of emissions of uniformly mixed pollutants, for which the value of the damage created by an emission is independent of the location of its source. It is shown later that the results also apply, with minor amendment, to non-uniformly mixing pollutants. Given that taxes on emissions are equivalent to subsidies (negative taxes) on emissions abatement, it will be convenient to deal explicitly with tax instruments, and refer to subsidy schemes only when there is a difference that matters.

Looking again at Figure 7.2, it is evident that there are several points at which a tax could be applied (just as there were several points of intervention for command and control regulations). We focus here on taxation of emissions. It is important to note that taxes on output of the final product, or on the levels of particular inputs (such as coal), will not have the same effect as emissions taxes, and will generally be less efficient in attaining pollution targets. This matter is examined in Problem 9 at the end of the chapter.

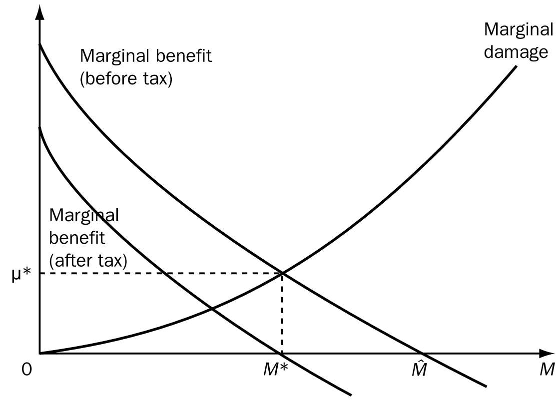


*Figure 7.3* An economically efficient emissions tax

A tax on pollutant emissions has for long been the instrument advocated by economists to achieve a pollution target. It is useful to distinguish between three cases:

1. the pollution target is the economically efficient level of pollution (the level which maximises social net benefits);
2. a specific target is sought, but it is set according to some criterion other than economic efficiency;
3. an emission reduction of some unspecified amount is sought.

We deal with each of these cases in turn. To attain the *efficient* level of pollution, it is necessary to have solved the net benefit maximisation problem discussed in the previous chapter. You should recall from that analysis that a shadow price implicitly emerges from that exercise, this price being equal to the monetary value of marginal damage at the efficient level of pollution. This is the rate at which the tax (or subsidy) should be applied per unit of emissions.

Figure 7.3 illustrates the working of an emissions tax. Note that the diagram uses aggregate, economy-wide marginal benefit and marginal damage functions (not those of individuals or single firms). If firms behave without regard to the pollution they generate, and in the absence of an emissions tax,

[10] Liability can also be viewed as an incentive-based instrument, although we do not pursue that interpretation any further here.

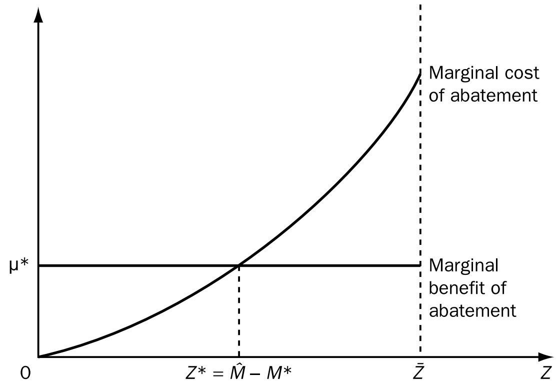


*Figure 7.4* The economically efficient level of emissions abatement

emissions will be produced to the point where the private marginal benefit of emissions is zero. This is shown as $\hat{M}$, the pre-tax level of emissions.

Now suppose an emissions tax was introduced at the constant rate $\mu^*$ per unit emission, the value of marginal damage at the efficient pollution level. Given this, the post-tax marginal benefit schedule differs from its pre-tax counterpart by that value of marginal damage. Once the tax is operative, profit-maximising behaviour by firms leads to a pollution choice of $M^*$ (where the post-tax marginal benefits of additional pollution are zero) rather than $\hat{M}$ as was the case before the tax. Note that levying an emissions tax at the rate $\mu^*$ creates just the right amount of incentive to bring about the targeted efficient emission level, $M^*$.[11]

It is sometimes more convenient to view the problem in terms of abatement, $Z$, rather than the level of pollution itself. This can be done by reinterpreting Figure 7.3. Viewed in this new light, the emission tax causes abatement to increase from zero (at $\hat{M}$) to its efficient level $Z^* = \hat{M} - M^*$ at the point $M^*$ on the horizontal axis. Alternatively, we can map the relevant parts of Figure 7.3 into abatement space, from which we obtain Figure 7.4. It is important to be clear about the relationships between these two diagrams. First, the curve labelled 'marginal cost of abatement' is just the mirror image of the (before-tax) marginal benefit curve in Figure 7.3; what firms privately forgo when they abate emissions is, of course, identical to the benefits they receive from emissions. The 'marginal benefit of abatement' to a representative firm is the tax rate applied, $\mu^*$. Each unit of abated emissions reduces the firm's total tax bill by that amount. As the tax rate is constant, the marginal benefit of abatement curve is horizontal. Secondly, note that we have truncated the two curves in Figure 7.4 at $Z = \bar{Z}$, where $\bar{Z}$ is identical in magnitude to $\hat{M}$. Confirm for yourself the reason for doing this. Finally, note that $Z^* = \hat{M} - M^*$, and so the distance from the origin to $Z^*$ in Figure 7.4 is equal to the horizontal distance between $\hat{M}$ and $M^*$ along the emissions axis in Figure 7.3.

In the absence of an emissions tax (or an abatement subsidy), firms have no economic incentive to abate pollution. (In terms of Figure 7.4, the marginal benefit of abatement lies at zero along the $Z$ axis.) Profit-maximising behaviour implies that firms would then undertake zero abatement, corresponding to emissions $\hat{M}$. However, when an emissions tax is levied (or, equivalently, when an abatement subsidy is available) an incentive to abate exists in the form of tax avoided (or subsidy gained). It will be profitable for firms to reduce pollution as long as their marginal abatement costs are less than the value of the tax rate per unit of pollution (or less than the subsidy per unit of emission abated). If the tax/subsidy is levied at the level $\mu^*$ the efficient pollution level is attained without coercion, purely as a result of the altered structure of incentives facing firms.

In the language of externalities theory, the tax eliminates the wedge (created by pollution damage) between private and socially efficient prices; the tax brings private prices of emissions (zero) into line with social prices ($\mu^*$). The tax 'internalises the externality' by inducing the pollution generator to behave as if pollution costs entered its private cost functions. Decisions will then reflect all relevant costs, rather than just the producer's private costs, and so the profit-maximising pollution level will coincide with the socially efficient level. Not only will the tax instrument (at rate $\mu^*$) bring about an efficient level of pollution but it will also achieve

[11] As shown in Appendix 7.1, a subsidy at the rate $\mu^*$ on units of pollution abated would have an equal short-run effect on emissions to a pollution tax at the rate $\mu^*$ on unabated units of pollution.

that target in a cost-effective way. Remember that cost-efficiency requires that the marginal abatement cost be equal over all abaters. Under the tax regime all firms adjust their firm-specific abatement levels to equate their marginal abatement cost with the tax rate. But as the tax rate is identical for all firms, so are their marginal costs.

Our discussion in this section so far has dealt with the case in which the EPA wishes to attain the economically efficient level of emissions, $M^*$. However, we saw in the previous chapter that the EPA may not have sufficient information for this to be feasible. Suppose that the EPA does have an emissions target, $\tilde{M}$, set perhaps on health grounds. To attain this (or indeed any other specific) emissions target, knowledge of the aggregate (economy-wide) marginal emissions abatement cost function would be sufficient. This should be clear by looking at Figure 7.4 again. For any target $\tilde{M}$, knowledge of the aggregate marginal cost of abatement function allows the EPA to identify the tax rate, say $\tilde{\mu}$, that would create the right incentive to bring about that outcome. Even though the target is not an efficient target, the argument used above about cost-efficiency remains true here: the emissions tax, levied at $\tilde{\mu}$, attains the target $\tilde{M}$ at least total cost, and so is cost-efficient. This result is rather powerful. Not only does the EPA not need to know the aggregate marginal pollution damage function, it does not need to know the abatement cost function of each firm. Knowledge of the aggregate abatement cost function alone is sufficient for achieving any arbitrary target at least cost. Compare this result with the case of command and control instruments; there, knowledge of every firm's marginal abatement cost function is required – a much more demanding information requirement.

Finally, let us deal with the third of the listed cases where an emission reduction of some unspecified amount is sought. Without knowledge of anything about abatement costs and benefits, the EPA could select some arbitrary level of emissions tax, say $\breve{\mu}$. Faced with this tax rate, profit-maximising firms will reduce emissions up to the point where marginal abatement costs are brought into equality with this tax rate. As all firms do this the emissions reduction is achieved at least cost, once again. Although the government cannot know in advance how much pollution reduction will take place, it can be confident that whatever level of abatement is generated would be attained at minimum feasible cost. Taxes (and subsidies by an equivalent argument) are, therefore, cost-efficient policy instruments. These results are demonstrated formally in Appendix 7.1, Parts 4 and 5.

We stated earlier that an emissions tax and an emissions abatement subsidy (at the same rate) have an identical effect in terms of pollution outcome in the short term (see Part 6 of Appendix 7.1). However, the two instruments do have some very important differences. Most importantly, the distribution of gains and losses will differ. Taxes involve net transfers of income from polluters to government, while subsidies lead to net transfers in the other direction (see Problem 4). This has important implications for the political acceptability and the political feasibility of the instruments. It also could affect the long-run level of pollution abatement under some circumstances. Some more discussion on this matter is given in Box 7.4.

To reinforce your understanding of this material in this section, you are recommended to work through Problem 10 at the end of this chapter. This uses an Excel workbook to simulate emissions reduction using command and control techniques, tax and subsidy instruments, and (to be discussed in the next section) transferable permits. Some information on practical experience with pollution taxes and abatement subsidies is given in Box 7.5.

### 7.4.2 Marketable emissions permits

As with command and control and tax/subsidy instruments, marketable permits (also known as tradable or transferable permits) can be applied at many points in the production-to-pollution process represented in Figure 7.2. Here we consider only one form: permits on the quantity of *emissions*. Marketable permit systems are based on the principle than any increase in emissions must be offset by an equivalent decrease elsewhere. There is a limit set on the total quantity of emissions allowed, but the regulator does not attempt to determine how that total allowed quantity is allocated among individual sources.

**Box 7.4 Are pollution taxes and emissions abatement subsidies equivalent?**

For an industry of a given size, an emission tax and an abatement subsidy levied or paid at the same rate are equivalent in terms of units of emissions abated. Thus, looking at Figure 7.3 again, a subsidy or a tax at the rate $\mu^*$ would reduce emissions from $\hat{M}$ to $M^*$ for a single firm with a given capital structure. As the industry is simply the sum of all firms, if the number of firms remains constant and the capital structure of each firm is unchanged, then the effects of taxes and subsidies are identical.

However, the two instruments are different in their effects on income distribution. A firm gains additional income from an abatement subsidy, as it will undertake abatement only when the unit abatement subsidy exceeds its marginal abatement cost. A tax on the other hand results in a loss of income to the firm as it pays the tax on all its emissions. To make this comparison more precise, look at Figure 7.5, the functions in which reproduce those in Figure 7.3.

An abatement subsidy will result in a payment to the firm equal to the areas $S_1 + S_2$, that is, $\mu^*$ multiplied by $(\hat{M} - M^*)$. However, by reducing emissions from $\hat{M}$ to $M^*$ the firm loses $S_2$ in profit on final output. The net gain to the firm is equal, therefore, to the area $S_1$. A tax levied at the rate $\mu^*$ on emissions $M^*$ will cost the firm $\mu^*M^*$, that is, the sum of the areas $S_3$, $S_4$, $S_5$ and $S_6$.

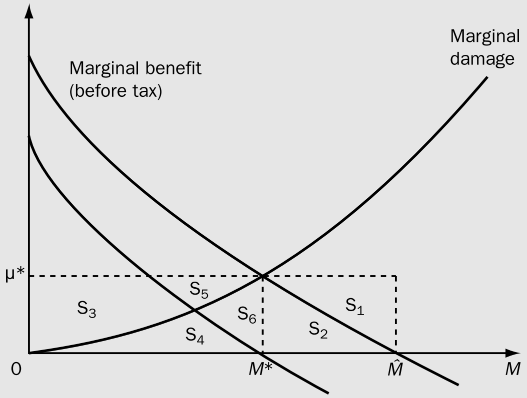


*Figure 7.5* Emissions tax and abatement subsidy schemes: a comparison

However, by reducing emissions from $\hat{M}$ to $M^*$ the firm also loses profit on reduced output, the area $S_2$. So the income effects are entirely different.

Let us explore this difference a little further. Recall that the tax paid is equal in value to $\mu^*M^*$, while the subsidy received is $\mu^*(\hat{M} - M^*)$. But $\mu^*(\hat{M} - M^*) = \mu^*\hat{M} - \mu^*M^*$. The second term on the right-hand side is the tax paid, and will depend on the amount of abatement undertaken. It is this second component which gives the firm an incentive to abate emissions. Recalling that $\mu$ is an outflow in a tax scheme and an inflow in a subsidy scheme, an outflow of $\mu^*M^*$ (with a tax) is identical to an inflow of $-\mu^*M^*$ (with a subsidy). The two incentive effects are identical, and it is this that forms the basis for the claim that the instruments are equivalent. However, the subsidy differs from the tax by the presence of the additional term, $\mu^*\hat{M}$, a fixed or lump-sum payment, independent of the amount of abatement the firm actually undertakes. In the long run such payments may alter industry profitability, and so alter the size of the industry itself. This lump-sum payment component of the subsidy may destroy the equivalence between the two instruments in terms of their effects on emissions abatement.

We are faced with the possibility that a subsidy might enlarge the industry, partially or wholly offsetting the short-run emissions reduction. It is not possible to be more precise about the final outcome, as that depends on several other factors, including whether or not government introduces other fiscal changes to counteract the income effects we have just described. A general equilibrium analysis would be necessary to obtain clear results. This is beyond our scope in this text, so we just note that the equivalence asserted above is not valid in all cases.

Finally, note another aspect of an abatement subsidy scheme. As one component of the subsidy payment depends on the uncontrolled level of emissions (that is, the component $\mu^*\hat{M}$), a firm has an incentive to misrepresent the uncontrolled level of emissions in order to obtain a favourable benchmark in terms of which the subsidy payments are calculated.

**Box 7.5 The use of economic instruments in OECD countries**

The use of economic instruments to achieve environmental goals has increased markedly since the 1970s. The number of applications has increased, as has the variety of instruments used. Revenues from environmentally related taxes in 2000 constituted about 7% of total OECD tax revenue, a figure that is growing steadily and which had accelerated at the end of the 1990s.

User charges and subsidies were being applied in the 1970s. Since then, emissions charges and taxes have become widespread, subsidies to encourage the installation or use of environmentally friendly capital equipment have become common, and several other incentive-based instruments have appeared for the first time, including deposit-refund systems and performance bonds. The use of marketable permits has began to appear, although it is as yet not widely spread. Table 7.5 lists the main categories of economic instruments and their usage in OECD economies. Box 7.6 considers several examples of the use of emissions taxes and emissions abatement subsidies.

Economic instruments are also widely used for natural resource management. Common applications are in the management of water quantity (typically abstraction charges or taxes), fisheries (taxes, fees and transferable quotas), forestry (charges and subsidies) and wetlands (financial assistance to owners). Economic instruments are also used to preserve soil and land quality, and to preserve species and wildlife (typically fees and permits). Several examples of resource management or conservation instruments are given in the resource harvesting chapters (17 and 18) later in the text.

*Table 7.5* Economic instruments used in OECD countries

| Country | Fees, charges and taxes | Tradable permits | Deposit-refund systems | Non-compliance fees | Performance bonds | Liability payments | Subsidies |
|---|---|---|---|---|---|---|---|
| Australia | • | • | • | | • | | • |
| Austria | • | | • | | | | • |
| Belgium | • | | | | | | |
| Canada | • | • | • | • | • | • | • |
| Czech Republic | • | | • | • | | | • |
| Denmark | • | • | • | | | • | • |
| Finland | • | | • | | | • | • |
| France | • | • | | | | | • |
| Germany | • | | | | | • | |
| Greece | • | | | • | | | • |
| Hungary | • | | • | • | | | |
| Iceland | • | | • | | | | |
| Ireland | • | | | | | | |
| Italy | • | | • | | | | |
| Japan | • | | | | | • | • |
| Korea | • | | • | • | | | |
| Mexico | • | | • | | | | |
| Netherlands | • | | • | | | | • |
| New Zealand | • | | | | | | |
| Norway | • | | • | • | | | • |
| Poland | • | • | • | • | | | • |
| Portugal | • | | | | | | |
| Spain | • | | | | | | |
| Sweden | • | | • | • | | • | • |
| Switzerland | • | • | | | | | • |
| Turkey | • | | • | • | | • | • |
| UK | • | | | | | | |
| USA | • | • | • | | • | • | • |

*Source*: OECD (1999)

Notes to Table 7.5:

1. Entries marked by • denote that the instrument category was stated to be used (or to have been used) by the country in question in response to a questionnaire-based survey of all (29) OECD economies in 1999. 24 countries responded. Non-respondent countries are those for which there are no entries in this table in any column except that labelled 'Fees, charges and taxes'
2. 'Charges' refer to requited emissions charges, user charges and product charges

**Box 7.6 Emissions tax and abatement subsidies in practice**[12]

The majority of emissions taxes in current use apply to the transport and energy sectors. A third important application is to waste management. Emissions fees were used in at least 20 OECD countries in 1999, and their use has been growing steadily since 1985. The OECD now lists approximately 200 examples of fees or taxes in the areas of air, water and noise pollution, and waste disposal. In some cases, tax revenues are earmarked for purposes of environmental improvement. In Germany and Italy, charges are used in conjunction with effluent standards: those firms which meet or better the standards are taxed at a lower rate per unit effluent than others.

Air-pollutant emissions charges are being used predominantly in Japan and a number of European countries. France has used charges as an incentive to install pollution abatement technology, with charges being repaid in the form of capital subsidies to firms adopting recommended control technologies. In 1998, France integrated several existing charges into a unified 'General Tax on Polluting Activities' (TGAP); the environmental agency is allocated a share of TGAP revenues for environmental improvement programmes. Sweden charges combustion plants for $NO_X$ emissions, with revenue being distributed among emitters in proportion to their share in total energy output. Hence the total cost of the system to emitters is zero, but each plant has an incentive to reduce its emissions-to-energy-output ratio. The regime appears to have led to significant falls in $NO_X$ emissions and to have spurred innovation in combustion technology. In Japan emissions levies are earmarked as a compensation fund for victims of air pollution; charge levels are dependent upon amounts of compensation paid out in previous years.

Several countries – including Australia, the Czech Republic, Hungary and Iceland – have systems of charges for ozone-depleting substances. Differential tax rates on leaded and unleaded petrol in the United Kingdom serve as an indirect charge on lead emissions, and Sweden has used differential charges and subsidies on cars and heavy vehicles to encourage the purchase of low-pollution engines and the adoption of catalytic converters. There are relatively high rates of tax on electricity and primary energy sources throughout Western Europe; while not being pollution taxes as such, they do have similar incentive effects by encouraging energy conservation and enhancing energy efficiency.

Although the European Union has abandoned plans for a common carbon tax, Denmark, Finland, Italy, the Netherlands, Norway, Sweden and the UK all currently use some form of energy tax which, to varying degrees, reflects the carbon content of fuels. However, in the great majority of countries where $CO_2$ (or other environmental) taxes have been implemented, some sectors have been exempted from the tax, or the tax rate is differentiated across sectors. This reduces the cost-effectiveness, and so raises the real cost, of the tax.

Water effluent charges are used in Australia, Belgium, Canada, the Czech Republic, France, Italy, Germany, Mexico, the Netherlands, Poland and several US states. Charge rates vary according to the quantity and quality of waste water. The UK has a landfill tax; this is examined in Box 7.7.

The USA makes little use of emissions taxes or charges. Exceptions include a tax on chlorofluorocarbons to help in the phasing out of these chemicals, and fees on sewage and solid and hazardous waste at landfills. Households typically pay by the gallon for sewage disposal, and waste haulage firms pay by the ton for solid waste disposal. However, household and business enterprises have traditionally paid *lump-sum* charges for solid waste disposal, and so *marginal* disposal costs are not passed on to the initial producers of waste, leading to significant efficiency losses. As more states move to volume-related charges (37 states now do this), volumes discarded have fallen and recycling rates have risen significantly (Anderson *et al.*, 1997). The United States has, though, made more extensive use of marketable emission permit instruments than have European economies (see Box 7.8).

Tax rates are typically set at levels insufficient to fully internalise external costs (EEA, 2000). Low rates of tax or subsidy imply correspondingly low levels of impact. In some cases charges have been high enough to have large incentive effects. The Netherlands, with relatively high rates, has shown large

[12] In this box we do not distinguish between taxes and fees or charges, using the terms interchangeably.

**Box 7.6** ***continued***

improvements in water quality. Sweden's use of differential taxes and subsidies, and the differential tax on unleaded petrol in the UK have been very effective in causing substitution in the intended directions. In some instances, the revenues from specific charges are earmarked for particular forms of environmental defence or clean-up expenditure – one example is the use of taxes on new paint purchases in British Columbia to support reprocessing and safe disposal of used paint.

Subsidies for attainment of environmental improvements are used widely. A few countries use subsidies that are proportionately related to quantities of air emissions or water effluent. It is far more common, though, for subsidies to be paid in the form of grants, tax allowances or preferential loans for capital projects that are expected to lead to environmental improvements (such as low-emissions vehicles, cleaner waste-treatment plants or the development of environmentally friendly products). These schemes are often financed from earmarked environmental funds. A comprehensive listing of such schemes can be found on the web page of OECD (1999).

*Sources*: Tietenberg (1990), Goodstein (1995), OECD (1999)

**Box 7.7 Landfill tax example**

A landfill tax was introduced in the UK in 1996. The tax, paid by landfill operators, is set at different rates for inactive waste such as bricks (£2 per tonne) and other waste (£7 per tonne). An element of tax neutrality is imposed by reducing employers' national insurance contributions to offset the costs of the landfill tax.

The tax is designed so that incentives exist to reduce waste flows. However, since its inception, operation of the tax has been plagued by concerns that waste has been disposed of illegally to avoid landfill tax charges. This illustrates the point that incentive-based instruments for environmental control may be ineffective unless there is careful monitoring and methods for ensuring compliance.

Charges levied on landfill operators are also found in the Czech Republic (since 1992). The tax is in two parts, the first being imposed on all landfill operators (with revenues recycled to municipal authorities for environmental protection activities). The second component – strictly speaking, a non-compliance fee – charges operators who fail to attain specified standards. Evidence suggests that the tax has markedly increased the proportion of sites attaining specified standards. A similar system operates in the Slovak Republic. It is more common for charges to be placed on generators of waste (rather than disposers of it), with applications in China, Estonia, Hungary, Poland and Russia.

There are two broad types of marketable emission permit systems – the 'cap-and-trade' system and the emission reduction credit (ERC) system. We shall analyse the cap-and-trade approach in some depth, and briefly consider the ERC system in Section 7.4.2.4 below.

A cap-and-trade marketable emission permits scheme *for a uniformly mixing pollutant* involves:[13]

- A decision as to the total quantity of emissions that is to be allowed (the 'cap'). The total amount of permits issued (measured in units of pollution) should be equal to that target level of emissions.
- A rule which states that no firm is allowed to emit pollution (of the designated type) beyond the quantity of emission permits it possesses.
- A system whereby actual emissions are monitored, and penalties – of sufficient deterrent power – are applied to sources which emit in excess of the quantity of permits they hold.
- A choice by the control authority over how the total quantity of emission permits is to be initially allocated between potential polluters.
- A guarantee that emission permits can be freely traded between firms at whichever price is agreed for that trade.

[13] We deal with marketable permits for non-uniformly-mixing pollutants in Section 7.5.3.

Marketable permit schemes differ from tax or subsidy schemes by working in terms of quantities rather than prices. But this feature is also true for command and control instruments such as quotas, licences and standards. The distinguishing feature is the transferability of permits between individual sources in the marketable permits case. Permit trading is not allowed in command and control licence systems.

It is the exchange process that generates the attractive qualities of the marketable permit system. In effect, tradability creates a market in the right to pollute. In that market, the right to pollute will have a value, given by the prevailing market price. So the decision to pollute generates an opportunity cost. By emitting an extra unit of the pollutant, one unit of permit is used up and so cannot be sold to another firm. The firm incurs a cost in emitting each unit of the pollutant, that cost being the current market permit price. Intuitively, this suggests that a marketable permit system should be equivalent (at least in some ways) to a tax or subsidy system, provided the permit price is equal to the tax or subsidy rate. As we shall see, this intuition is correct.

Let us consider how an equilibrium price might emerge in the market for permits. Suppose that permits have been allocated at no charge to firms in some arbitrary way. Once this initial allocation has taken place, firms – both those holding permits in sufficient number to cover their desired emission levels and those not holding sufficient for that purpose – will evaluate the marginal worth of permits to themselves. These valuations will differ over firms.

Some firms hold more permits than the quantity of their desired emissions (in the absence of any control). The value of a marginal permit to these firms is zero.[14] Others hold permits in quantities insufficient for the emissions that they would have chosen in the absence of the permit system. The marginal valuations of permits to these firms will depend upon their emission abatement costs. Some will have high marginal abatement costs, and so are willing to pay high prices to purchase emissions permits. Others can abate cheaply, so that they are willing to pay only small sums to purchase permits; their marginal permit valuation is low.

Indeed, it is not necessarily the case that a firm which holds fewer permits than its desired emissions level will *buy* permits. It always has the option available to reduce its emissions to its permitted level by undertaking extra abatement. The firm may find it preferable to sell permits (rather than buy them) if the price at which they could be sold exceeds its marginal abatement cost.

In any situation where many units of a homogeneous product are held by individuals with substantially differing marginal valuations, a market for that product will emerge. In this case, the product is tradable permits, and the valuations differ because of marginal abatement cost differences between firms. Therefore, a market will become established for permits, and a single, equilibrium market price will emerge, say $\mu$. Notice that trading does not alter the quantity of permits in existence, it merely redistributes that fixed amount between firms.

In equilibrium marginal abatement costs will be equal over all firms. It is this property of the system which ensures that transferable marketable permits, like taxes and subsidies, achieve any given target at least cost. Moreover, another equivalence arises. If the total quantity of permits issued is $M^*$ *and* that quantity is identical to the level of emissions which would emerge from an emissions tax (or an abatement subsidy) at the rate $\mu^*$ then a marketable permit scheme will generate an equilibrium permit price $\mu^*$. In effect, the marketable permit system is an equivalent instrument to either emissions taxes or emissions abatement subsidies. We demonstrate this result algebraically in Part 7 of Appendix 7.1.

#### 7.4.2.1 The initial allocation of permits

The implementation of a marketable permits system requires that the EPA select a method by which the total allowable quantity of permits (the cap) is initially allocated among sources. Simplifying matters somewhat, we can envisage that it must choose one of the following:

[14] If permits were storable or 'bankable' so that they could be used in the future, their worth would be positive (rather than zero) as there will be some positive probability that they could be used later when the firm would otherwise have insufficient permits to cover desired emissions. But we shall leave this complication to one side for now.

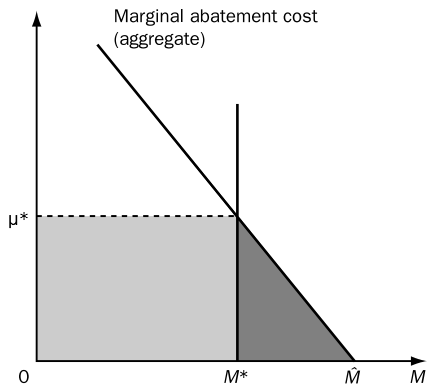


*Figure 7.6* The determination of the market price of emissions permits

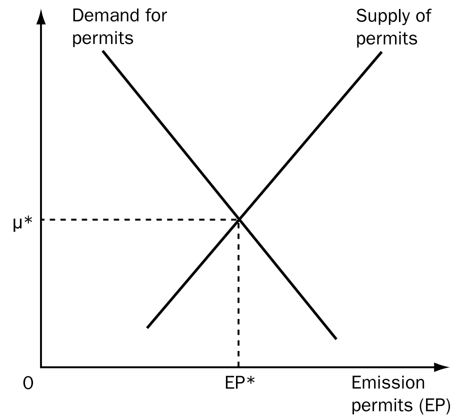


*Figure 7.7* The determination of the market price of emissions permits: free initial allocation case

- the EPA sells all permits by auction;
- the EPA allocates all permits at no charge (which in turn requires that a distribution rule be chosen).

We shall now investigate how the market price of permits is determined in each of these two cases.

#### 7.4.2.2 Determination of the equilibrium market price of permits

*Case 1: Auctioned permits*

Suppose that the permits are initially allocated through a competitive auction market. Individual firms submit bids to the EPA. When ranked in descending order of bid price, the resulting schedule can be interpreted as a market demand curve for permits. Assuming that no strategic behaviour takes place in the bidding process, this demand curve will be identical to the aggregate marginal abatement cost function.

The market equilibrium permit price is determined by the value of the aggregate marginal abatement cost at the level of abatement implied by the total number of issued permits.[15] This is illustrated in Figure 7.6. The demand curve for permits is the aggregate marginal abatement cost function for all polluting firms. The total number of permits (allowed emissions) is $M^*$. Given this quantity of permits, the market price for permits will be $\mu^*$. Firms collectively are required to reduce emissions from $\hat{M}$ to $M^*$.

*Case 2: Free initial allocation of permits on an arbitrary basis*

Alternatively, the EPA may distribute the permits at no charge, and allow them to be subsequently traded in a free market. The initial allocation is unlikely to correspond to the desired (that is, profit-maximising) holdings of permits (and in aggregate, of course, is likely to be less than total desired emissions). Some firms will try to buy additional permits from others, while others will try to sell some of their initial holding. Buyers will typically be firms with relatively high marginal abatement costs, who hope to purchase additional quantities at a price less than their marginal abatement cost. Sellers will be those in an opposite position, hoping to sell some permits at a price greater than their marginal abatement cost.

In a well-functioning competitive market, the market price that would emerge in this case would be identical to that which would be established if permits were sold at a competitive auction. This is portrayed in Figure 7.7. Note that the quantity traded, EP*, is less than the number of permits issued by the EPA ($M^*$), because trades only take place as holdings are adjusted to desired levels.

[15] It is assumed here that all permits are sold at one price (the highest single price consistent with selling all permits).

It is clear that the method by which permits are initially allocated has no bearing on the amount of abatement that takes place; that depends only on the total number of permits issued. What is, perhaps, less evident is that the method of initial allocation also has no effect on the equilibrium permit price.

There is one important qualification to these remarks about permit price determination. We have assumed that the market behaves as if it were perfectly competitive. But if the polluting industry in question is dominated by a small number of firms, or if for any reason the quantity of trading is small, strategic behaviour may take place. This could happen both in permit auctions and where firms are adjusting permit holdings from their initial allocations to their profit-maximising levels. Strategic behaviour may cause the market price of permits to diverge from its competitive level.

A simple numerical illustration (which extends an example used earlier in the chapter) will help to strengthen understanding about the way that this instrument operates. Consider the information shown in Table 7.6. We suppose that the EPA selects an emissions cap – and so a total permit allocation – of 50 units. The pollutant is emitted by just two firms, A and B, and emissions abatement can only be undertaken by these firms. The EPA decides arbitrarily to allocate half of total permits to each firm, so prior to trading A and B are each allowed to emit 25 units of the pollutant. As in our earlier discussion, we assume that in the absence of any control system A would choose to emit 40 units and B 50 units.

*Table 7.6* Emissions abatement data for firms A and B

| | A | B | A + B |
|---|---|---|---|
| Uncontrolled emissions | 40 | 50 | 90 |
| Uncontrolled abatement | 0 | 0 | 0 |
| Efficient emissions | 15 | 35 | 50 |
| Efficient abatement | 25 | 15 | 40 |
| Initial permit allocation | 25 | 25 | 50 |
| Final permit allocation | 15 | 35 | 50 |

Given the initial permit allocations, A must reduce emissions by 15 units and B by 25 units. It can be seen from Figure 7.8 (which reproduces exactly the abatement cost functions used previously in Figure 7.1) that A has a marginal abatement cost of 45 and B a marginal abatement cost of 125.

The fact that firm A has lower marginal abatement cost than firm B after the initial permit allocation implies that the total abatement of 40 units of emission is not being achieved at least cost. Moreover, B places a much higher value on an incremental permit than does A (125 as compared with 40). Thus the two will find it mutually beneficial to trade with one another in permits. What will be the outcome of this trade? If the market behaved as if it were a competitive market, an equilibrium market price of 75 would emerge. At that price, firm B (the high-cost abater) would buy permits and A (the low-cost abater) would sell permits. In fact, A would buy 10 permits from A at 75 each, because for each of those

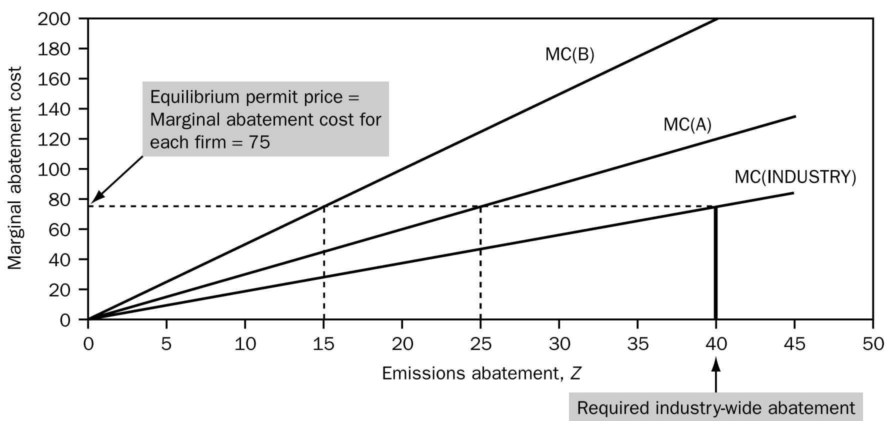


*Figure 7.8* Efficient abatement with two firms and marketable permits

10 permits, it would be paying less than it would cost the firm to abate the emissions instead. Conversely, B would sell 10 permits to A at 75 each, because for each of those 10 permits, it would be receiving more than it would cost the firm to abate the emissions instead.

Trading finishes at the point where A has 15 permits (10 less than its initial allocation) and B has 35 (10 more than its initial allocation). Marginal control costs are equalised across polluters, and the total cost of abating emissions by 40 units has thereby been minimised. The permit system will, therefore, have identical effects on output and emissions as an optimal tax or subsidy system, and will be identical in terms of its cost-effectiveness property.

One other feature shown in Figure 7.8 should be noted. The line labelled MC(Industry) is the industry-wide (or aggregate) marginal cost of abatement schedule. It is obtained by summing horizontally the two firm's marginal abatement cost functions, and is given by[16]

$$\text{MC(Industry)} = \frac{15}{8}Z$$

The equilibrium permit price is found as the industry marginal cost (75) at the required level of industry abatement (40). Note that as the required abatement rises, so will the equilibrium permit price.

#### 7.4.2.3 Marketable permit systems and the distribution of income and wealth

In a perfectly functioning marketable permit system the method of initial allocation of permits has no effect on the short-run distribution of *emissions* between firms. But it does have significant effects on the distribution of *income and wealth* between firms. If the permits are sold by competitive auction, each permit purchased will involve a payment by the acquiring firm to the EPA equal to the equilibrium permit price. A sum of money equal to $\mu^*$ multiplied by $M^*$ will thus be transferred from businesses to government. This is shown by the lighter shaded area in Figure 7.6.

In addition to this, the emissions restrictions will impose a real resource cost (rather than a financial transfer) on firms. In terms of Figure 7.6 again, firms collectively are required to reduce emissions from $\hat{M}$ to $M^*$ and so the real resource costs of the abatement are given by the area of the shaded triangle to the right of $M^*$; that is, the sum of marginal abatement costs over the interval $\hat{M}$ to $M^*$. If firms must initially buy the permits from the government at the price $\mu^*$ then they will collectively face a further financial burden shown by the lighter shaded area in the diagram.

Note that the transfer of income from the business sector to the government when successful bids are paid for is not a real resource cost to the economy. No resources are being used, there is simply a transfer of income between sectors of the economy. Whenever we discuss least-cost methods of abatement in this chapter, you should note that it is the real resource costs that are being minimised, not any transfer costs such as those just referred to.

If, on the other hand, the EPA distributes permits at no charge, there is no transfer of income from businesses to government. However, there will be transfers between firms. Some will buy from others and some will sell to others. So some firms will gain financially while others lose. The pattern and magnitude of these within-industry transfers will depend on the formula used to make the initial permit allocation.

But even here there is still a real resource cost to the business sector, equal once again to the triangular shaded area in Figure 7.6. That burden is the same whatever initial allocation system is used. Taking all these remarks together, it is clear that the free allocation system is more attractive to polluting firms than the auction sale of permits.

The fact that there are different net income effects means that we must introduce the same qualification we made earlier (in comparing taxes with subsidies) about long-run effects. An industry may contract in the long run if permits must be initially purchased; this effect will not be present when they are distributed at no charge.

[16] To obtain this, first invert the two firm's functions, giving $Z_A$ = (1/3)MC and $Z_B$ = (1/5)MC. Next sum the two inverted equations to give $Z$ = ((1/3) + (1/5))MC = (8/15)MC. Finally, invert this summed expression to obtain MC = (15/8)$Z$.

#### 7.4.2.4 The emission reduction credit (ERC) form of marketable permit system

Previous paragraphs have referred to a cap-and-trade permit system. A few comments are in order about the alternative ERC system. In an ERC approach, a 'business-as-usual' scenario is taken to estimate a baseline profile of relevant emissions. Emissions by any particular source above its anticipated baseline volume are subject to some non-compliance penalty. However, if a source emits less than its calculated baseline level, it earns a corresponding amount of emission reduction credits. Such credits can be sold to other sources that anticipate exceeding their baseline emission level.[17] The purchased ERCs constitute an entitlement to exceed baseline emissions without penalty.

The US emission permits scheme is a modified form of this ERC system. There, marketable permits operate in conjunction with more conventional standards or licence schemes. The United States Environmental Protection Agency (US EPA) establishes national ambient air quality or permissible water pollutant concentration standards. To attain these standards, controls – required abatement technologies or ceilings on emissions flows – are imposed on individual polluting sources. This is the conventional command and control approach that has characterised pollution control in most countries in the twentieth century. The novelty arises in the next component of the programme.

If any polluter succeeds in reducing emissions by a greater amount than is required by the standard it must satisfy, it obtains emission reduction credits of that quantity. The firm which acquires these emission reduction credits can engage in trades, selling some or all of its ERC to other firms, which then obtain a legal entitlement to emit pollutants beyond the standard which the USEPA has imposed on them. Put another way, each firm is legally entitled to emit a quantity of pollutants up to the sum of its standard entitlement plus any ERC it has acquired. Each ERC is, thus, in effect, a transferable or marketable emissions permit.

The American ERC trading system has a number of other distinctive features:

The *offset policy* allows existing firms to expand, or new firms to enter, areas in which emission standards have not been met in the aggregate provided that they acquire sufficient quantities of ERC. In other words, growth can take place provided corresponding emissions reductions take place by existing firms within that area.

The *bubble policy* treats an aggregate of firms as one polluting source (as if they were enclosed in a single bubble) and requires that the bubble as a whole meets a specified standard. If that is achieved, it does not matter whether an individual source within the bubble fails to meet the firm-specific standard imposed on it.

*Emissions banking* allows firms to store ERC for subsequent use or sale to others.

Some additional information on the complexities of marketable permit schemes that have been used in practice is given in Box 7.8. The examples considered there also include permit schemes in which what is being 'permitted' is something other than pollution emissions.

## 7.5 Pollution control where damages depend on location of the emissions

We now consider instruments designed to attain pollution stock (rather than emission) targets for non-uniformly-mixing stock pollutants (non-UMP). Previous analysis has shown that in this case the spatial location of emissions is of central importance. It will be convenient to deal with the particular example of air pollution in an 'airshed' that contains several spatially distinct receptor areas and many emission sources. However, our results will apply, with suitable change of terminology, to any non-UMP.

We saw earlier that one way in which the EPA may handle these issues is by controlling *ex ante* the location of polluters and people affected by pollution. Indeed, in the very long run, the best way of dealing with this problem is likely to be zoning: prohibiting

[17] If banking is allowed, they may also be used by the source at a later date.

**Box 7.8 Marketable permits in practice**

The United States seems to have been the first country to adopt the use of marketable permits to attain environmental goals. In the case of *emissions* control, these have covered $SO_2$ and ozone-depleting substances (ODS), mobile-source pollutants (HC and $NO_X$), lead in petrol, and water quality (BOD). Marketable permit systems can now also be found in Australia (saline discharges into rivers), Canada (ODS, and pilot schemes for $NO_X$ and VOC in Ontario), Denmark ($CO_2$ power plant emissions), Poland (VOC), Switzerland ($NO_X$ and VOC) and several individual US states ($NO_X$ and $SO_2$ and the use of woodstoves and fireplaces in mountainous areas).

There are also examples of marketable permit schemes for purposes other than emissions control. Often, these consist of marketable extraction, harvesting or development rights for a variety of natural resources. Examples include the Australian system of water abstraction rights, construction or development permits for land management in the USA, France (housing in the Alps) and New Zealand (housing density), and a large variety of permit systems for the harvesting of renewable resources (e.g. transferable fishing or logging quotas; several of these are described in Chapters 17 and 18).

The actual extent to which marketable emissions permit programmes have been used is limited, but has undergone considerable growth in recent years. It has been used to reduce the lead content in petrol, to control production and use of chlorofluorocarbon ozone-depleting substances, and in the 'Emissions Trading Program' for the control of volatile organic compounds, carbon monoxide, sulphur dioxide, particulates and nitrogen oxide. Details of these programmes can be found in surveys by Cropper and Oates (1992), Tietenberg (1990), Hahn (1989, 1995), Hahn and Hester (1989a, b), Opschoor and Vos (1989) and Goodstein (1995). The passage of the 1990 Amendments to the Clean Air Act has seen the United States introduce a major system of marketable permits to control sulphur emissions.

Most economists expect emissions trading to confer large efficiency gains relative to the use of command and control instruments alone. These gains arise from the reductions in overall abatement costs that trading permits. Recall from our previous discussions that high-cost abaters do less abatement and low-cost abaters do more abatement when trading of permits or ERC is allowed. Tietenberg's assessment of the performance of the emissions permit trading schemes is

- The programme has unquestionably and substantially reduced the costs of complying with the Clean Air Act. Most estimates place the accumulated capital savings for all components of the programme at over $10 billion. This does not include the recurrent savings in operating costs. On the other hand the programme has not produced the magnitude of cost savings that was anticipated by its strongest proponents at its inception.
- The level of compliance with the basic provisions of the Clean Air Act has increased. The emissions trading programme increased the possible means for compliance and sources have responded accordingly.
- The vast majority of emissions trading transactions have involved large pollution sources.
- Though air quality has certainly improved for most of the covered pollutants, it is virtually impossible to say how much of the improvement can be attributed to the emissions trading programme.

Tietenberg, in Markandya and Richardson (1992), pp. 269–270

A survey by Cropper and Oates confirms the view that the use of transferable permit programmes, and other market incentive schemes based on taxes or subsidies, has been limited in scale, but they assess that interest in and acceptability of market-based incentive instruments is growing:

> effluent charges and marketable permit programs are few in number and often bear only a modest resemblance to the pure programs of economic incentives supported by economists. . . . As we move into the 1990's, the general political and policy setting is one that is genuinely receptive to market approaches to solving our social problems. Not only in the United States but in other countries as well, the prevailing atmosphere is a conservative one with a strong predisposition towards the use of market incentives wherever possible, for the attainment of our social objectives.
>
> Cropper and Oates (1992), pp. 729, 730

An important new development was initiated at Kyoto, Japan in 1997. The industrialised countries, in agreeing to a programme of greenhouse gas emissions limits, decided that the rights to emit pollutants could be traded between nations. This scheme, which is still in the process of being implemented, is discussed at length in Chapter 10.

*Sources*: Tietenberg (1990), Goodstein (1995), OECD (1999) new sources from being set up in, or near to, the airshed, and requiring existing sources to move away from the receptor areas. But what should the EPA do when the location of polluters and people is already determined, and moving either is not a feasible option?

When the location of sources is regarded as being fixed, pollution control must work by regulating in some way the emissions from those sources so as to meet the relevant air quality standards.[18] As we have been doing throughout this chapter, it is assumed here that targets have already been set. In this case, standards will consist of maximum allowable concentration rates of the stock pollutant in each of the relevant receptor areas. These targets may be 'efficient' targets (those we analysed in Chapter 6) or they may not. To the authors' knowledge, no targets for non-UMP have ever been set in terms of economic efficiency. So it will be sensible to deal with the case of arbitrary specific targets. For simplicity, we take the target to be the same for all receptors. Finally, we assume that in pursuit of its objectives the EPA seeks to reach targets at least cost.

Let us consider each of the following three instruments in turn:

1. non-transferable emissions licences allocated to each source (a command and control approach);
2. emissions taxes or emissions abatement subsidies;
3. marketable emissions permits.

### 7.5.1 Using non-transferable emissions licences

The use of non-transferable emissions licences is simple in principle. All that is required is for the EPA to calculate the maximum allowable emissions from each source so that the pollution target is reached in every receptor area, *and* at minimum possible overall cost. That is, the EPA needs to solve a cost-minimisation problem. Licences can then be allocated to each source in the quantities that emerge from the solution to that problem.

In order to obtain clear, analytical results, it is necessary to take the reader through the maths of this problem. That is done in Appendix 7.1. In the main text here, we just indicate the way in which the problem is set up, and interpret the main results obtained in Appendix 7.1. An Excel workbook (*Ambient instruments.xls*) provides a worked numerical example of the problem we are investigating.

As a prelude to doing this, it will be convenient to recap the notation we use for non-UMP. The airshed being considered contains $J$ spatially distinct pollution receptors (indexed $j = 1, 2, \ldots, J$) and $N$ distinct pollution sources (indexed $i = 1, 2, \ldots, N$). The transfer coefficient $d_{ji}$ describes the impact on pollutant concentration from source $i$ in receptor $j$. Pollution at location $j$, $A_j$, is the sum of the contributions to pollution at that location from all $N$ emission sources:

$$A_j = \sum_{i=1}^{N} d_{ji} M_i \tag{7.1}$$

where $M_i$ is emissions from source $i$. Section 6.6 provided much of the theoretical background for the case of non-UMP, but there is one major difference of emphasis between the approach we took there and the approach we adopt here. In Chapter 6, our interest was in target choice. To find the efficient emissions target, we maximised a net benefit function. Therefore, the solutions to that exercise give us the net benefit maximising level of emissions (for each source).

However, in this chapter our interest is not in target choice but rather in instrument choice. It is assumed that targets (for pollutant stocks in each receptor area) have already been set. As far as licences are concerned, our task is to find the level of emissions from each source that minimises the overall cost of reaching those targets. For tax (subsidy) instruments, our goal is to find the tax (subsidy) rate or rates that will reach those targets at least cost. We shall also be interested in how a marketable permit system could be designed in this case.

Let $A_j^*$ denote the EPA's target pollutant concentration at receptor $j$. (The symbol $A$ can be thought

[18] The terms 'targets' and 'standards' are being used synonymously here.

of as ambient air quality, another expression for the concentration rate of some relevant air pollutant.) For simplicity we suppose that the target for each receptor area is the same, so that $A_j^* = A^*$ for all $j$. The overall goal of the EPA is that in no area should the pollutant concentration exceed $A^*$. That is,

$$A_j = \sum_{i=1}^{N} d_{ji} M_i \leq A^* \quad \text{for } j = 1, \ldots, J \tag{7.2}$$

Next suppose that the EPA adopts one single criterion in pursuing its objective. It wishes to achieve the overall target (given in equation 7.2) at least cost. The solution (as we show in Part 8 of Appendix 7.1) requires that

$$\mathrm{MC}_i = \mu_1^* d_{1i} + \mu_2^* d_{2i} + \ldots + \mu_J^* d_{Ji},\ i = 1, 2, \ldots, N \tag{7.3}$$

where $\mathrm{MC}_i$ denotes the marginal abatement cost of firm $i$. We shall interpret equation 7.3 in a moment. Meanwhile, note that the systems 7.2 and 7.3 constitute $N + J$ equations which can be solved for the cost-minimising values of the $N + J$ unknowns ($N$ emissions levels and $J$ shadow prices).

To implement a non-transferable licence system to achieve the pollution targets at least cost, the $N$ values of $M_i^*$ need to be calculated, and licences allocated to firms accordingly. Note that even if firms have identical marginal abatement cost functions, they will *not* do equal amounts of emission abatement. This can be seen from the fact that the transfer coefficients on the right-hand side of 7.3 will vary from firm to firm. Hence the value of the whole expression on the right-hand side of 7.3 will differ between firms, and so their marginal abatement costs must differ too. That implies doing different amounts of abatement.

This may be compared with the condition that we found earlier for a uniformly mixing pollutant,

$$\mathrm{MC}_i = \mu^*,\ i = 1, 2, \ldots, N$$

which means that the marginal cost of emissions abatement is equal over all pollution sources. Hence, if firms had identical abatement cost functions they would do identical amounts of abatement. The intuition behind the result that firms will abate to different amounts where they emit non-UMP is simple. Emissions from some sources have more damaging consequences than emissions from others, because of the way in which emissions become distributed over the area of concern. Those sources whose emissions lead to relatively high damage should have relatively low emissions.

### 7.5.2 Using emissions taxes or emissions abatement subsidies

We now turn to consider a tax (or subsidy) instrument. This requires a bit more care in interpreting equation system 7.3. The $\mu_j^*$ terms that appear in each of the $N$ equations are shadow prices. There is one of these for each receptor area. Each denotes the monetary value of a worsening of the pollution stock by one unit in that area. The $d_{ij}$ coefficients tell us how many units pollution increases by in receptor $j$ if emissions from source $i$ rise by one unit. So for example $\mu_2^* d_{2i}$ gives the monetary value of damage that accrues in area 2 from an additional unit of emissions in source $i$. By summing these values over all source areas (that is, $\mu_1^* d_{1i} + \mu_2^* d_{2i} + \ldots + \mu_J^* d_{Ji}$) we find the total value of damage caused in all receptor areas by an additional unit of emission from $i$. Cost-efficiency requires that each firm pays a tax on each unit of emission, $t_i$, (or receives a subsidy on each unit abated, $s_i$) equal to the value of that damage, so we have

$$t_i = s_i = \mu_1^* d_{1i} + \mu_2^* d_{2i} + \ldots + \mu_J^* d_{Ji}$$

Note that the tax (subsidy) rate will now not be the same for each firm. This is just what we would expect for non-UMP as damage varies according to the location of emission source.

There is one important corollary of this. As tax or subsidy instruments require that rates are unique to each pollution source, one of the attractive features of these instruments (that a single rate can be applied over all polluters) no longer applies. Indeed, a single tax rate would *not* lead to a cost-effective abatement programme in this case.

If the EPA were determined to use a tax instrument, nonetheless, and tried to calculate the source-specific tax rates, it would require exactly the same amount of information as a command and control system does. In particular, it would need to know the marginal abatement cost function for every firm.

Hence a second desirable property of a tax instrument – that it does not need knowledge of an individual firm's costs – also disappears. All in all, one would expect much less use to be made of pollution tax or subsidy instruments in the case of non-uniformly-mixing air, water or ground pollution than with a uniformly mixing pollutant.

### 7.5.3 Using marketable emissions permits

How would marketable permits work in this case? The system – known as an ambient marketable permits or spatially differentiated system – would operate as follows:

1. Each receptor site will have a pollution concentration target. As before, we assume that this is the same for all receptors, $A^*$.
2. For each receptor site, the EPA must calculate how many units of emission can be allowed to arrive at that site before the pollution target is breached. More formally, it must calculate how many 'emissions permits' there can be that will allow firms to decrement (that is, worsen) ambient concentrations at that site.
3. These permits are issued to pollution sources, either by competitive auction or by free initial allocation ('grandfathering' if this is done proportionally to previous unregulated emission levels).
4. A pollution source is prohibited from making an emission to any receptor site above the quantity of permits it holds for emissions to that site. Each firm will, therefore, be required to hold a portfolio of permits to worsen concentrations at specific receptor areas.
5. A market for permits will emerge *for each receptor area*. Each polluting source will trade in many of these markets simultaneously. The results of these trades will determine a unique equilibrium price in each market.
6. Permits for each receptor area are freely marketable on a one-to-one basis, but this does not apply to permits for different receptors.

Note that 'emissions permits' have a special meaning in this context. They are not unrestricted rights to emit. Rather, they are rights to emit such that pollutant concentrations will worsen by a particular amount at a particular place. So, for example, if I want to emit one unit, and that will worsen pollution by 3 units at receptor 1 and by 4 units at receptor 2, I must buy a permit to worsen pollution (by those amounts) in each of the two markets.

How does this relate to equation 7.3? The $J$ shadow prices $\mu_j^*$ correspond to the equilibrium permit prices in each market. At the least-cost solution, a firm will equate the marginal cost of emissions abatement with the marginal cost of not abating the emission. The right-hand side of equation 7.3 gives this latter cost, which is a weighted sum of these permit prices. The weights attached to the permit price for receptor $j$ will be the impact that one unit of emissions has on pollutant concentration at site $j$. Thus the right-hand side gives the cost to the firm, in permit prices paid, for one unit of its emissions.

Clearly, the administration of an ideal least-cost marketable permit system is hugely demanding. However, it does have one major advantage over both command-and-control and tax/subsidy instruments: the EPA does not have to know the marginal abatement cost function of each firm in order to achieve the pollution targets at least cost. This is the major reason why emissions permits have attracted so much attention from economists, and why they are being introduced in practice in a form similar to that outlined above.

There are as yet no actual examples of systems that match this ideal form exactly. Existing permit systems are only approximations to the ideal type. The most important departure in practice is the absence of separate markets for permits for each receptor. (Systems in practice tend, instead, to have markets for each type of pollution generator.) You should be able to see that the absence of separate receptor markets may substantially increase the true cost of achieving pollution targets.

The extent to which an ideal least-cost marketable permit scheme would attain ambient standards at lower cost than some alternative instruments has been analysed by several authors. We outline one of these studies (Krupnick, 1986) in Box 7.9. Krupnick's study also highlights another matter of considerable importance: abatement costs can rise very sharply as the desired targets are progressively tightened.

**Box 7.9 Costs of alternative policies for the control of nitrogen dioxide in Baltimore**

Nitrogen dioxide ($NO_2$) is a good example of a non-uniformly-mixing pollutant. Alan Krupnick (1986) investigated the cost of meeting alternative one-hour $NO_2$ standards in the Baltimore area of the United States. He compared a variety of control programmes applied to 200 large emission point sources in the area. He identified 404 separate receptor areas in the region. Krupnick considered three alternative standards applied for each receptor area: 250, 375 and 500 $\mu g/m^3$ control.

Simulation techniques are used to estimate total abatement costs for each of several different policy instruments. We deal here with four of the cases that Krupnick investigated:

- the least-cost instrument: a spatially differentiated ambient-pollution marketable permits scheme of the type discussed in the text;
- a type-specific fee: an effluent charge with charges differentiated by source type (but not by receptor areas impacted);
- a uniform fee: an effluent charge not differentiated by source type (nor location of impact);
- a hybrid instrument, labelled RACT/least-cost: a mixture of command and control and incentive instruments. The RACT part takes the form of a technology standard (‘Reasonably Available Control Technology’) which is imposed on all firms. For firms that fail to meet (weaker) national air-quality standards, market incentives are used to induce further emissions reductions (the least-cost part).

The results of Krupnick’s simulations (for two ambient targets) are shown in Table 7.7. Numbers not in parentheses refer to the stricter target of 250 $\mu g/m^3$, those in parentheses the weaker target of 500 $\mu g/m^3$. These targets were selected in view of the fact that uncontrolled emissions led to high ambient pollution levels of around 700–800 $\mu g/m^3$ at several receptor sites, and technology studies suggest that targets stricter than around 190 $\mu g/m^3$ are unobtainable given the presence of the existing point sources.

Comparing first the costs of attaining different targets, Krupnick notes that ‘compliance costs rise steeply as the standard is tightened, regardless of the policy simulated. In the least-cost case, costs rise by a factor of 25 (from \$66 000 to \$1.633 million) when standards are halved (from 500 to 250 $\mu g/m^3$.’ The smaller proportionate increase in the hybrid case (RACT/least-cost) is due to the fact that the technology controls imposed by RACT give the firms little additional room for manoeuvre for further cost reductions when the standard is made stricter.

Notice that the emissions reduction is relatively small for the least-cost control compared with others. This happens because the target being sought is not a given total emissions reduction but a maximum ambient pollution standard over the whole area. Several of the instruments are inefficient (in abatement cost terms) because they operate in a more uniform manner than the spatially differentiated least-cost permit method. In so doing, the optimal distribution of abatement effort is not being applied, and excessive amounts of control are being adopted on many pollution sources.

For the type-specific fee, control costs are not much larger than for the least-cost method (and are identical for the weaker control). A fee that distinguishes between different types of polluter does seem able to mimic fairly well a proper spatially differentiated permit (or tax) approach. This is reassuring, as type-specific fees are likely to be used in practice instead of least-cost ambient permit methods as a result of their much greater simplicity. In contrast, note that when a uniform fee is imposed to achieve the stricter ambient standard (and where uniformity means that no effort is made to relate the charge to impact of emissions on ambient levels at various places) control costs increase very dramatically. A uniform fee can result in the largest emission reduction, but without doing any better in terms of ambient standards, and at hugely additional cost. Note, finally, that a single market emissions permit system would have an identical effect to that of a uniform fee. Spatially differentiating permit markets offers huge cost savings in principle.

*Table 7.7* Simulation results for the cost of meeting two ambient targets

| | Emissions reduction (%) | | Abatement costs \$US millions/year | |
|---|---|---|---|---|
| Least cost (ambient permits) | 32 | (6) | 1.663 | (0.066) |
| Type-specific fee | 34 | (6) | 1.719 | (0.066) |
| RACT/least cost | 42 | (36) | 2.200 | (1.521) |
| Uniform fee | 73 | (21) | 14.423 | (0.224) |

*Source*: Adapted from Krupnik (1986), Tables II and III

## 7.6 A comparison of the relative advantages of command and control, emissions tax, emission abatement subsidy and marketable permit instruments

In this section, we bring together a set of results obtained earlier in the chapter, and introduce a few additional results; all these are of benefit in assessing the relative merits of alternative pollution control instruments.

### 7.6.1 Cost-efficiency

We established earlier several results relating to cost-efficiency. To summarise, an emissions tax, emissions abatement subsidy or marketable permit system can achieve any emissions target at least cost. A command and control (CAC) regulation instrument may, but will not usually, be cost-efficient. In order to be cost-efficient, the EPA must know each polluter's marginal cost of abatement function so that an emission control can be calculated for each firm that will equalise marginal abatement costs. It is very unlikely that this requirement will be met. The conclusion we draw from this is that a command and control quantity regulation approach is inefficient relative to a tax, subsidy or marketable permit scheme, and so will achieve any specified target at a higher real cost. Some empirical evidence on this is presented in Box 7.10.

For a non-UMP, the remarks above need to be qualified. Cost-effective command and control systems, as before, require knowledge of individual firms' marginal cost of abatement functions. But so too do tax and subsidy instruments in this case. In general, only transferable permit schemes do not require that knowledge. This accords permit systems great potential advantages over others.

### 7.6.2 Monitoring, administering and enforcing compliance costs

Little or nothing has been said so far about the costs associated with monitoring, administering and enforcing compliance for each instrument. Yet these

**Box 7.10 The costs of emissions abatement using command and control and market-based instruments**

A substantial literature now exists on the comparative costs of attaining emissions abatement targets using traditional quantity or technology regulations – what we call command and control (CAC) instruments – and so-called market instruments (particularly emissions taxes, abatement subsidies and marketable/transferable emissions permits). Much of this literature derives from experience in the USA with these two categories of instrument. Tietenberg (1990) provides an admirable account of recent evidence on these costs. Table 7.8 reproduces one of Tietenberg's tables, showing the ratio of costs under CAC approaches to the least-cost controls (using market instruments) for air pollution control in the United States. We have examined one of these studies – that by Krupnick (1986) – in more detail in Box 7.9.

Although they can be 'best' instruments in some circumstances, such direct controls are often extremely costly. Tietenberg (1984) finds that the CAC approach costs from twice to 22 times the least-cost alternative for given degrees of control. These ratios suggest that massive cost savings might be available if market instruments were to be used in place of CAC. In his 1990 paper, Tietenberg reports estimates that compliance with the US Clean Air Act through market instruments has led to accumulated capital savings of over $10 billion. It should be pointed out, however, that most studies compare actual CAC costs with those theoretically expected under least-cost market-based instruments. In practice, one would not expect market instruments to operate at these theoretical minimum costs, and so the ratios we quoted above overstate the cost savings that would be obtained in practice by switching from CAC techniques.

Three arguments underlie the tenet that market-based incentive approaches are likely to be more efficient than regulation and control. First, markets are effective in processing information; second, market instruments tend

**Box 7.10 *continued***

to result in pollution control being undertaken where that control is least costly in real terms; and third, market-based approaches generate dynamic gains through responses over time to their patterns of incentives.

However, stringent conditions are necessary for markets to guarantee efficient outcomes. Policy instrument choice takes place in a 'second-best' world, where results are much less clear. The absence of markets (including those for externalities and public goods), asymmetric information, moral hazard and other instances of market failure, all point to possible benefits of CAC-based public intervention or to the inappropriateness of complete reliance on markets and market instruments. (See Fisher and Rothkopf (1989) for an excellent survey.)

A European example is given in the file *Agriculture.doc* in the *Additional Materials* for Chapter 7. A study by Andreasson (1990) examines the real resource costs of three different policies for reducing nitrate fertiliser use on the Swedish island of Gotland: non-marketable quotas on fertiliser use, a tax on nitrogenous fertiliser and a marketable permit system. Some additional references to studies which attempt to quantify the costs of attaining pollution standards using various instruments are given in the recommendations for further reading.

*Table 7.8* Empirical studies of air pollution control

| Study | Pollutants covered | Geographic area | CAC benchmark | Ratio of CAC cost to least cost |
|---|---|---|---|---|
| Atkinson and Lewis | Particulates | St Louis | SIP regulations | 6.00[a] |
| Roach *et al.* | Sulphur dioxide | Four corners in Utah | SIP regulations Colorado, Arizona, and New Mexico | 4.25 |
| Hahn and Noll | Sulphates standards | Los Angeles | California emission | 1.07 |
| Krupnick | Nitrogen dioxide regulations | Baltimore | Proposed RACT | 5.96[b] |
| Seskin *et al.* | Nitrogen dioxide regulations | Chicago | Proposed RACT | 14.40[b] |
| McGartland | Particulates | Baltimore | SIP regulations | 4.18 |
| Spofford | Sulphur dioxide | Lower Delaware Valley | Uniform percentage regulations | 1.78 |
| | Particulates | Lower Delaware Valley | Uniform percentage regulations | 22.0 |
| Harrison | Airport noise | United States | Mandatory retrofit | 1.72[c] |
| Maloney and Yandle | Hydrocarbons | All domestic DuPont plants | Uniform percentage reduction | 4.15[d] |
| Palmer *et al.* | CFC emissions from non-aerosol applications | United States | Proposed standards | 1.96 |

*Notes*:
CAC = command and control, the traditional regulatory approach.
SIP = state implementation plan.
RACT = reasonably available control technologies, a set of standards imposed on existing sources in non-attainment areas.
[a] Based on a 40 μg/m$^3$ at worst receptor.
[b] Based on a short-term, one-hour average of 250 μg/m$^3$.
[c] Because it is a benefit–cost study instead of a cost-effectiveness study the Harrison comparison of the command and control approach with the least-cost allocation involves different benefit levels. Specifically, the benefit levels associated with the least-cost allocation are only 82% of those associated with the command-and-control allocation. To produce cost estimates based on more comparable benefits, as a first approximation the least-cost allocation was divided by 0.82 and the resulting number was compared with the command-and-control cost.
[d] Based on 85% reduction of emissions from all sources.
*Source*: Tietenberg (1990), Table 1

costs could be quite substantial. If they are large, and if they differ significantly between instruments, these costs are likely to have an important bearing on which type of instrument is least-cost for achieving some target. One reason for the prevalence of minimum technology requirements as a pollution control instrument may be that these costs are low relative to those of instruments that try to regulate emissions output levels.

### 7.6.3 Long-run effects

From the point of view of the EPA, instrument selection will depend on the degree to which the amount of pollution control varies with the passage of time for any particular instrument. An important consideration concerns whether or not the long-run effect is markedly different from the short-run effect. The long-run effect of an instrument depends mainly on two things: net income effects and technological innovation effects. We consider each of these in turn.

#### Net income effects

Changes in net income arising from the operation of a pollution control instrument can affect the long-run industry size. We noted earlier that subsidy schemes may have the (environmentally) undesirable property of increasing the long-run size of the targeted industry through positive income effects. Similar issues were raised when we were comparing alternative methods of initially allocating marketable permits.

Of course, it is possible in principle to design control regimes that are revenue-neutral. For example, firms in a subsidised industry may be required to make lump-sum payments which sum to the total value of subsidies. This would preserve the incentive effects of subsidy systems without allowing long-run effects arising from income changes. However, it may be politically difficult to implement such a scheme, and there may be reasons why government does not wish to match receipts and payments in such a way.

#### Technology effects

A second route through which long-run effects may transmit is via induced impacts on the rate of technological innovation. There are two aspects to this. One concerns what are sometimes called *dynamic efficiency* effects. These arise from the pattern of incentives to innovate generated by a pollution control instrument. A common argument in this regard is that command and control instruments have poor long-run properties because they generate weak incentives for innovation (see, for example, Jaffe and Stavins, 1994). The binary nature of many such instruments (you reach the target or you do not reach it) creates a discrete switch in behaviour: once a required target has been obtained there is no longer any incentive to go further.

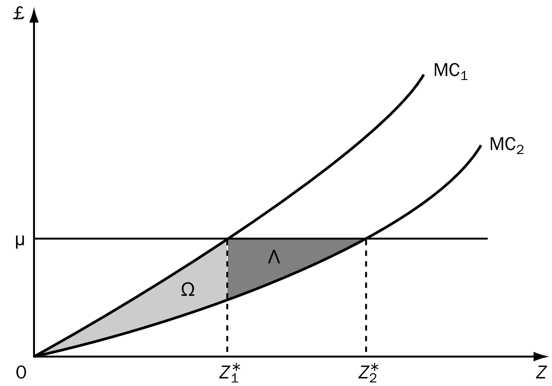


*Figure 7.9* Dynamic incentives under emissions tax controls

In contrast (it is argued) an emissions tax (or abatement subsidy) will generate a dynamically efficient pattern of incentives on corporate (and consumer) behaviour. The incentive structure operates to continually reward successful environmentally friendly innovation. In a market-based scheme, every unit of emissions reduction is rewarded by a tax saving. The key issue here is what incentives firms face in developing pollution-saving technology or developing new, environmentally cleaner products. Under a emissions tax scheme, these incentives may be strong, as we show in Figure 7.9.

Area Ω is the saving that would result if marginal costs were lowered from $MC_1$ to $MC_2$ and the emissions level were unchanged. But if marginal cost were lowered in this way, the firm's profit-maximising emissions abatement level would rise from $Z_1^*$ to $Z_2^*$, and so an additional saving of Λ would accrue to the firm. The firm has an incentive to develop new technology to abate emission if the total costs of developing and applying the technology are less than the present value of the savings $\Omega + \Lambda$ accumulated over the life of the firm.[19] In contrast, in a CAC regulatory system, dynamic incentives are weaker or non-existent. As we said above, if a target is set in (non-marketable) quantitative terms, then once that target has been met there is little or no further incentive on the polluter to reduce emissions.

But there is a second aspect that weakens the force of these arguments. Some researchers believe that technological change can be driven from above. Suppose that the EPA identifies best-practice environmentally friendly technology, and imposes this as a requirement on firms through minimum acceptable technology regulations. Not only will this have a direct effect on spreading technology diffusion, but the indirect effects may be powerful too. Barriers due to frictions, lack of information, and other market imperfections that may lead firms to be overcautious or unable to act voluntarily no longer bite in the face of imposed requirements. Moreover, these changes have catalytic effects which set in motion spurts of innovation as learning effects occur. These kinds of arguments are likely to have most relevance for technological innovation and diffusion in developing economies.

It is difficult to arrive at unequivocal conclusions from all this. However, a reasonable conclusion must be that, in some circumstances at least, technology-based controls and other command and control instruments will have superior long-run properties to market-based instruments.

### 7.6.4 Double dividend

In the previous chapter, we noted the possibility that some environmental regulation schemes may generate a so-called double dividend. It seems likely that the availability and size of a double dividend will vary from one circumstance to another, and on which instrument is being used. A sensible choice of instrument should consider these matters.

### 7.6.5 Equity/distribution

Finally, we note that the distributional consequences of a pollution control policy instrument will be very important in determining which instruments are selected in practice. Different instruments for pollution control have different implications for the distribution of income within an economy. We have already examined the direct business financial gains and losses (which are, of course, exactly mirrored by offsetting government financial losses or gains). It is also necessary to think about the consequences for income and wealth distribution in society as a whole. For example, an emissions tax imposed upon fossil fuels will indirectly affect final consumers who purchase goods that have large energy input. Individuals for whom heating comprises a large proportion of their budget may well experience quite large falls in real income. Indeed, many kinds of 'green taxes' are likely to have regressive effects upon income distribution.

It is important to distinguish between income shifts that are merely redistributive and do not correspond to any real resource gains and losses to the economy, and real income changes which do imply real resource changes for the economy as a whole. The latter arise because pollution control does involve real costs. Of course, by having less pollution, there are benefits to be had as well, which in a well-designed pollution control programme should outweigh these real costs. Nevertheless, the beneficiaries and losers will not be the same individuals, and it is this that one is concerned with when discussing the equity or fairness of an instrument.

It should also be noted that emissions taxes (and other environmental controls) have important implications for the relative competitiveness of national economies. (See Chapter 10 for more on this.) Some analysts have advocated a switch from taxes on labour and capital to taxes on emissions to avoid excessive tax burdens, and schemes have been proposed to penalise nations that attempt to gain competitive advantage by not introducing emissions taxes. Good discussions of these issues are to be

[19] Note that the optimal tax rate would change as new technology lowers control costs, so matters are a little more complicated.

found in Bertram *et al*. (1989), Brown (1989), Grubb (1989a), Hansen (1990), Kosmo (1989) and Weizsäcker (1989).

As we noted earlier, where a particular instrument has an adverse financial effect on one sector of the economy, it is open to the government to use compensating fiscal changes to offset those changes so that the distribution of income and wealth between individuals is not systematically changed. For example, the financial transfers implied by a emissions tax scheme could be compensated by lump-sum payments to firms or by abatement subsidy payments. And income transfers from poorer groups facing higher energy bills, for example, could be compensated for by other fiscal changes.

The main point here is that additional tax revenues received by government could be distributed to groups adversely affected by the initial policy change. However, the difficulties in designing distributionally neutral packages are immense. Where compensation is paid to individuals or groups for whom the tax incidence is considered excessive, the form of compensation should be designed not to alter behaviour, otherwise the efficiency properties of the instrument will be adversely affected. This implies lump-sum compensation should be used where possible. Compensation schemes of this form rarely happen in practice. Nevertheless, decision makers do have this option; whether they choose to exercise it is another matter.

## Summary

- An instrument that attains a pollution target at least cost is known as a cost-effective instrument.
- A least-cost control regime implies that the marginal cost of abatement is equalised over all firms undertaking pollution control.
- Bargaining processes might bring about efficient outcomes (and so might lead to the attainment of targets without regulatory intervention).
- The likelihood of efficient bargaining solutions to pollution problems being achieved is reduced by the presence of bargaining costs, and if bargaining would take place over a public (as opposed to a private) good.
- Pollution control instruments can be classified into a set of broad classes, the most important of which are command and control instruments and economic incentive-based instruments.
- In many – but not all – circumstances, economic incentive-based instruments are more cost-effective than command and control instruments.
- The long-run effects of pollution control instruments can be very different from their short-run effects, because of net income effects and impacts on the rate and direction of technological change.
- Where a pollutant is not uniformly mixing, the relative advantages of incentive-based instruments are considerably reduced. Some forms of marketable permit systems appear to offer the best prospect of attaining ambient pollution targets cost-effectively.
- Our discussion of the properties and relative advantages of various instruments that could be used to attain environmental policy targets has taken place under the implicit assumption that some single authority has the ability to implement and administer a control programme. But many pollution problems spill over national boundaries. Given that the world does not have a single government, how can policy targets and instruments be devised, introduced, administered and monitored for global or international pollution problems? This question warrants separate attention in Chapter 10.

## Further reading

Where a reference is underlined below, it is available online; the URL is given in the References.

Baumol and Oates (1988) is a classic source in the area of environmental regulation. The whole book is relevant but it is quite difficult and formal. The theoretical basis for a political economy of environmental regulation is investigated in Boyer and Laffont (1999). Tietenberg (1992, chapters 14 to 20) provides an extensive and primarily descriptive coverage of specific types of pollution and the control techniques applied to each. Other good general accounts of pollution control policy are to be found in Fisher (1981, chapter 12), which discusses the work of Ronald Coase and the roles of wealth and bargaining power, Common (1995), Hartwick and Olewiler (1986) and Goodstein (1995). Fisher and Rothkopf (1989) consider the justification for public policy in terms of market failure. A possibility, that we touch upon in the next chapter, is that public intervention itself generates substantial costs. These costs may be sufficiently large to prevent intervention delivering positive net benefits. This notion of 'government failure' is analysed in Weimer and Vining (1992). Laffont and Tirole (1993, 1996) discuss the innovation incentive effects of permits when number is limited.

There are several national and international agencies that produce periodic surveys of environmental protection instruments and their effectiveness. Among these are various parts of the United Nations Organisation, the European Union, the United States EPA and the OECD. An extensive listing can be found on the Chapter 7 Links web page. References that the reader may find useful include OECD (1995), which surveys the use of environmental taxes and other charges used for environmental protection in the OECD countries; Anderson *et al.* (1997), US experience with economic incentives instruments; OECD (1997d), evaluating economic instruments for environmental policy; OECD (1999) for a detailed account of instruments used – and their effectiveness – in OECD countries; EPA (1999), economic incentives for pollution control in the USA; EPA (2001), US experience with economic incentives; EEA (2001), which considers ways of improving official environmental reporting; and EEA (2000), an online survey of environmental taxes in the EU.

Pearce and Brisson (1993) discuss the use of command and control instruments in the UK. Bohm (1981) considers deposit refund systems. Helm (1993, 1998) discusses possible reform of environmental regulation in the UK. Smith (1998) investigates taxation of energy. Portney (1990) analyses air pollution policy in the USA, and Portney (1989) assesses the US Clean Air Act. Crandall (1992) provides an interesting analysis of the relative inefficiency of a standards-based approach to fuel efficiency in the United States. Kolstad (1987) examines the inefficiency losses associated with using undifferentiated taxes or other charges when economic efficiency requires that charges be differentiated across sources. Krupnick's (1986) paper on nitrogen dioxide control in Baltimore, discussed in the chapter, repays reading in the original.

Dales (1968) is the paper generally credited with having established the notion that marketable permits may be used for pollution control, and Montgomery (1972) derived the efficiency properties of marketable permits. For accounts of the use of market-based pollution control instruments see Hahn (1984, 1989), Hahn and Hester (1989a, b), Opschoor and Vos (1989) and Tietenberg (1990, 1992). Jorgensen and Wilcoxen (1990a, b, c) analyse the impact of environmental regulation upon economic growth in the United States (but note that these papers are relatively difficult).

The following references deal with air pollution emissions trading programmes in developing countries: Ellerman (2001), $SO_2$ emissions in China; Blackman and Harrington (1999); Benkovic and Kruger (2001); Montero *et al.* (2000), Chile; and several papers in the *Journal of Economic Perspectives* (Summer 1998, Vol. 12, no 3). Some general accounts of air emissions problems and policies in India are found in Bose *et al.* (1997, 1998). Cowan (1998) considers the use of economic instruments for water pollution and abstraction.

Enforcement issues and incentive compatibility (to be discussed in the next chapter) are analysed in Heyes (1998) and Laplante and Rilstone (1996). For a detailed analysis of issues concerning compensation in connection with distribution effects of tax changes, see Hartwick and Olewiler (1986, chapter 12), who also analyse the consequences of subsidies and taxes in the short run and the long run. The role and importance of non-convexities are discussed in Fisher (1981, p. 177), Portes (1970) and Baumol and Oates (1988). Second-best aspects of taxation, and possible double dividends from environmental policy, are discussed in Cremer and Gahvani (2001) and Parry *et al*. (1999).

The seminal text on non-point pollution is Russell and Shogren (1993). Others on this topic include Dosi and Tomasi (1994), Braden and Segerson (1993), Laffont (1994), Millock *et al*. (1997), Romstad *et al*. (1997), Segerson (1988) and Shogren (1993). For water pollution see Segerson (1990) and Ribaudo *et al*. (1999), and for non-point pollution from agriculture Vatn *et al*. (1997).

Useful accounts of instruments used in fisheries management include OECD (1997c) and the regular OECD publication *Review of Fisheries*, which covers changes in fishery management systems. Discussion of the idea of a safe minimum standard of conservation can be found in Bishop (1978) and Randall and Farmer (1995). The 'Blueprint' series (see, for example, Pearce, 1991a) provides a clear and simple account of the new environmental economics policy stance, in a rather ideological style. Finally, a number of texts provide collections of papers, several of which are relevant to pollution control policy: these include Bromley (1995) and, at a more rigorous level, the three 'Handbooks' edited by Kneese and Sweeney (1985a, b, 1993).

## Discussion questions

1. Suppose that the EPA obtains damages from polluting firms in recompense for the damage caused by the pollution. Should the EPA distribute the moneys recovered from such damage settlements to the pollution victims? (Hint: consider, among other things, possible changes in victim behaviour in anticipation of such compensation.)
2. Consider a good whose production generates pollution damage. In what way will the effects of a tax on the output of the good differ from that of a tax on the pollutant emissions themselves? Which of the two is likely to be economically efficient? (Hint: think about substitution effects on the demand side and on the supply side.)
3. Evaluate the arguments for the use of market or incentive-based instruments versus 'command and control' instruments in the regulation of environmental externalities under conditions of certainty.
4. Discuss the scope for the allocation of private property rights to bring the privately and socially optimal levels of soil pollution into line.
5. Discuss the distributional implications of different possible methods by which marketable permits may be initially allocated.
6. Distinguish between private and public goods externalities. Discuss the likelihood of bargaining leading to an efficient allocation of resources in each case.
7. Use diagrams to contrast pollution tax instruments with marketable emission permit systems, paying particular attention to the distributional consequences of the two forms of instrument. (Assume a given, target level of pollution abatement, and that permits are initially distributed through sale in a competitive market.)
8. Discuss the efficiency properties of a pollution tax where the tax revenues are earmarked in advance for the provision of subsidies for the installation of pollution abatement equipment.
9. Suppose that a municipal authority hires a firm to collect and dispose of household waste. The firm is paid a variable fee, proportional to the quantity of waste it collects, and is charged a

fee per unit of waste disposed at a municipal waste landfill site. Households are not charged a variable fee for the amount of waste they leave for collection, instead they pay an annual fixed charge. Comment on the economic efficiency of these arrangements and suggest how efficiency gains might be obtained.

10. An interesting example of a regulatory failure relates to electricity generating stations in the UK. Several thermal power stations in the UK were required to install flue-gas desulphurisation (FGD) plant in order to meet a national standard for sulphur emissions. The power stations fitted with FGD plant are not compensated for sulphur abatement. Electricity is purchased for the national grid on a competitive bidding system. The stations fitted with FGD are unable to compete on cost with other stations without that equipment, and as a result are withdrawn entirely from the grid at some times and operate below capacity at others.

    Explain why this situation is socially inefficient, and suggest a means by which this inefficiency could be avoided.

## Problems

1. Suppose that an EPA must select one instrument from two available. Two criteria matter: (a) $P$, the probability of the instrument attaining its target; (b) $C$, the proportionate saving in abatement cost incurred in using that instrument (relative to the cost using the highest-cost instrument). The EPA calculates a weighted sum (score) for each instrument, and chooses that with the highest score. Assume that the instruments have the following values for $P$ and $C$:

   Instrument 1: $P = 0.9$, $C = 0.0$

   Instrument 2: $P = 0.7$, $C = 0.2$

   (i) Write an Excel spreadsheet to illustrate how the instrument choice varies with changes in the relative weights (between zero and one) attached to the two criteria. Also explore how instrument choice varies as the magnitudes of $P$ and $C$ for each instrument vary.

   (ii) Use an algebraic formulation of this problem to obtain expressions that allow these results to be shown analytically.

2. Using the Excel workbook *Leastcost.xls*, demonstrate that the cost penalty from sharing abatement equally between the two firms rather than using the least-cost distribution of abatement is larger the greater is the difference in the firms' abatement cost functions (as measured by the value of the slope parameter in the abatement cost functions).

3. The Coase theorem claims that a unique and efficient allocation of resources would follow from rational bargaining, irrespective of how property rights were initially allocated. Demonstrate that the distribution of net gains between bargaining parties will, in general, depend upon the initial distribution of property rights.

4. Show that a pollution tax on emissions and a subsidy to output producers for each unit of pollution reduction would, if the rates of subsidy were identical to the pollution tax rate, lead to identical outcomes in terms of the levels of output and pollution for a given sized industry. Explain why the distribution of gains and losses will usually differ, and why the long-run level of pollution abatement may differ when the industry size may change.

5. In all discussions of pollution abatement costs in this chapter, the fixed costs of pollution abatement were implicitly taken to be zero. Do any conclusions change if fixed costs are non-zero?

6. Demonstrate that in the simple special case of a uniformly mixing flow pollutant, in which the value of the damage created by the

emission is independent of the location of the emission source or the time of the emission, the tax rate should be uniform over all polluters for the tax to be an efficient instrument (that is, it will be applied at the same rate per unit of pollution on all units of the pollutant).

7. Our discussion in this chapter has shown that if the control authority does not know the marginal damage function, it will not be able to identify the economically efficient level of pollution abatement, nor the efficient tax or subsidy level. Demonstrate that
   (a) knowledge of the pollution abatement schedule alone means that it can calculate the required rate of tax to achieve any target level it wishes,
   (b) if it knew neither the marginal damage nor the marginal abatement cost schedules, then it could arbitrarily set a tax rate, confident in the knowledge that whatever level of abatement this would generate would be attained at minimum feasible cost.

8. You are given the following information:
   (a) A programme of air pollution control would reduce deaths from cancer from 1 in 8000 to 1 in 10 000 of the population.
   (b) The cost of the programme is expected to lie in the interval £2 billion (£2000 million) to £3 billion annually.
   (c) The size of the relevant population is 50 million persons.
   (d) The 'statistical value' of a human life is agreed to lie in the interval £300 000 to £5 million.

   If the only benefit from the programme is the reduced risk of death from cancer, can the adoption of the programme be justified using an economic efficiency criterion?

9. In controlling emissions, there is an important difference between a command and control instrument and a tax instrument. Both require that the polluter pay the cost of attaining the emission reduction target. However, the tax instrument imposes an additional charge (on the emissions which remain at the target level of pollutions); this is not paid under a command and control regime. The failure to incorporate damage costs into the price of the product can generate distortions or inefficiencies in the economy. Kolstad (2000), from which this problem is drawn, gives an example in the paper manufacturing industry. Suppose that paper can be produced using pulp either from recycled paper (which is non-polluting) or from virgin timber (which is polluting). Compare the operation of a CAC instrument with a tax instrument applied to the manufacture of pulp from virgin timber, and show how this distorts (creates an inefficiency) in paper production.

10. This exercise involves using an Excel file to undertake some simulations regarding the relative costs of alternative instruments, and to interpret and comment on your results. Instructions for the exercise are given in *Pollution2.doc*; the Excel file is *Pollution2.xls*. Both of these can be found in the *Additional Materials* for Chapter 7.

## Appendix 7.1 The least-cost theorem and pollution control instruments

This appendix is structured as follows. In Part 1, we define the notation used and set the scene for what follows. Then in Part 2 we derive a necessary condition for pollution control to be cost-effective: that is, to attain any given target at least cost. An EPA has several instruments available for attaining a pollution (or pollution abatement) target. Here we consider three classes of instrument: quantitative regulations (a variant of command and control) in Part 3; an emissions tax (Parts 4 and 5); an emissions abatement subsidy (Part 6); and transferable emissions permits (Part 7). Collectively, Parts 3 to 7 take the reader through what an EPA would need to know, and how it could operate each of those instruments, in order to achieve a target at least cost. Finally in Part 8, we generalise previous results to the case of a non-uniformly-mixing pollutant.

## Part 1 Introduction

There are $N$ polluting firms, indexed $i = 1, \ldots, N$. Each firm faces a fixed output price and fixed input prices, and maximises profits by an appropriate choice of output level ($Q_i$) and emission level ($M_i$). Emissions consist of a uniformly mixing pollutant, so that the source of the emission is irrelevant as far as the pollution damage is concerned.

Let $\hat{\Pi}_i$ be the maximised profit of the $i$th firm in the absence of any control over its emission level and in the absence of any charge for its emissions. This is its unconstrained maximum profit level. At this unconstrained profit maximum the firm's emission level is $\hat{M}_i$.

Let $\Pi_i^*$ be the maximised profit of the $i$th firm when it is required to attain a level of emissions $M_i^* < \hat{M}_i$. This is its constrained maximum level of profits. To reduce emissions, some additional costs will have to be incurred or the firm's output level must change (or both). The constrained profit level will, therefore, be less than the unconstrained profit level. That is, $\Pi_i^* < \hat{\Pi}_i$.

We next define the firm's abatement costs, $C$, as constrained minus unconstrained profits:

$$C_i = \hat{\Pi}_i - \Pi_i^*$$

Abatement costs will be a function of the severity of the emissions limit the firm faces; the lower is this limit, the greater will be the firm's abatement costs. Let us suppose that this abatement cost function is quadratic. That is

$$C_i = \alpha_i - \beta_i M_i^* + \delta_i M_i^{*2} \tag{7.4}$$

We illustrate this abatement cost function in Figure 7.10. Note that that the abatement cost function is defined only over part of the range of the quadratic function. Abatement costs are zero when the emission limit is set at $\hat{M}_i$, the level the firm would have itself chosen to emit in the absence of control. Abatement costs are maximised when $M_i^* = 0$, and so the firm is prohibited from producing any emissions.

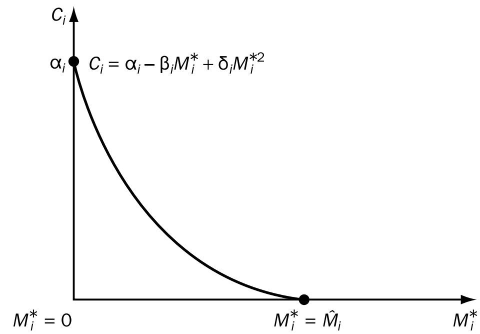


*Figure 7.10* The firm's abatement cost function

Two things should be said about equation 7.4. First, as each parameter is indexed by $i$, abatement costs are allowed to vary over firms. Second, the arguments that follow do not depend on the abatement cost function being quadratic. We have chosen that functional form for expositional simplicity only.

## Part 2 The least-cost theorem

We now consider the problem of an environmental protection agency (EPA) meeting some standard for total emissions (from all $N$ firms) at the least cost. Let $M^*$ denote the predetermined total emission target. In the expressions that follow, the $M_i^*$ variables are to be interpreted as endogenous, the values for which are not predetermined but emerge from the optimising exercise being undertaken. The problem can be stated as

$$\text{Min} \sum_{i=1}^{N} C_i \text{ subject to } M^* = \sum_{i=1}^{N} M_i^* \tag{7.5}$$

The Lagrangian for this problem is

$$\begin{aligned} L &= \sum_{i=1}^{N} C_i + \mu\left(M^* - \sum_{i=1}^{N} M_i^*\right) \\ &= \sum_{i=1}^{N} (\alpha_i - \beta_i M_i^* + \delta_i M_i^{*2}) - \mu\left(M^* - \sum_{i=1}^{N} M_i^*\right) \end{aligned} \tag{7.6}$$

The necessary conditions for a least-cost solution are

$$\frac{\partial L}{\partial M_i^*} = -\beta_i + 2\delta_i M_i^* + \mu^* = 0, \quad i = 1, 2, \ldots, N \tag{7.7}$$

and

$$\frac{\partial L}{\partial \mu} = -M^* + \sum_{i=1}^{N} M_i^* = 0 \qquad (7.8)$$

Equations 7.7 and 7.8 give $N + 1$ equations in $N + 1$ unknowns. Solving these simultaneously gives each firm's emission limit, $M_i^*$ (which now should be regarded as the *optimised* emissions limit for the firm), and the optimised shadow price of the pollution constraint (the Lagrange multiplier) $\mu^*$. Since $\mu^*$ is constant over all firms, it can be seen from equation 7.7 that a least-cost pollution abatement programme requires that the marginal cost of abatement be equal over all firms.

There is a tricky issue relating to signs in equation 7.7. Notice that an increase in $M_i^*$ corresponds to a relaxation of a pollution target (a decrease in required abatement) so the term $(-\beta_i + 2\delta_i M_i^*)$ is the marginal cost of a *reduction in pollution abatement* being required of firm $i$. It will therefore be a negative quantity. This can be verified by looking at the slope of the $C_i$ function in Figure 7.10.

By multiplying equation 7.7 through by minus one, we obtain

$$\beta_i - 2\delta_i M_i^* = \mu^* \qquad (7.7')$$

Here the term on the left-hand side $(\beta_i - 2\delta_i M_i^*)$ is the firm's marginal cost of an increase in pollution abatement, a positive quantity. It follows from 7.7′ that $\mu^*$ is also a positive quantity. This is consistent with the text of this chapter and the previous one, and matches, for example, the graphic in Figure 7.4.

## Part 3 Least-cost pollution control using quantitative regulation

If the EPA knew each firm's abatement cost function (that is, it knew $C_i$ for $i = 1, \ldots, N$), then for any total emission standard it seeks, $M^*$, the system of equations 7.7 and 7.8 could be solved for $M_i^*$ for each firm. The EPA could then tell each firm how much it could emit. The total quantity of emissions would, from equation 7.8, be reached exactly, and the target would, as the above theorem shows, be attained at least cost.

## Part 4 Least-cost pollution control using an emissions tax

As an alternative to setting quantitative emissions controls on each firm, an emission tax could be used. If the EPA knew each firm's abatement cost function, then for any total emission standard it seeks, $M^*$, the system of equations 7.7 and 7.8 could be solved for the value of the shadow price of the pollution constraint, $\mu^*$. Note that, unlike $M_i^*$, this shadow price is constant for each firm. The EPA could then set a tax at a rate of $t^*$ per unit of emissions and charge each firm this tax on each unit of pollution it emitted. Profit-maximising behaviour would then lead each firm to produce $M_i^*$ emissions, the least-cost solution.

To see why this should be so, note that in the absence of any quantity constraint on emissions, profit-maximising behaviour in the face of an emissions tax implies that the firm will minimise the sum of its abatement costs and pollution tax costs. That is, the firm chooses $M_i$ to minimise $CT_i$, the total of its abatement and tax costs:

$$CT_i = C_i + tM_i = \alpha_i - \beta_i M_i + \delta_i M_i^2 + t^* M_i$$

The necessary condition is

$$\frac{\partial CT_i}{\partial M_i} = -\beta_i + 2\delta_i M_i^* + t^* = 0, \quad i = 1, 2, \ldots, N \qquad (7.9)$$

Clearly, if $t^*$ in equation 7.9 is set equal to $\mu^*$ in equation 7.7, the necessary conditions 7.7 and 7.9 are identical, and so the tax instrument achieves the total emissions target at least cost.

## Part 5 What role is there for a tax instrument where each firm's abatement cost functions are not known?

In general, the EPA will not know abatement costs. However, if an arbitrarily chosen tax rate, say $\bar{t}$, is selected, and each firm is charged that rate on each unit of emission, then *some* total quantity of emissions, say $\bar{M}$, will be realised at least cost. Of course, that amount $\bar{M}$ will in general be different from $M^*$. Only if $\bar{t} = t^*$ will $\bar{M}$ be identical to $M^*$. An iterative, trial-and-error process of tax rate change may enable the EPA to find the necessary tax rate to achieve a specific target.

### Part 6 Least-cost pollution control using an emissions-abatement subsidy

Another method of obtaining a least-cost solution to an emissions target is by use of abatement subsidies. Suppose a subsidy of $s^*$ is paid to each firm on each unit of emissions reduction below its unconstrained profit-maximising level, $\hat{M}_i$. Then profit-maximising behaviour implies that the firm will maximise total subsidy receipts less abatement costs. That is, the firm maximises

$$\begin{aligned} CS_i = s(\hat{M}_i - M_i) - C_i = s(\hat{M}_i - M_i) \\ - (\alpha_i - \beta_i M_i + \delta_i M_i^2) \end{aligned}$$

The necessary condition is

$$\frac{\partial CS_i}{\partial M_i} = \beta_i - 2\delta_i M_i^* - s = 0, \quad i = 1, 2, \ldots, N \tag{7.10}$$

which, after multiplying through by $-1$, is identical to equation 7.9 if $s = t$. So, once again, if $s$ in equation 7.10 is set equal to $\mu^*$ in equation 7.7, the necessary conditions 7.7 and 7.10 are identical, and so the subsidy instrument achieves the total emissions target at least cost. Moreover, this result demonstrates that in terms of their effects on emissions, a tax rate of $t$ per unit of emissions is identical to a subsidy rate of $s$ per unit of emissions abatement, provided $s = t$.

### Part 7 Least-cost pollution control using transferable emissions permits

Suppose that the EPA issues to each firm licences permitting $L_i^0$ units of emissions. Firms are allowed to trade with one another in permits. The $i$th firm will trade in permits so as to minimise the sum of abatement costs and trade-acquired permits:

$$\begin{aligned} CL_i &= C_i + P(L_i - L_i^0) \\ &= \alpha_i + \beta_i M_i + \delta_i M_i^2 + P(L_i - L_i^0) \end{aligned} \tag{7.11}$$

where $P$ is the market price of one emission permit. Given that $L_i$ is the quantity of emissions the firm will produce after trade we can write this as

$$\begin{aligned} CL_i &= C_i + P(L_i - L_i^0) \\ &= \alpha_i - \beta_i L_i + \delta_i L_i^2 + P(L_i - L_i^0) \end{aligned} \tag{7.12}$$

The necessary condition for minimisation is

$$\frac{\partial CL_i}{\partial L_i} = -\beta_i + 2\delta_i L_i^* + P = 0, \quad i = 1, 2, \ldots, N \tag{7.13}$$

which can be interpreted as the firm's demand function for permits.

If the EPA sets a total emissions target of $M^*$ then M* is the total supply of permits and

$$M^* = \sum_{i=1}^{N} L_i^0 = \frac{\partial L}{\partial \mu} = \sum_{i=1}^{N} L_i \tag{7.14}$$

Now compare equations 7.13 and 7.14 with equations 7.7 and 7.8. These are identical if $P = \mu^*$ (remembering that $L_i = M_i^*$). Moreover, comparison of equation 7.13 with equations 7.11 and 7.12 shows that $P = t = s$. So by an initial issue of permits (distributed in any way) equal to the emissions target, the EPA can realise the target at least cost. Moreover, it can do so without knowledge of individual firms' abatement cost functions.

### Part 8 Least-cost abatement for a non-uniformly-mixing pollutant

The target of the EPA is now in terms of ambient pollution levels rather than emission flows. Specifically the EPA requires that

$$A_j = \sum_{i=1}^{N} d_{ji} M_i \le A_j^* \quad \text{for } j = 1, \ldots, J \tag{7.15}$$

The problem for the EPA is to attain this target at least cost. We deal with the case where the same ambient target is set for each receptor area. This problem can be stated as

$$\text{Min} \sum_{i=1}^{N} C_i \text{ subject to } A_j = \sum_{i=1}^{N} d_{ji} M_i \le A^* \quad \text{for } j = 1, \ldots, J \tag{7.16}$$

The Lagrangian for this problem is

$$L = \sum_{i=1}^{N} C_i - \mu_1 \left( A^* - \sum_{i=1}^{N} d_{1i} M_i \right) - \ldots - \mu_J \left( A^* - \sum_{i=1}^{N} d_{Ji} M_i \right) \tag{7.17}$$

where $C_i = \alpha_i - \beta_i M_i + \delta_i M_i^2$

The necessary conditions for a least-cost solution are

$$\frac{\partial L}{\partial M_i^*} = -\beta_i + 2\delta_i M_i + \sum_{j=1}^{j=J} (\mu_j^* d_{ji}) = 0,$$
$$i = 1, 2, \ldots, N \tag{7.18}$$

and

$$\frac{\partial L}{\partial \mu_j} = -A^* + \sum_{i=1}^{N} d_{ji} M_i = 0 \quad \text{for } j = 1, \ldots, J \tag{7.19}$$

The system of equations 7.18 and 7.19 consists of $N + J$ equations which can be solved for the $N + J$ unknowns ($M_i^*$, $i = 1, \ldots, N$ and $\mu_j^*$, $j = 1, \ldots, J$).

Equation 7.18 can be written as

$$-\beta_i + 2\delta_i M_i = -\sum_{j=1}^{j=J} (\mu_j^* d_{ji}), \quad i = 1, 2, \ldots, N \tag{7.20}$$

Then after multiplying through by −1, using $MC_i$ to denote the $i$th firm's marginal cost of abatement, and expanding the sum on the right-hand side, we obtain

$$MC_i = \mu_1^* d_{1i} + \mu_2^* d_{2i} + \ldots + \mu_J^* d_{Ji},$$
$$i = 1, 2, \ldots, N \tag{7.21}$$

The pair of equations 7.20 and 7.21 can be compared with the solution for the uniformly mixing pollution case, equation 7.7 multiplied by −1.