George Yarrow

Regulatory Policy Research Centre
Hertford College, Oxford

NOTES ON ENERGY PRICES AND REGULATORY POLICY

1. Introduction

Energy prices are a much favoured target of public policy, as a glance at the history of the energy sector in Britain quickly shows. Today, not only are electricity and gas subject to price controls in franchised monopoly markets, but a levy is applied to electricity generated from fossil-fuel generating sets and the proceeds are passed over to the operators of nuclear power stations. Unlike many goods and services, electricity and gas supplied to domestic premises are not currently subject to Value Added Tax, although this position is scheduled to change in 1994.

The furore over the forthcoming imposition of VAT on household electricity bills is an indication of some of the political pressures on energy prices. The proximate cause of the extension of the tax base was a deteriorating position in Britain's public finances (i.e. a widening gap between overall public expenditure and tax revenues). Secondary arguments in favour of the extension concerned incentives to reduce energy consumption. On the other side of the coin, consumers adversely affected by the price hike have opposed the move, and particular opposition has been generated by the distributional effects of the expected price increases, particularly on the old and the poor.

Two immediate lessons can be learned from the VAT episode in Britain. First, energy pricing is a politically sensitive issue, and the political economy of public policy in this area tends to be a good deal more complex than might be suggested by a textbook on the microeconomics of the energy sector that concentrates on efficiency issues alone. Factors such as the macroeconomic situation and the balance of political forces at a specific time will tend to be highly influential in the conduct of policy.

Second, it is rather more difficult than economists sometimes suggest to negate distributional effects
when making policy changes. Indeed, as a rough rule of thumb, when considering the economics of regulation it is most useful to think first of distribution and only secondly of efficiency. In the VAT extension case, the British Government has argued that those most seriously affected by the proposed price hike will be compensated in one way or another. Thus, in raising the general price level, the increased energy prices will trigger automatic increases in state pensions and in state social security benefits which are indexed to prices. The compensation is, however, necessarily approximate -- pensioners consuming different amounts of energy will receive the same upratings to incomes -- and some adverse redistributions are inevitable.

Before proceeding, let me look at the efficiency/distribution breakdown in a more technical way. Suppose in some general sense it is possible to calculate an efficient level of energy prices and that, somehow or other, the efficient level of prices has been achieved. Consider next the effects of deviations of prices from that level. Provided the change in prices is not too large, the approximate effect on efficiency is zero.

To see this, let efficiency be a function of price, say E(p). Then the change in efficiency consequent on a price change δp is:

$$\delta E = (dE/dp)\delta p + \text{terms in } \delta p^2, \delta p^3, \text{etc.}$$

But at the efficient price level the rate of change of efficiency with respect to price, dE/dp, is zero (this is a necessary condition for maximum efficiency) and, for small δp, the higher order terms in the expansion will be very small (in mathematical terms, of the second order).

The accuracy of the approximation depends not only on the size of the price change, but also on the convexity of the relationship between efficiency and price. The less convex (i.e. the more flat) the function the less sensitive will be efficiency to any given price change. Convexity of the relationship depends in turn on the shapes of underlying social cost and social benefit relationships. For example, convexity is lower when these relationships are relatively price inelastic.

In comparison with efficiency effects, the distributional effects of price changes around the efficient point will be much greater. If price is increased by δp then, to a first-order approximation, unless compensated, consumers will lose by an amount equal to $q.\delta p$, where q is the volume traded, and suppliers will gain by an equivalent amount. Again speaking fairly roughly, the economic value of distributive effects is therefore proportional to dp and the economic value of efficiency effects is proportional to δp^2. If, then, the price change in question was around 10%, it would not be surprising to find that distributive effects were around ten times larger than efficiency effects.

Since regulation is driven by political decisions, political decisions tend to be driven by interest group pressures, and interest group pressures tend to be driven by the impacts of regulation on the groups concerned (and not by some generally held conception of the public good), it is well to keep in mind some of these basic lessons of regulatory economics.

2. Monopoly regulation versus environmental regulation

Having stressed the importance of distributional factors, let me nevertheless first consider some of the efficiency issues that arise in relation to energy pricing. In doing so, it is immediately apparent that there is potentially a major conflict of philosophies between those parts of public policy that are concerned with problems of monopoly and competition and those parts that are concerned with the environmental effects of energy production.

In a nutshell, the traditional concern of monopoly regulation has been that outputs are inefficiently low, whereas the concern from the environmental side is that outputs are inefficiently high. This is illustrated by the ways in which the two problems are handled in standard expositions of theory.

Monopoly. A profit-seeking firm with market power chooses an output where marginal revenue equals marginal cost. Since marginal revenue is below price (the demand curve is downward sloping) it follows that price exceeds marginal cost. The value of an extra unit of output (price) therefore exceeds the cost of producing it, which implies that output is lower than is optimal (or, in the alternative formulation, that price is too high).

Environment. In maximising profit the firm takes no account of environmental damages. Social marginal costs exceed private marginal costs, and in working with the latter the firm tends to produce too high a level of output.

It is important to note at this point, that introductory treatments of environmental costs tend to work within the model of an otherwise competitive market. In this case, the finding that marginal social costs exceed marginal private costs leads directly to the conclusion that output is too low. Similarly, when examining monopoly and deciding whether the outcome is efficient or not, the benchmark for comparison in the price equals marginal cost equilibrium.

What is happening here is that the problems are being considered one at a time, assuming that all other conditions for optimality in resource allocation are satisfied. Put another way, in both cases we are within one policy change of a first-best allocation of resources. However, while this may be a useful way of first learning economic concepts, it is not a sensible way of analyzing regulatory policies. Policy analysis is invariably located in a second-best world. Suppose, then, that we consider both issues together, within in a common framework. What can be concluded in this case?

3. Public policy as vertical control

For anyone familiar with the economic literature on vertical supply arrangements and vertical integration, I would recommend use of this framework when considering the question just posed. The government can be regarded as a supplier of "environmental services" to the firm, a perspective that corresponds to the notion that implicit property rights lie with the state. These environmental services include the disposal of waste gases by emission into the atmosphere, for which services the firm may or may not be required to pay.

Let us suppose for the moment that the firm is required to pay in the form of a "tax" per unit of its output (on the assumption, which can be relaxed in other contexts, of a fixed
relationship between output and use of the environment). The firm then builds this tax into its private unit cost function and charges a markup which depends upon demand conditions.

In this case, if the tax is set so as to reflect the best estimate of the marginal environmental damage caused by an additional unit of output, then the monopoly markup will ensure that price is above the socially efficient price and output is below the socially efficient output. There is, however, no reason to assume that the environmental tax will be set in this way, for now we are dealing with a second-best world where the pricing instrument should be used in a way that takes account of predictable distortions elsewhere.

In fact the environmental charge should be a sufficient instrument for achieving the socially efficient price and output levels on the market. To take the simplest case, let the elasticity of demand be constant, so that the monopolists markup over marginal cost is also constant, say m. Then market price, p, will be:

$$p = m[c + t],$$

where c is the firms private marginal cost of production (assumed constant for simplicity) and t is the environmental charge. Hence, any given market price can be achieved by setting:

$$t = [p/m] - c.$$

Note that in some cases this may be negative, so that output is subsidised. Speaking roughly, this arises when environmental damages are relatively low (i.e. p is not a lot greater than c) and monopoly power is high (i.e. the markup m takes a high value). The result here is fairly intuitive: when the monopoly problem is severe output will be too low, even allowing for social costs, and subsidy is one means of inducing greater production.

Under the given, simplified conditions, one policy instrument -- a tax or subsidy on output -- is sufficient to induce the efficient equilibrium. However, in dealing with the efficiency issue, the use of only one policy instrument means that the government is unable simultaneously to control the distributive effects of intervention.

Suppose for example that the target price is p and that the parameters are such that the required value of t to achieve this price is positive. That is, output (q) is taxed rather than subsidised. Then, in this situation, we find that:

$$\text{supernormal profit} = [p - c - t]q$$

$$= pq - pq/m$$

$$\text{government revenue} = tq$$

$$= pq/m - cq.$$

The government is therefore unable to appropriate all the surplus of revenues over private costs that is generated by the vertical chain [= pq - cq].

If, now, we add a second policy instrument the picture changes. One possibility would be a lump sum tax on the firm equal to the value of monopoly profits. This would be like public ownership in that all profit goes to the state. The other obvious possibility is control of downstream prices, and hence control of the markup m (similar to resale price maintenance or control of maximum prices in a vertical supply agreement). Thus, as can be seen from the above expressions, as the markup m decreases the share of total surplus appropriated by the state increases and the share of the surplus appropriated by the firm decreases.

From this perspective, then, price controls on privately-owned energy utilities are chiefly a distributional instrument aimed at preventing shareholders reaping the benefit of measures designed to raise energy prices on environmental grounds.

4. Regulatory competition

Within an overall policy framework that includes environmental charges or other environmental instruments, there is nothing intrinsically wrong with utility price capping that encourages the firm to increase its output. The reason is simply that such tendencies can be offset by the other instruments.

Consider, for example, price-cap regulation that specifies that the average revenue of the firm does not exceed a certain value in a given year (in the UK this would be determined by an RPI-X+Y formula). On this basis the firm is able to increase its revenue if it is able to increase its output and, if the incremental revenue exceeds the incremental cost the firm will have incentives to expand outputs.

Now this incentive for expansion has been criticised on the grounds that it encourages increased environmental damages. There has therefore been a move toward reducing its force within the context of monopoly regulation. That is, there has been pressure on regulatory bodies set up to deal with issues of monopoly and competition to extend their interests into the field of environmental regulation.

One way in which these pressures have been accommodated is via movements away from capping prices or average revenues and toward capping total revenues. The argument is as follows. If the regulated utility's maximum allowed total revenue is fixed according to an RPI-X+Y formula the firm will no longer be able to increase its revenue by expanding output. Indeed, marginal revenue is always zero, and therefore below marginal cost, so the firm will have incentives actually to reduce output.

A total revenue constraint is an extreme option, but alternatives include controls that are combinations of caps on total and average revenues. For example, total revenue could be constrained to lie below $F + p^*q$, so that:

$$pq < F + p^*q.$$

Under this system the price control operates like a two part tariff and the incremental revenue to the firm is p^*. Arguably, p^* should be set equal to marginal cost for first-best incentives.

It is, however, advisable to proceed with caution along this path of reducing incentives for output expansion via regulatory price-cap formulae. While it is true that the adjustments discussed preserve the cost-reducing incentives of RPI-X regulation, it should be remembered that weakened incentives to expand tend to encourage more monopolistic behaviour in other dimensions. For example, they will weaken competition between electricity and gas utilities in ways that could have negative environmental consequences by weakening the drive to find more efficient ways of using electricity (gas) in order to capture business from suppliers of gas (electricity).

Given that monopoly regulation exists because competition is insufficiently strong, measures that serve to relax competitive pressures should be treated with some scepticism. Returning to the wider policy context shows that competition simplifies, rather than complicates, environmental regulation.

5. Prices versus standards

The above discussion has been in terms of a framework in which the environmental policy instrument was a charge or tax, and this is an approach much favoured by economists. In practice, however, policies based upon imposed standards are much more common. Examples of the standards approach are regulations requiring all firms in an industry to cut back emissions by a fixed percentage or to reach a given common level of emissions. In contrast, the charges approach might employ a per unit tax on some measure of emissions from a plant.

If regulators had full information about costs of abatement and levels of damages, implementation of optimal environmental regulation would be a relatively trivial exercise. Moreover, it would be of little consequence which policy instrument -- taxes or standards -- was chosen to achieve the desired target levels of environmental quality. As with other types of government regulation, however, the real economic problems start to emerge only when it is recognised that such information will not easily be available.

Economists have tended to urge greater reliance upon the charges or pollution tax approach on the basis that it has advantages which include:

i. achievement of given environmental quality at least cost because the charge induces higher levels of abatement where control costs are lower; and

ii. the charge provides incentives to seek more efficient control technologies so as to obtain continuing reductions in pollution charges.

In practice, however, the charges/tax approach often has significant weaknesses. Among the most important is that, given the regulator's lack of detailed knowledge about abatement costs, the level of emissions that can be expected to occur at a given tax rate is highly uncertain. Consequently there will be uncertainty about the likely level of damages and, particularly if the relationship between emissions and incremental damages is highly non-linear, so that damages accumulate rapidly past certain levels of emission, this uncertainty can be very costly.

Indeed, much environmental regulation can be described as precautionary in that it is designed to reduce the probability of certain highly damaging events. And, because it controls quantities rather than prices, the imposition of standards frequently achieves this "insurance" objective more effectively than a tax-based approach.

One intermediate solution that has the desirable insurance effects of standards and the incentive effects of charges/taxes is the marketable permits approach. This involves the creation of permits to use environmental inputs -- such as the right to discharge a certain amount of a waste gas into the atmosphere -- which are then either auctioned off or otherwise allocated to firms. The maximum total level of pollution is determined by the number of permits issued, but incentives for cost efficiency and technical progress are retained because firms can buy and sell the rights in an open market. If, as a result of new scientific information, the government decides at a later date that environmental controls should be tightened, it will be able to enter the market and buy back permits, thus reducing the total number in circulation (a procedure that would be akin to open market operations by a central bank in periods of tightening monetary policy).

For the marketable permits approach to work well, it is:

i. necessary to have a reasonably clear unit of measurement for the polluting activity, so that rights can be defined precisely; and

ii. important that there be a stable relationship between damages and the unit of pollutant.

In many cases these conditions are not satisfied. For example, the damages caused by particular emissions may be geographically specific and contingent upon a wide range of other circumstances.

However, where the environmental problem is of larger, perhaps global, scale -- so that damages are less localised and specific -- the marketable permits approach has much to recommend it. It combines effective insurance properties -- frequently the key regulatory objective -- with market incentives for efficiency and freedom to trade in "property" rights. It also fits quite well with the quota rationing schemes that tend to be the outcome of international negotiations. Thus, initial national quotas could be divided into marketable permits which could be traded both domestically and internationally. The governments of wealthier countries that are more willing to use resources to improve environmental quality could purchase permits in the open market in order to take them out of circulation, thereby simultaneously reducing global pollution, increasing the wealth of other nations, and, by increasing the price of permits and hence the opportunity cost of polluting activities, strengthening incentives for further technical progress in control processes.

6. Jurisdictional issues

In one way or another, policy responses to environmental problems involve a change either in property rights themselves or in the management of existing rights. Where pollution is contained within national boundaries, appropriate actions can be taken by the relevant national government; where the problem is more localised still it may be dealt with at a local government level, because this is all that is needed to internalise the economic externality. That is, if both polluters and those adversely affected by the pollution are within the same jurisdiction, the controlling authority will be responsible to both sets of parties and should, in theory, take account of both sets of interests in its decisions. (In practice, of course, the influence of the different interests may well differ radically).

A number of environmental problems, however, involve external effects that cross national boundaries, and these include some of the most pressing issues of the day. For example, the polluters may be located in country A whereas most damage may be caused to country B. Viewing matters in terms of national economic welfare, the government of country A will have incentives to enforce less strict environmental standards than would be optimal from the viewpoint of maximising the welfare of the combined populations of the two nations. By the same token, the government of country B has incentives to push for more stringent controls than would be globally optimal, since the costs of tighter controls would be borne by the inhabitants of country A. Thus, the environmental regulation problem is further exacerbated and is likely to lead to international conflicts.

International externalities are just another example of distributional issues in environmental regulation. While cost-benefit analysis proceeds on the basis that the sum totals of monetary costs and benefits should be compared with each other, in practice any change in regulations will, as already stressed, lead to a shift in the distribution of income. Gainers and losers from any change will often be different groups, and opposition to the change can be expected from the latter irrespective of whether or not there is a net gain in economic welfare as a whole. Even where the problem is relatively localised, therefore, conflicts are likely to arise. (Consider, for example, the conflicts that have occurred in response to decisions concerning the expansion of London's airports or the routing of rail links to the channel tunnel.)

The distributional consequences are particularly acute at the international level as a consequence of the wide variations in living standards across the globe. The value of increased environmental quality -- and hence the monetary damage caused by any degradation in quality -- tends to be an increasing function of income. Hence, other things being equal, a richer country will prefer stricter environmental controls than will a poorer country. It is therefore possible that stricter controls that improve resource allocation at the global level would actually reduce economic welfare in less developed nations. Further, in practice the equity problem is deepened by the fact that most of the past global pollution is attributable to the higher income economies.

It is inevitable, therefore, that future progress on global issues will require compensatory payments to less developed nations (otherwise those countries will have few incentive to contribute to the alleviation of international external effects). Radical proposals will be necessary to avoid endless and ineffective negotiations. The creation of marketable permits may, for at least some problems, be one of the solutions that emerges over the longer term. These allow control of the global level of pollution whilst encouraging efficiency in meeting that target. Moreover, distributional issues can be reduced via the initial allocation of permits.

The international dimension does, however, involve more than just distributional effects. Once broad policy targets have been set at the global level, it is generally the responsibility of national governments to implement appropriate policies. This raises problems of enforcement, and there is no guarantee that the targets will be met. Again, marketable permits may offer a partial solution in that allocations could be withdrawn from any country found to be in violation. The deviant could therefore be punished by the loss of property rights with a positive economic value on world markets.

7. The example of global warming

The greenhouse effect provides a good illustration of some the points discussed in the earlier sections. Current concerns centre on the contribution of CO_2 emissions to global warming. This is a production externality to the extent that the cost of the assimilating capacity of the environment to the producer is less than the cost to society as a whole.

The precise relationships between emissions and damages are, however, highly complex, and involve a number of uncertainties. First, there is uncertainty over how much emitted carbon dioxide stays in the atmosphere and how much is absorbed by the oceans. Second, there are other greenhouse gases that can affect temperatures: CFCs account for a much smaller proportion of greenhouse gas volume but have substantially stronger effects than CO_2. Third, existing atmospheric models do not yet encompass all possible feedback effects (i.e. responses of other variables). Thus, there can be effects on cloud formation, surface ice volumes, ocean circulation, etc. which could mitigate or exacerbate direct greenhouse effects. Fourth, there are significant and uncertain time lags between any increases in gas concentrations and the full emergence of the warming effect.

It follows that, even ignoring the difficulties of placing monetary values on the consequences, measurement of the damage function is beset with problems. The range of estimates of the effects of, say, a doubling of CO_2 concentrations on average temperatures is quite wide, and the consequences of global warming on the scales indicated as possible is highly uncertain. At the aggregate level, it is possible to arrive at economic valuations of the effects that range from very <u>positive</u> numbers (i.e. global warming would have net <u>beneficial</u> effects) to very large negative numbers.

It is possible that, in these circumstances, policy is likely to be strongly influenced by "worst case" evaluations: in particular, environmental policy may be driven by a concern to provide some <u>insurance</u> against highly unfavourable contingencies. Thus, central estimates of the likely monetary effects of CO_2 emissions may not be the most important measures in the analysis. And, given lack of knowledge about the precise structure and parameters of the major ecological subsystems (and hence about their stability properties), "worst case" estimates almost inevitably imply substantial damages.

It is also highly relevant that any global warming that occurs can be expected to have substantial distributional effects. Thus, two of the most sensitive issues concern effects on sea levels and agriculture, which will have differing impacts on different countries. For example, any northward movement of the grain belt would have a significant impact on global economic and political balances. The regional information necessary to predict winners and losers is still extremely thin, but potentially adverse effects on some of the major industrialised countries, and on the

US in particular, will strengthen environmental concerns arising from the "income effect" (i.e. the valuation of environmental quality rises with income) noted earlier.

It is to be expected therefore that policymakers will continue to be under increasing pressure to respond to scientific concerns about greenhouse gases. The degree of pressure will probably be most sensitive to new scientific findings relevant to "worst case" scenarios, and the responses are unlikely to be based on the finely balanced calculations of costs and benefits. Thus, the instruments initially chosen by governments to convince electorates that they are "doing something" to provide insurance can be expected to be relatively crude, particularly given the problems of achieving international agreements (more refined control systems, such as marketable permits to discharge CO_2, may come later). For example, the Toronto Conference on "The Changing Atmosphere" (June 1988) put forward recommendations that included:

i. A 20% reduction in carbon dioxide emissions in industrialised countries by 2005, half of which should come from increased energy efficiency and half from modified energy supply.

ii. National levies on fossil-fuel consumption to finance a world atmosphere fund.

It is clear, however, that such measures would be ineffective over the longer term without controls that applied more generally throughout the world. Thus, for example, greater energy conservation coupled, say, with a move towards greater reliance on non-fossil fuel energy sources in the industrialised world could have perverse effects on fossil-fuel combustion elsewhere by putting downward pressure on world prices.