ECONOMICS OF NATURAL RESOURCES AND THE ENVIRONMENT

David W. Pearce

and R. Kerry Turner

The Johns Hopkins University Press
Baltimore

4 · THE OPTIMAL LEVEL OF POLLUTION

4.1 POLLUTION AS EXTERNALITY

The economic definition of pollution is dependent upon *both* some *physical* effect of waste on the environment *and* a human reaction to that physical effect. The physical effect can be biological (e.g. species change, ill-health), chemical (e.g. the effect of acid rain on building surfaces), or auditory (noise). The human reaction shows up as an expression of distaste, unpleasantness, distress, concern, anxiety. We summarise the human reaction as a *loss of welfare*. As Chapter 2 indicated, terms such as 'utility' or 'satisfaction' are, for our purposes, synonymous with welfare.

We now need to distinguish two possibilities for the economic meaning of pollution. Consider an upstream industry, which discharges waste to a river, causing some loss of dissolved oxygen in the water. In turn, suppose the oxygen reduction causes a loss of fish stock in the river, incurring financial and/or recreational losses to anglers downstream. If the anglers are not compensated for their loss of welfare, the upstream industry will continue its activities as if the damage done downstream was irrelevant to them. They are said to create an *external cost*. An external cost is also known as a *negative externality*, and an *external diseconomy*. If we were considering a situation where one agent generates a positive level of welfare for a third party, we would have an instance of an *external benefit* (*positive externality*, or *external economy*).

An external cost exists when the following *two* conditions prevail:

1. An activity by one agent causes a *loss of welfare* to another agent.
2. The loss of welfare is *uncompensated*.

Note that *both* conditions are essential for an external cost to exist. For example, if the loss of welfare is accompanied by compensation by the agent causing the externality, the effect is said to be *internalised*. This distinction will be made clearer shortly.

4.2 OPTIMAL EXTERNALITY

The first fundamental feature of the different definitions of externality has already been noted: the physical presence of pollution does not mean that 'economic' pollution exists. The next observation is equally important, but much less easy to understand – *even if 'economic' pollution exists it is unlikely to be the case that it should be eliminated.* This proposition can be demonstrated using Figure 4.1.

In Figure 4.1, the level of the polluter's activity, Q, is shown on the horizontal axis. Costs and benefits in money terms are shown on the vertical axis. MNPB is 'marginal net private benefits'. A formal derivation of MNPB, in the context where the polluter is a firm, is given in Appendix 4.1. But an intuitive explanation is also possible. The polluter will incur costs in undertaking the activity that happens to give rise to the pollution, and will receive benefits in the form of revenue. The difference between revenue and cost is *private net benefit*. MNPB is then the marginal version of this net benefit, i.e. the extra net benefit from changing the level of activity by one unit. MEC is the 'marginal external cost', i.e. the value of the extra damage done by pollution arising from the activity measured by Q. It is shown here as rising with output Q. We consider other possible shapes for MEC in Appendix 5.2.

We are now in a position to identify the *optimal level of externality*. It is where the two curves intersect, i.e. where MNPB = MEC. Why is this? We first offer an intuitive explanation. Since the two curves are marginal curves, the areas under them are 'total' magnitudes. The area under MNPB is the polluter's total net private benefit, and the area under MEC is total external cost. On the assumption that the polluter and sufferer are equally deserving – i.e. we do not wish to weight the gains or losses of one party more than another's – *the aim of society could be stated as one of maximising the sum of benefits minus the sum of costs.* If so, we can see that triangle OXY is *the largest area of net benefit obtainable*. Hence, Q^*

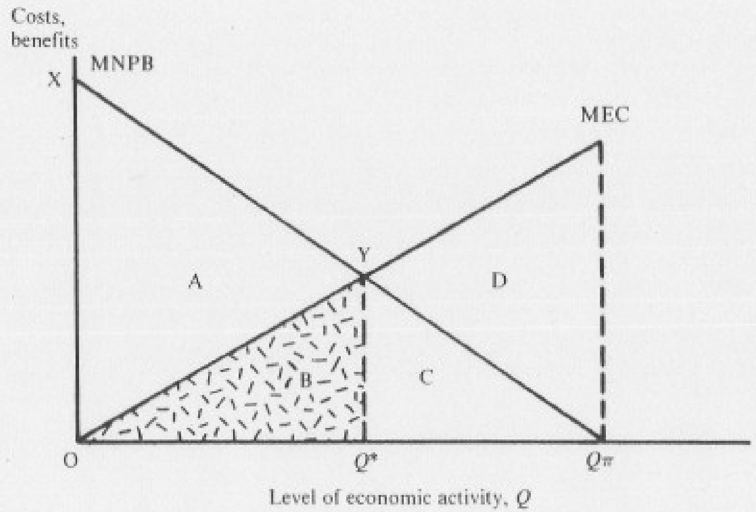

Figure 4.1 Economic definition of optimal pollution.

is the optimal level of activity. It follows that the level of physical pollution corresponding to this level of activity is the optimal level of pollution. Finally, the optimal amount of economic damage corresponding to the optimal level of pollution Q^* is area OYQ^* – area B in Figure 4.1. Area OYQ^* is known as *the optimal level of externality*.

This result can also be derived formally. At Q^*

$$\text{MNPB} = \text{MEC} \tag{4.1}$$

but (from Appendix 4.1)

$$\text{MNPB} = \text{P} - \text{MC} \tag{4.2}$$

where MC is the marginal cost of producing the polluting product. Hence

$$\text{P} - \text{MC} = \text{MEC} \tag{4.3}$$

or

$$\text{P} = \text{MC} + \text{MEC} \tag{4.4}$$

Now, MC + MEC is the sum of the marginal costs of the activity generating the externality. It is *marginal social cost* (MSC). Hence, when

$$\text{MNPB} = \text{MEC}, \; P = \text{MSC} \tag{4.5}$$

'Price equals marginal social cost' is the condition for *Pareto optimality*. We do not demonstrate this here – any undergraduate microeconomics or welfare economics text should contain a proof.

4.3 ALTERNATIVE DEFINITIONS OF POLLUTION

Popular literature on pollution, and sometimes the scientific literature too, speaks of 'eliminating' pollution. The above discussion explains why the typical economic prescription does not embrace this idea. In Figure 4.1 the elimination of pollution can only be achieved by not producing the polluting good at all. But, the laws of thermodynamics imply that there can be no such thing as a non-polluting product. Hence to achieve zero pollution we would have to have zero economic activity. Calls for 'no pollution' thus appear illogical.

The situation is not quite as extreme as this, however. We need to modify Figure 4.1 in an important respect if we are to try to make compatible the economist's and the scientist's presciptions about desirable levels of pollution. In Chapter 2 we saw that the natural environments which receive waste products can be characterised as having a certain 'assimilative capacity' – they can receive a certain level of waste, degrade it and convert it into harmless or even beneficial products. If the level of waste, W, is less than this assimilative capacity, A, then some externality will still occur as the process of degradation and conversion takes place. But if W exceeds A a further process of degradation will also occur, for A itself will be impaired. Disposing of waste to environments that cannot handle it simply reduces the capacity of that environment to deal with more waste.

To some extent we can capture this idea of assimilative capacity by observing that the MEC curve in Figure 4.1 should really have its origin at some positive level of economic activity Q_A. Below this level, the only kind of externality will be 'temporary' – the environment will eventually return to normal once the waste degradation process has taken place. On the assumption that we can ignore this temporary externality for the moment, the MEC curve appears as in Figure 4.2. (Note that MEC begins at Q_A only if people

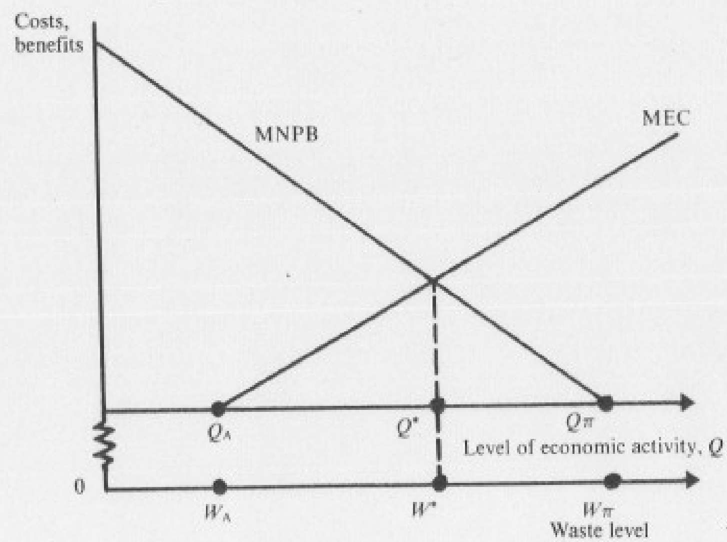

Figure 4.2 Optimal pollution levels with positive assimilative capacity.

notice the physical effects then. Otherwise it can begin even further to the right along the horizontal axis. In the extreme, if people do not care about the physical effects of the waste flows there is no MEC curve at all.)

Figure 4.2 does not alter any of the analysis about the economically optimal level of externality. The findings of the previous section stand. But we can now see that the idea of 'zero pollution' is not, after all, quite so silly as if first appeared. Zero pollution is still non-optimal, as Figure 4.2 shows, but it does not entail zero economic activity. In a static world the difference between the economist's optimum and the scientist's prescription is likely to be significant. As we shall see later in this text, once dynamic considerations are introduced the difference is not so marked, and may not exist at all.

Figure 4.2 also shows how the level of economic activity relates to the level of waste emitted. Assuming waste is directly proportional to the level of activity we can simply translate any amount of Q into some corresponding level of W. Just as Q^* is the optimal level of economic activity, so W^* is the optimal level of waste-producing pollution. Later we shall have occasion to modify this picture: if the polluter adopts pollution abatement equipment, Q can increase without the corresponding W – recall that the First Law of Thermodynamics still dictates that W will be proportional to Q – affecting the environment. Basically, some of the W is 'redirected' so that it does not affect the environment. Once again, we see that the 'zero pollution' prescription has some foundation – zero waste is an impossibility, but zero quantities of waste affecting the environment is less fanciful.

Finally, Figures 4.1 and 4.2 are basic to most of the analyses in the chapters that follow. It will therefore pay the reader to study them carefully. Because the subsequent analysis is generally not affected by the starting point of the MEC curve we will, for notational convenience, tend to use the MEC curve shown in Figure 4.1. When it is necessary to introduce the effects of positive assimilative capacity, we will adopt Figure 4.2.

4.4 TYPES OF EXTERNALITY

We are now in a position to define some further terms. In terms of Figure 4.1,

Area B	= the optimal level of externality
Area A + B	= the optimal level of net *private* benefits for the polluter
Area A	= the optimal level of net *social* benefits
Area C + D	= the level of *non-optimal* externality which needs to be removed by regulation of some sort
Area C	= the level of net private benefits that are socially unwarranted
Q^*	= the optimal level of economic activity
$Q\pi$	= the level of economic activity that generates maximum *private* benefits

Figure 4.1 thus demonstrates a very important proposition: in the *presence of externality there is a divergence between private and social cost.* If that divergence is not corrected the polluter will continue to operate at a point like $Q\pi$ in Figure 4.1. At $Q\pi$, private benefit is maximised at A + B + C, but external cost is B + C + D. So, net social benefit = A + B + C – B – C – D = A – D, which is clearly less than A, the net social benefits when the polluter's activity is regulated to Q^*.

Externality level C + D is said to be *Pareto relevant* because its removal leads to a 'Pareto improvement', i.e. a net gain in social benefits. Externality level B is *Pareto irrelevant* because there is no need to remove it.

4.5 WHO ARE THE POLLUTERS?

We have deliberately refrained from classifying polluters. The typical 'image' is that polluters are firms. But it is also the case that polluters are individual people – car drivers create noise and cause accidents, people who play radios in and out of doors cause noise nuisance, and so on. Indeed, the general combinations are as follows:

Externality Generator	*Externality Sufferer*
Firm	Firm
Firm	Individuals
Individuals	Firm
Individuals	Individuals
Government	Firm
Government	Individuals

The inclusion of government as a creator of externality acknowledges that governments often generate external effects through poor legislation and rules.

4.6 CONCLUSIONS

1. Scientists tend to define pollution differently to economists.
2. For the economist, pollution is an *external cost* and occurs only when one or more individuals suffer a *loss of welfare.*
3. Even then, economists do not typically recommend the *elimination* of externality because they argue that the *optimal externality* is not zero.
4. The idea of 'zero pollution' is not, however, absurd. At least two considerations make it more reasonable than it appears at first sight. These are (a) the fact that the environment tends to have a positive assimilative capacity, and (b) the fact that it is possible, to some extent, to divorce economic activity from waste flows

affecting the environment by introducing pollution abatement.

5. It is wrong to think of 'polluters' only as firms: individuals pollute. So do governments.
6. Caveat – the analysis in this chapter has assumed *perfect competition*. As we shall see, some of the conclusions do not hold when we relax this assumption.

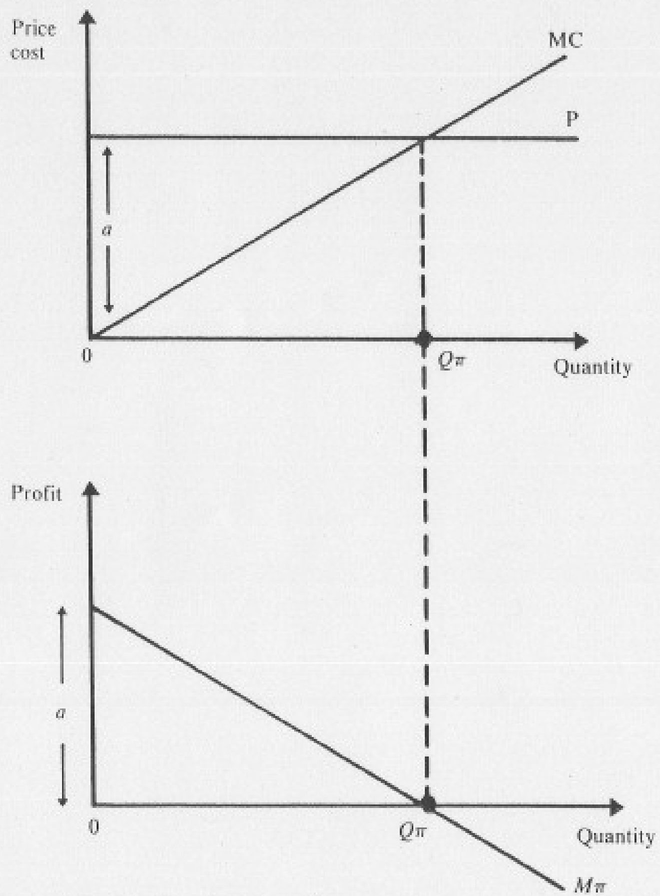

Figure A4.1 Deriving the MNPB curve

APPENDIX 4.1: DERIVING A MARGINAL NET PRIVATE BENEFIT CURVE

Chapter 4 introduced MNPB in a general way. To give it more formal meaning we can look at how it is derived in the context of the theory of the firm. Figure A4.1 shows a demand and marginal cost curve for a *perfectly competitive firm*. (The type of competition is important – we shall see later in the text that the definition of MNPB given here does *not* hold for imperfectly competitive conditions.) By subtracting marginal cost (MC) from price (P), we derive a *marginal profit* curve ($M\pi$). $M\pi$ shows the extra profit made by expanding output by one unit. Clearly, total profits, the area under $M\pi$, are maximised when $M\pi = 0$. Profit is equivalent to the *net benefit* obtained by the firm. Hence, marginal profit is formally equivalent to marginal net private benefits.

5 · THE MARKET ACHIEVEMENT OF OPTIMAL POLLUTION

5.1 PROPERTY RIGHTS

Chapter 4 demonstrated that a socially optimal level of economic activity does not coincide with the private optimum if there are external costs present. The issue arises therefore of how to reach the social optimum. Some form of intervention by government would seem to be necessary. Before looking at the various forms of regulation that might be applied, it is important to probe a little further to be sure that markets will not 'naturally' achieve the optimal level of externality.

It is the contention of one school of thought that even if markets may not secure the optimum amount of externality, they can be very gently 'nudged' in that direction without the necessity for full-scale regulatory activity involving taxes or standard-setting. This basic idea was first propounded in a paper by Ronald Coase (1960). To understand the argument we have first to establish the concept of 'property rights'.

Despite the apparent meaning of the phrase, a property right relates to the right to *use* a resource. This might mean the right to cultivate crops on land that is owned, the right to use one's own house, and the right to use the natural environment in a particular way. Such rights are rarely, if ever, absolute: they are circumscribed in some way by the generally accepted rules of society. The right to cultivate land does not usually carry with it the right to grow opium poppies or even giant hogweed (which is capable of causing quite severe skin irritation). The rights are said to be 'attenuated'. Note that 'property' has a much wider meaning than in everyday language – it can refer to any good or resource. Similarly, the environment is a resource and hence 'property'.

Rights can be *private*, i.e. owned by readily identifiable individuals, or *communal* where the use of the property in question is shared with others. The latter kind of property is known as *common property*. Before the enclosures of land in England, grazing land was often common property: many individuals could graze their animals on the land. In a great many developing countries, land is owned communally. We consider in Chapters 16 and 17 whether the way in which property rights are held helps to explain the process of natural resource degradation, but for the moment we are interested in the general concept of property rights.

5.2 THE POTENTIAL FOR MARKET BARGAINS IN EXTERNALITY

Figure 5.1 repeats the basic optimal externality diagram in Chapter 4. Recall that, left unregulated, the polluter will try to operate at $Q\pi$ where his profits are maximised. But the social optimum is at Q^*. The workings of the market and the goal of a social optimum appear to be incompatible.

Now consider a situation in which *the sufferer has the property rights*. What this means is that the sufferer has the right *not* to be polluted and the polluter does not have the right to pollute. In that case the starting point is surely the origin in Figure 5.1. The sufferer will prefer that no pollution at all takes place and, since he has the property rights, his view will hold the biggest sway. But now consider whether the two parties – polluter and sufferer – might 'bargain' over the level of externality. Suppose the issue is whether to move to point d or not. If they moved to d, the polluter would gain Oabd in total profit, but the sufferer would lose Ocd. But since Oabd is greater than Ocd, there is potential for a bargain. Very simply, the polluter could offer to *compensate* the sufferer by some amount *greater* than Ocd, and less than Oabd. The polluter will still have a net profit. Moreover, the sufferer would be better off: although he would lose Ocd, he would gain more than that in compensation. If such a bargain could be struck, the move to d would be seen to be an improvement for both parties (such a move is known as a 'Pareto improvement' since at least one party is better off and no party is any worse off). But if the move from O to d is a social improvement so is the move to e (simply repeat the argument). Indeed, so is a further

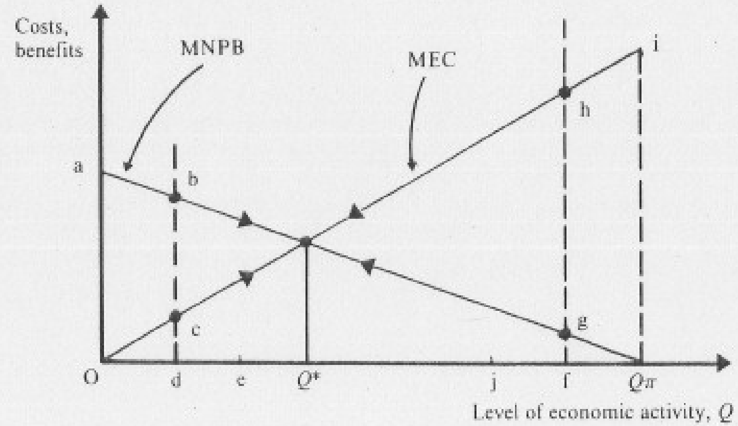

Figure 5.1 Optimal pollution by bargaining.

move to Q^*. But any move to the right of Q^* is not feasible because the polluter's net gains then become less than the sufferer's losses – hence the polluter cannot compensate the sufferer to move beyond Q^*. Thus, if we start at O and the property rights belong with the sufferer, there is a 'natural' tendency to move to Q^*, the social optimum.

Now imagine that the property rights are vested in the polluter. The starting point is $Q\pi$ because that is the point to which the polluter will go given that he has every right to use the environment for his waste products. But it is now possible for the two parties to come together again and consider the move from $Q\pi$ *back* to f. But this time the sufferer can compensate the polluter to give up a certain amount of activity. Since the sufferer would have to tolerate a loss of fhi$Q\pi$ if the move to f does not take place, he will be willing to offer any amount less than this to make the move. The polluter will be willing to accept any amount greater than fg$Q\pi$, the profits he will have to surrender. The potential for a bargain exists again and the move to f will take place. But if the move to f is a social improvement, so is the move from f to j and from j to Q^*. Hence Q^* is once again the level of activity to which the system will gravitate.

So long as we can establish a bargain between polluter and polluted, the market will, on the above argument, take us to Q^* which is the social optimum. The potential importance of the argument can now be seen, for *regardless of who holds the property*

rights, there is an automatic tendency to approach the social optimum. This finding is known as the 'Coase theorem', after Coase (1960). *If it is correct, we have no need for government regulation of externality, for the market will take care of itself.*

5.3 CRITICISMS OF THE COASE THEOREM

Clearly the theorem is of considerable potential importance since it removes the necessity of government regulation of pollution problems (and also threatens to render the next few chapters redundant!) But, despite its elegance, there are many problems with the Coase theorem. We consider the main criticisms only.

The state of competition

Chapter 4 was careful to point out that the analysis of optimal externality assumed perfect competition. It was on this basis that we saw that

MNPB = P – MC

and, hence,

(MNPB = MEC) entails (P = MSC)

In terms of the bargaining approach, what is being assumed is that MNPB is the polluter's *bargaining curve*. It is this to which he refers when deciding how much to pay, or how much to accept, in compensation. But suppose that perfect competition does not prevail. Then P – MC is no longer the bargaining curve because it will not be equal to MNPB. If the polluter is a firm, it should be fairly evident that his bargaining curve is his marginal profit curve (see Appendix 4.1) and, under imperfect competition, this is equal to *marginal revenue* minus marginal cost, i.e.

MNPB = MR – MC

Under imperfect competition, MR is not equal to P because the demand curve is above the marginal revenue curve. It follows that the bargaining solution does not apply under imperfect competition.

How serious this is as a criticism depends on two things. First it depends on how different we think the real world is from perfect competition. While some economists would argue that the amount of competitive 'imperfection' (or monopoly) is not very great, our view is that perfect competition is a convenient fiction for constructing economic models, but it is remote from describing the real world. Thus, the existence of imperfect competition provides the basis for a serious criticism of the Coase theorem. The second point is more complicated and is dealt with more formally in Appendix 5.1. The possibility exists that the bargaining curve of the polluter can be defined as one relating jointly to the interests of polluters and consumers. They need then to bargain with the sufferers of the pollution. While the approach is technically correct, it requires a rather fanciful involvement of producers (polluters), consumers and sufferers all in one bargain. It does not therefore seem at all realistic.

The absence of bargains and the existence of transaction

The second criticism of the Coase theorem is that we are probably all rather hard-pressed to think of real-world examples of such bargains taking place. It is true that some electricity-generating authorities 'bargain' with the local population to accept nuclear power stations or waste disposal facilities, perhaps offering cash compensation or a contribution to local facilities. There are also examples of international bargains between countries that suffer pollution and countries that create it, but they typically involve common property resources, and we deal with that issue later. But Chapter 2 indicated that externality is likely to be pervasive because of the materials balance principle. We should therefore be able to point to many such bargains rather than to isolated examples. The fact that we do not observe many examples of the bargains taking place suggests that there are either obstacles to them, or that the Coase theorem is not rooted in real-world economics.

The response of those who believe in the market bargain approach is that there are indeed obstacles to bargaining in the form of *transactions costs*. Such costs include those of bringing the parties together, organising often widely distributed and difficult-to-identify sufferers, the actual bargain itself and so on. If the transactions costs are so large that any *one* party's share of them outweighs the expected benefits of the bargain, that party will withdraw from the bargain, or not even commence it. Moreover, it seems likely that transactions costs will fall on the party that does not have the

property rights. But transactions costs are real costs – we have no reason for treating them differently to other costs in the economy. Thus, if transactions costs are very high all we appear to be saying is that the costs of the bargain outweigh any benefits. In that case it is *optimal* that no bargain occurs.

Carried to this level the argument quickly becomes redundant, for what it says is that bargains will either take place or they will not. If they do, then the amount of externality emerging will be optimal (by the Coase theorem). If they do not take place, it is also optimal for it simply means that transactions costs exceed expected net benefits from the bargain. We have an unfalsifiable theory about optimal externality. It says that all the externality we observe is optimal externality and hence there is no need to do anything about it. But the proof involves non-falsifiable statements and hence the argument is non-falsifiable.

Nonetheless, the transactions costs argument serves to remind us of some important caveats in any recommendation about regulation of externality:

1. Simply because we observe externality it does not mean that something should be done on grounds of economic efficiency – we might be observing Pareto-irrelevant externality (Chapter 4). This kind of mistake is in fact very common, as with statements to the effect that 'all' pollution should be eliminated, or tobacco smoking should be prohibited and so on.
2. The existence of high transactions costs might explain why government intervention occurs. For high transactions costs do not entail that the externality is optimal at all – instead it may simply be that government intervention is cheaper and can achieve optimality.

Letting T = transactions costs, B = the gain from the bargain for the party bearing the transactions costs, and G = the cost of government intervention, we might summarise the possibilities as follows:

- If $T < B$, a bargain might take place (see below for reasons why they might not occur in this context).
- If $T > B$, a bargain will not occur, but some other regulatory approach might occur.
- If $T > G < B$, government regulation is likely to occur, and it will be efficient.

Finally, note that while transactions costs may leave some of the bargaining theory intact, their existence means that the optimal level of activity is no longer invariant with the allocation of property rights. It will matter who bears the transactions costs.

Identifying the bargaining parties

Even if transactions costs are less than the benefits to be obtained from a bargain, no bargain may take place. Many pollutants are long-lived – they stay in the environment for long periods of time and may affect people years, decades or even hundreds of years from now. If so, the people who are going to be affected by the pollution may not yet exist, and it is then not possible to speak of the two parties coming together to bargain. Toxic chemicals, radioactive waste, ozone layer depletion and global carbon dioxide pollution all fit this category, among many others. At best, some groups in the present generation would have to bargain *on behalf* of future generations. The idea of future generations having such representatives is of course not fanciful – many regulations reflect that kind of interest – and typically we expect governments to take on this role. But the contexts involved are usually common property ones and the outcome is usually some attenuation of the rights of polluters.

A further problem of identifying the polluters and the sufferers arises in cases of *open access resources*. An open access resource is one owned by nobody (common property resources are owned by an identifiable *group*). In such cases it is not clear who would bargain with whom since no one individual has an incentive to reduce his or her access to the resource.

Lastly, even in conventional pollution contexts it is often difficult to say who the polluters and sufferers are. Sufferers may be unaware of the source of pollution from which they suffer, or even unaware that damage is being done. This is often the case for air pollutants and water pollutants. Indeed, this situation seems likely to characterise the majority of pollution situations. The costs of generating the information for the sufferers needs to be added to the costs of transacting any bargain. The likelihood of bargains being socially efficient even if they occurred is also remote given the need to identify damage done and its distribution among sufferers. Of course, this kind of problem will arise for regulatory solutions as well. Governments have to find information on damage.

Common property contexts

We noted earlier that property rights can be private or communal. In the communal case a kind of mutual bargain among users of the property can occur. Each user agrees to restrict his usage of the resource in the interest of its longer-term sustainable use for the community as a whole, and for later generations. This is called a *cooperative* solution to a problem of *assurance*. Each individual needs assurance that others will also behave in a cooperative fashion, otherwise there will be a temptation to 'break ranks' and seek the maximum private gain. Despite a voluminous theoretical and empirical literature on such 'game theoretic' situations, it is not easy to say why some common property contexts are subject to cooperative solutions and others break down. But from the bargaining theory point of view the important point to note is that each user of the common property is the polluter (or resource user) *and* each individual user is also the beneficiary. In terms of the previous diagrams, MNPB and MEC 'belong' to the same people. Rational cooperative individuals will therefore net out the costs and benefits to arrive at their own personal Q^* so that the sum of the individual positions will be the social optimum. Nonetheless it can pay an individual to move beyond Q^* if he or she judges they can 'get away with it' and make fairly large short-term gains at the expense of the other users now and in the future.

Threat-making

One other problem with the bargaining solution is that it offers potential for making an economic activity out of threat-making. If a sufferer compensates a polluter because the polluter has the property rights, it is open to other 'polluters' to enter the situation and to demand compensation. Threat-making is hardly a rational use of scarce economic resources. Possibly the situation can be corrected by carefully defining who is entitled to property rights, e.g. by denying them to potential threat-makers, but it has to be acknowledged that compensation schemes for potential polluters have suffered this difficulty. In some countries it is possible to receive government cash for *not* engaging in cultivation, the idea being to protect environmentally valuable land and reduce agricultural surpluses. It seems likely that some farmers could say that they are going to farm

8 · MARKETABLE POLLUTION PERMITS

8.1 THEORY OF MARKETABLE PERMITS

The idea of pollution permits was introduced by J.H. Dales (1968). As with standard-setting, the regulating authority allows only a certain level of pollutant emissions, and issues permits (also known as pollution 'consents' or certificates) for this amount. However, whereas standard-setting ends there, the pollution permits are tradeable – they can be bought and sold on a permit market.

Figure 8.1 illustrates the basic elements of marketable permits. MAC is the marginal abatement cost curve which, as Chapter 6 showed, can also be construed as the MNPB function if the only way of abating pollution is to reduce output. The horizontal axis shows the level of emissions and the number of permits: the easiest assumption to make is that one permit is needed for each unit of emission of pollution. The optimal number of permits is OQ^* and their optimal price is OP^*. That is, the authorities, if they seek a Pareto optimum, should issue OQ^* permits. S^* shows the supply curve of the permits: their issue is regulated and is assumed not to be responsive to price.

The MAC curve is in fact the demand curve for permits. At permit price P_1, for example, the polluter will buy OQ_1 permits. He does this because, in terms of control strategies, it is cheaper to abate pollution from Q_2 back to Q_1 than to buy permits. To the left of Q_1, however, it is cheaper to buy permits than to abate pollution. MAC is thus the demand curve for permits.

8.2 THE ADVANTAGES OF MARKETABLE PERMITS

Why do the permits have to be marketable? There are six main attractions of marketability.

1. Cost minimisation

Figure 8.2 repeats Figure 8.1, but omits the MEC curve. It also shows the overall MAC curve as being the sum of the individual polluter's MAC curves. We assume just two polluters for simplicity. This aggregation is legitimate because it was shown above that the MAC curve is the demand curve for permits: adding the curves up is therefore the same as aggregating any set of demand curves. By reference to the individual MAC curves of the two polluters we can see how many permits are purchased. Polluter 1 buys OQ_1 permits, and polluter 2 buys OQ_2 permits at price P^*. Note that the higher

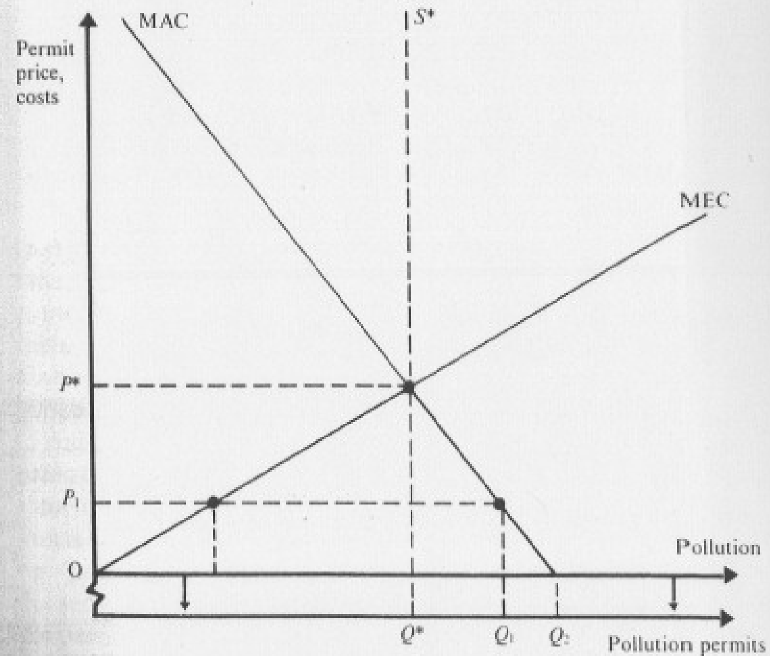

Figure 8.1 The basic analytics of marketable permits.

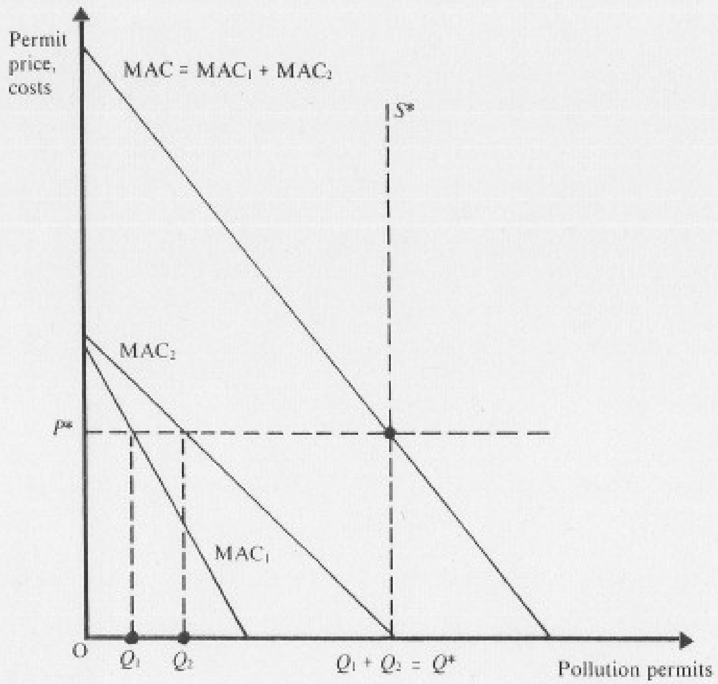

Figure 8.2 Cost minimisation with marketable permits.

cost polluter (2) buys more permits. This gives us a clue to the cost-effectiveness of permits. Polluters with low costs of abatement will find it relatively easier to abate pollution rather than buy permits. Polluters with higher costs of abatement will have a greater preference for buying permits than for abating pollution. Since polluters have different costs of abatement there is an automatic market – low-cost polluters selling permits and high-cost polluters buying them. By giving the polluters a chance to trade, the total cost of pollution abatement is minimised compared to the more direct regulatory approach of setting standards. Indeed, what we have is an analogue of the Baumol–Oates theorem about taxes being a minimum-cost way of achieving a standard (see Section 6.7).

2. New entrants

Suppose new polluters enter the industry. The effect will be to shift

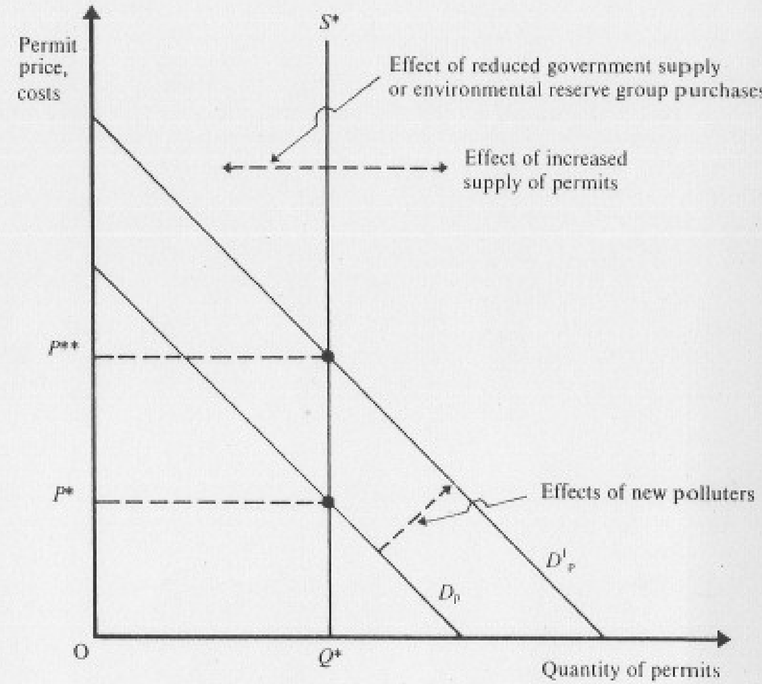

Figure 8.3 Changing the supply and demand for permits.

the aggregate pollution permit demand curve to the right, as in Figure 8.3. As long as the authorities wish to maintain the same level of pollution overall, they will keep supply at S^* and the permit price will rise to P^{**}. The new entrants will buy permits if they are high abatement cost industries, otherwise they will tend to invest in pollution control equipment. Once again, the overall cost minimisation properties of the permit system are maintained. But suppose the authorities felt that the increased demand for permits should result in some relaxation in the level of pollution control. Then they could simply issue some new permits, pushing the supply curve S^* to the right. Alternatively, if they felt that the old standard needed tightening they could enter the market themselves and buy some of the permits up, holding them out of the market. The supply curve would shift to the left. In short, the permit system opens up the possibility of varying standards with comparative ease to reflect the conditions of the day. The authority would simply engage in market operations, rather like a central bank buys and sells securities to influence their price.

3. *Opportunities for non-polluters*
Although it is not regarded as an intended feature of the permit system, there is another intriguing feature of them. If the market in permits is truly free, it will be open to anyone to buy them. An environmental pressure group, concerned to lower the overall level of pollution, could enter the market and buy the permits, holding them out of the market, or even destroying them. Such a solution would be efficient because it would reflect the intensity of preference for pollution control, as revealed by market willingness to pay. The danger with this idea is, of course, that a government might react adversely to a situation in which the level of pollution it had decided was optimal or acceptable was being altered by people who disagreed with it. They might simply issue new permits each time the environmental group bought the permits.

4. *Inflation and adjustment costs*
Permits are attractive because they avoid some of the problems of pollution taxes. As we saw in Chapter 6, even where a standard is set and taxes are used to achieve it, there are risks that the tax will be mis-estimated. With permits it is not necessary to find both the desirable standard and the relevant tax rate; it is necessary only to define the standard and find a mechanism for issuing permits. Moreover, if there is inflation in the economy, the real value of pollution taxes will change, possibly eroding their effectiveness. Because permits respond to supply and demand, inflation is already taken care of. Taxes also require adjustment because of entry to, and exit from, the industry. Permits, as we have seen, adjust readily to such changes, whereas taxes would require adjustment.

5. *The spatial dimension*
We have tended to assume that there are just a few polluters and that the points at which the pollution is received (the 'receptor points') are also few in number. In practice we are likely to have many emission sources and many receptor points. If we are to set taxes with at least a broad relationship to damage done, it will be necessary to vary the taxes by source since different receptor points will have different

assimilative capacities for pollution. Additionally, there are likely to be *synergistic* effects. That is, several pollutants may combine to produce aggregate damages larger than the sum of the damages from single pollutants. This raises the spectre of a highly complex and administratively burdensome system. To a considerable extent permits avoid this spatial problem. To investigate this further we need to look briefly at different types of permit systems.

6. *Technological 'lock-in'*
Permits are also argued to have an advantage over charges systems with respect to 'technological lock-in'. Abatement expenditures tend to be 'lumpy'; to increase the level of effluent removal, for example, it is frequently necessary to invest in an additional type of abatement process. Adjustments to changes in charges are therefore unlikely to be efficient unless the changes in the charge can be announced well in advance and can be backed by some assurance that a given charge level will be fairly stable over the short and medium term. The charge approach also risks underestimating abatement costs. For example, if the aim is to achieve a given standard, then, together with the regulating authority's assessment of abatement costs, this will determine the relevant charge. If the authority is wrong about the abatement costs, however, the charge could be set too low in the sense that polluters will prefer to pay it than to invest in abatement equipment, thus sacrificing the desired standard. This reluctance of polluters to invest in equipment will be strengthened by the previously discussed 'lumpiness' factor. A permit system generally avoids this problem of lumpy investment, the authority's uncertainty about abatement costs, and polluters' distrust of charges. This is so because the permits themselves are issued in *quantities* equal to the required standard, and it is prices that adjust. The consequences of an underestimate of abatement costs in the presence of permits is simply that the price of permits is forced up (since the demand for them is determined by abatement costs, as we saw), whereas the environmental standard is maintained (Rose-Ackerman, 1977).

8.3 TYPES OF PERMIT SYSTEMS

The literature has tended to classify three types of permit system. The *ambient permit system* (APS) works on the basis of permits defined according to exposure at the receptor point. Quality standards might vary according to the receptor point: there is no need for each receptor point to have the same ambient quality standard. Under an APS, then, permits have to be obtained from the market in permits at the receptor point. This means that the trade in permits will not be on a one-for-one basis; it will be necessary to trade on the basis of the number of permits required to allow a given amount of pollution concentration at the receptor point. Each polluter, then, may face quite complex markets – different permit markets according to different receptor points, and hence different prices.

The *emissions permit system* (EPS) is much simpler. It simply issues permits on the basis of source emissions and ignores what effects those emissions have on the receptor points. Within a given region or zone, then, the polluter would have only one market to deal with and one price, the price of a permit to emit pollutants in that area. Trade in permits is on a one-for-one basis.

The APS has obvious complications for the polluters and may well be an administrative nightmare for the regulators as well. The EPS is simpler but has other problems. By not discriminating according to receptor points it is unlikely to discriminate between sources on the basis of the damage done. It will therefore be inefficient. Put more formally, the price of permits will not approximate the marginal external cost. Second, any one area is likely to experience some concentration of pollution in specific small areas – so-called 'hot spots' – where actual concentrations exceed the standard. Because the EPS is emission-based across a wider area, it will not take account of this failure to observe the standard at all points. The simple technique of re-defining the area so that the hot spot is contained within a narrower zone to which the standard applies really amounts to turning the EPS into an APS, and we are back to the complexities of many markets and prices. The EPS also works on the basis of a one-for-one trade within the defined zone – there is no trade outside the zone. With the APS, however, all receptor points are taken into account. EPS could thus result in damage outside the zone being ignored.

To overcome these difficulties a third system has been proposed. This is the *pollution offset* (PO) system. Under the PO system, the permits are defined in terms of emissions, trade takes place within a defined zone, but trade is not on a one-for-one basis. Moreover, the standard has to be met at all receptor points. The exchange value of

the permits is then determined by the effects of the pollutants at the receptor points. The PO system thus combines characteristics of the EPS (permits are defined in terms of emissions, and there is no trade outside the defined area) and the APS (the rate of exchange between permits is defined by the ambient effects).

Which is the best system? Tietenberg (1985) has reviewed much of the evidence. His review suggests that EPS is more expensive than APS in terms of the total abatement costs likely to be involved. But the APS is also judged to be a largely unworkable system because of its complexity. How then does EPS fare in comparison to the more traditional standard-setting, or 'command-and-control' systems? The evidence is varied and is not easy to compare as the two systems might have different amounts of emission control because of difficulties in the spatial configuration of the requirements to meet the standard. The PO system was not evaluated.

8.4 PERMIT TRADING IN PRACTICE

There is some experience of pollution permit trading in the United States. The Clean Air Act (1970) established National Ambient Air Quality Standards (NAAQSs) which were to be implemented by the individual states under State Implementation Plans (SIPs). The Act marked the introduction of federal control, through the Environmental Protection Agency (EPA), over what had previously been a state responsibility alone. The SIP for each state had to indicate to EPA how the state would implement the ambient standards for all pollutants other than 'new sources' which were controlled directly by standard-setting by EPA.

In 1977 the Clean Air Act was amended to allow for the fact that many states were not meeting the ambient standards. Areas not meeting the standards were declared to be *non-attainment* regions. Stringent regulations were applied to these regions. All 'reasonably available control technologies' (RACTs) had to be applied to existing plant, and there had to be 'reasonable further progress' in achieving annual reductions so that the standard could be achieved. New sources were subject to construction permits which were conditional on the use of the 'lowest achievable emission rate' (LAER), the lowest emission rate demonstrated to have been achieved elsewhere. In the area where standards had been met, the focus switched to prevention of significant deterioration (PSD), i.e. to ensuring that the areas did not deteriorate.

The other main change in 1977 was the introduction of an emissions trading programme. Basically this operates through an *emission reduction credit*. Suppose a source controls emissions more than it is required to do under the standard set. Then it can secure a credit for the 'excess' reduction. The credit could then be traded in several ways. The first way is through a policy of *offsets*. These can be used in non-attainment areas, allowing new sources to be established, and which thus add to emissions, provided there is a credit somewhere else in the region. The new source effectively buys the credits from existing sources, the overall pollution level is not increased, and new industry is not unduly deterred from setting up in non-attainment regions that would otherwise suffer a loss of income and employment.

The second way is through a *bubble policy*. A 'bubble' is best thought of as an imaginary glass dome covering several different sources of pollution, either several points within one plant, or several different plants. The aim is not to let the overall emissions from the imaginary bubble exceed the level required by the standard-setting procedure. If any one point exceeds the RACT standard, for example, it can be compensated for by securing emission reduction credits elsewhere within the bubble.

The third procedure utilises *netting*. This is similar to the bubble, but relates to sources undergoing modification and which wish to avoid the rigours of being classified as a new source and subjected to the stricter standard (LAERs). Again, so long as plant-wide emissions do not increase, the modified source can increase emissions if there are emission reduction credits to offset the increase.

Lastly there is *banking* whereby sources can store up emission reduction credits for use later in a netting, bubble or offset context.

These components have a clear affinity with the permit trading systems discussed previously. The actual progress of these legislative features of the US policy is complex and varied. An overall evaluation of the policy is difficult, but several general observations stand out. First, trading has tended to result in better air quality, although there are exceptions. Second, there appear to have been significant cost savings. Third, the offset policy probably has assisted regions which would otherwise have suffered economically because of firms being unable to set up in non-attainment regions. Fourth,

administrative costs have been high. Fifth, it is probable that abatement technology introduction has been stimulated by the policy. By 1986 the total number of bubbles in existence was thought to be about 250; 3,000 offset transactions were reported. The amount of netting appears not to be known and banking has had a very limited impact.