جزوه زبان تخصصی رشته **اقتصاد انرژی**

(کارشناسی ارشد)

پیام نور

۱۳۹۵

The Economics OF Energy

Volume I & II

Edited by Paul Stevens

An Elgar REFRENCE collection

CHETENHAM, UK. Northampton MA, USA

Introduction

Paul Stevens

Economics is about posing questions. In the 1960s, Joan Robinson used to tell her Cambridge undergraduates, 'The function of economists is to make sure the right questions are being asked ... any damn fool can provide the answers.' The implicit assumption in her typically assertive statement was that it would be the politicians who would provide the answers. Energy economics is no exception to this adage. The subject is full of key questions, some helpful and some quite misleading. Furthermore, the general importance of energy, whether at strategic, environmental or other levels, tends to aggravate the number of 'damn fools' involved in seeking answers. However, before further elaboration, it is worth remembering the comment of another great economist, Edith Penrose, who used to remark that if you were going to ask questions you should at least have the courtesy to explain why they are of interest. Hence it is incumbent upon this author to explain why the subject matters before explaining the subject.

1 Why Energy Matters

Energy economics has a long history as a branch of economics but its position in the pecking order has been driven very much by crises, real or perceived.[1] It was concern over coal supplies and British economic power which prompted Jevons to write about British coal in the last century (Jevons, 1866). It was the first oil shock of 1973 which pushed energy back up the agenda of economists. More recently it has been environmental concerns which have boosted interest in the subject. Since the mid 1980s there has been a huge expansion in the literature on the economics of energy and the environment. The problem facing the editor of this collection on energy economics is that to include a few articles on the issue would not do justice to the subject, while to include a comprehensive coverage would swamp the work. Much of the literature is recent, and it is too early to determine what may or may not be eventually influential. The problem is solved by leaving out the economics of energy and the environment. Those interested are directed to Common, 1988; Oates, 1992; and Tietenberg, 1997.

Energy is important and does matter, but is subject to periodic bouts of complacency caused by apparently plentiful supplies and declining prices. Energy's importance arises from a number of dimensions. Energy is a key input into the economic process which transforms factors of production into goods and services (**Nordhaus, 1980**).[2] This role as a factor input under the generic of 'land' was emphasized with the advent of the Industrial Revolution. At a very basic level this was about access to energy to expand the production of energy-intensive activities (**Humphrey and Stanislaw, 1979**; **Rosenberg, 1980**, Smil, 1994b; Fouquet and Pearson, 1998; Rosenberg, 1988). Energy is also a key input into that rather elusive economic concept – 'the standard of living' (**Dilnot and Helm, 1987**; Burney, 1995). Ever since the discovery and use of fire, energy's contribution to human wellbeing has included both the material provision of warmth, light and improved food, and the provision of leisure (for example as a focus for story telling), and spirituality, as an object of worship. Societies' use of energy varies enormously, reflecting basic and fundamental differences of geography, climate, culture and economic systems, and the interpretation of such differences can be a statistical minefield (Desai, 1978; Smil, 1994a; Park and Labys, 1994; **Pagá and Gürer, 1996**).

In addition to being a contribution to production or utility functions, energy, and especially oil since 1973, has been a key factor in the macro economy (Fried and Schultze, 1975; **Corden, 1976**; Mathiesen, 1981; Pereira et al., 1987; Heal and Chichilnisky, 1991; Mitra, 1994; **Mork, 1994**; Lee and Ratti, 1995; Balabanoff, 1995b; Beenstock, 1995; Huntington, 1998). This is true at an international level. Oil remains the largest single category in international trade whether measured by volume or value. Movements in the oil price have been an important influence in the world economy, and some have attributed changes in the price as a causal influence on global business cycles.

At a national level, energy (and especially oil) also matters. Energy importers have faced serious balance of payments problems and resource transfer impacts from the oil shocks which have led to extensive (and often disastrous) forays into policy solutions to manage import dependency and fears of ever greater energy scarcity. Energy exporters have also suffered in terms of their economic performance. What economic theory insists should have been a blessing in a world facing a dual gap of capital and foreign exchange constraints often proved to be a curse. It distorted the role of government in an economy, led to severe attacks of 'Dutch Disease' and often resulted in hopeless attempts at resource-based industrialization (Gelb, 1986; **Neary and van Wijnbergen, 1986**; Stevens, 1986; Rowthorn and Wells, 1987; Auty, 1988; Jazayeri, 1988; Mabro, 1988; Philip, 1994).

Even at local and regional levels, energy matters. Energy tends to be characterized by large, highly capital-intensive projects which inevitably are intrusive on local communities, either at an economic (Hallwood, 1990; Upton, 1996) or an environmental level. In many parts of the world this local dimension is increasingly being related to issues associated with human rights and this may well be the next big area of interest in energy economics in the 21st century.

Finally, oil has had a profound influence on international politics. It has been the cause of war and rumours of war, and some claim it has been central to the winning and losing of wars (Sampson, 1975; Doran, 1977; Nore and Turner, 1980; Yergin, 1991; Kohl, 1991; Tempest, 1993). One member of the US Congress is alleged to have said that if Kuwait had produced carrots there would have been virtually no reaction to Saddam's brutalization of Kuwait and probably no brutalization in the first place. Apparent global dependence upon the oil reserves of the Middle East make this pivotal role unlikely to change in the immediate future.

The energy sector also matters since it has certain characteristics which arguably emphasize market failures. This perhaps makes the energy sector less amenable to simple market forces than other sectors of an economy. The energy sector (as will be explained) is frequently characterized by very large economies of scale, scope and density. In energy, big is beautiful, at least in the eyes of accountants concerned with profit and loss. It is highly capital intensive, involving projects with long lead times, in which entry and exit tend to be rather expensive.[3] The most extreme example now is oil refining: in the US, Europe and Japan, environmental legislation makes it prohibitively expensive to close plant. Oil refineries will not die – they will simply cease to be used. Links in the energy supply chain also have strong elements of natural monopoly, notably grid delivery systems for gas and electricity. The result of all these factors is that the energy sector faces serious barriers to the creation of effective competition. This is not to say competition in energy is unattainable. Rather that it is more difficult to achieve than in the case of the normal sorts of widgets which preoccupy economists. Similarly, the production and consumption of energy generates very significant externalities, most obviously in the field of environmental concerns but also with respect to energy's strategic importance. It is even possible to argue that while exclusion from access to energy is clearly technically feasible, it is politically and morally dubious. This gives energy the characteristics associated with a public good. Thus, given much greater market failure, perhaps energy requires greater government intervention than is currently fashionable among the many obsessed with the virtues of market forces. As Section 4.3.1 points out, this is a controversial area.

This section has tried to address briefly the question why energy matters. There are many different ways of then approaching the actual subject of energy economics, the study of which has been complicated by tremendous outpourings in the literature, especially since 1973. Some are referred to in this introduction and a selection are reproduced in these two volumes. Some of the literature is good, simple economics which illuminates issues and problems. Unfortunately this is not true of all the economics literature on energy. The first oil shock encouraged many in the profession who knew little or nothing of the intricacies of the energy industries to apply theoretical concepts, and from these to draw conclusions. It was a case of not letting facts get in the way of good analysis, but those who failed to get to grips with the facts were often extremely wide of the mark, as will be discussed in the section on depletable resources. In such cases, economics was a positive hindrance to understanding the energy industries (Stevens, 1995b). Within the various energy industries themselves, this has left an unfortunate legacy of indifference and suspicion of academic economics. This is a great pity, since the real power of economics is to explain highly complex systems in a sensible and accessible way. Whatever else the energy industries are, they are highly complex, involving a mixture of economics, politics, science and engineering. Economics is ideally placed to illuminate the issues.

This introduction is intended to survey the subject, and pursues the simple paradigm upon which economics is based. Unlimited demands for energy (at least definitionally) mean energy resources are scarce. The issues of interest are how choices are made to allocate these scarce resources, and the implications of the allocation. Section 2 is devoted to issues concerned with energy demand. Section 2.2 considers a sub-issue of demand conservation, in particular why conservation is so difficult to achieve by market means hence aggravating the seemingly unlimited demand for energy. Section 2.3 contains a brief sample of the huge amount of work on modelling in energy economics. The location of this section within demand issues reflects the fact that much of the modelling started in the area of demand, resulting from both a well established methodology and relatively good data. The section strays into other areas of energy modelling, however, including supply and markets.

The next stage is to consider the resource dimension, and Section 3 considers various supply side issues. In particular, it seeks to address the economics of exhaustible resources (which in the opinion of this author has been one of the worst examples of the application of economic theory with insufficient reference to fact). Other supply issues addressed concern fiscal issues and issues of costs. Transport of energy is also considered in this section.

The scarcity which arises from the unlimited demands for energy forces choice, and the next three sections – 4.1 Market Players, 4.2 Markets and Prices, 4.3 Markets Versus Governments – are concerned with various aspects of how markets solve this problem of choice. Finally, Section 5 is devoted to a particular aspect of energy economics, namely its role in developing countries.

The choice of articles for these volumes has been difficult. Much of the best literature in energy economics is down to earth and practical. It is often a description of what has actually happened in the industries and markets. Complex theory has tended to be unhelpful because the requirement of mathematical tractability has often forced simplifying assumptions which cannot be subsequently dropped, leaving the analysis unrealistic. The most effective use of economic theory has been used to order thoughts and arguments and to allow simple approaches to highly complex issues. With the exception of the work of Hotelling – which as will be argued has proved to be extremely unhelpful – little theory has been developed explicitly for the subject, although many useful general principles have been deduced by some from observing the behaviour of the various energy industries. For all these reasons it is difficult to identify seminal articles, and the resulting collection is a rather idiosyncratic selection. The volumes present two types of article. The first provides general surveys of issues to cover the broad range of questions relevant to energy economics. The second then tries to provide a sample of best-practice answers to illustrate how it might be done. Inevitably these articles do not provide comprehensive coverage and in large part reflect the editor's own research interests and prejudices. This in part explains a certain emphasis on oil issues.

2 Demand Side Issues

2.1 The Nature of Energy Demand

The technical nature of energy is important for analytical purposes, and some understanding is needed if the analysis is to make sense.[4] Energy can be measured either by the energy content of a source in calorific terms, as calories or British thermal units (Btu), or by form in terms of weight or volume. Clearly aggregated energy data must be measured in calorific terms – usually oil or coal equivalence. Table 1 provides a rough guide to the various conversion possibilities, although it must be noted that sources use different conversion factors and that the differences can be significant. For example, the Chinese authorities in 1990 used a multiplier of 0.71 to convert tonnes of raw coal to standard coal equivalent. However, it can be argued (Smil, 1994a) that 0.64 would be more appropriate given the poor quality and treatment of Chinese coal. This statistical difference comes to some 82 million tonnes which is equivalent to the total coal consumption of the UK and France in the same year (BP Amoco, 1999).

The energy content of fuels is also complicated by a distinction between the laboratory content of energy (a joule is measured by 'the energy conveyed by one watt of power for one second' (Slesser, 1982, page 155)) and its value 'at the burner tip'. Two conversion stages are involved. The first is to convert primary energy into useable energy. Primary energy is defined as the hydrocarbons, biomass and primary electricity.[5] Oil refining (which converts crude oil into oil products) provides such an example, as does the washing and sorting of coal or the

Table 1 Some Simple Conversion Factors

One tonne of oil equivalent equals approximately:
Heat Units
= 10 million kilocalories
= 42 gigajoules
= 40 Million Btu
Solid Fuels
= 1.5 tonnes of hard coal
= 3 tonnes of lignite
Gaseous Fuels
= 1,111 cubic metres of gas
= 39,200 cubic feet of gas
Electricity
= 12 megawatt-hours
= Produces about 4000 gigawatt-hours of electricity in a modern power station.

Source BP Amoco, 1999

stripping out of liquids from natural gas. The useable energy is then converted into useful energy in the form of light, heat or work. The amount of useful energy delivered 'at the burner tip' will depend upon the efficiency of the conversion appliance. Thus two cans of gasoline which theoretically contain the same amount of useable energy will produce very different levels of useful energy (in terms of miles driven) if one is put in a small, efficient, modern car, and the other in a large, gas-guzzling, poorly maintained old car. Such issues significantly compound the complexities associated with energy pricing and the operations of energy markets.[6]

Many sources of energy data are available for economic analysis. Among the most commonly and widely used are the BP Amoco *Statistical Review of World Energy*, published anually, and the data sets provided by the International Energy Agency (IEA) based in Paris. In addition, each industry tends to have its own sources of statistical information. The IEA provides unparalleled levels of detail both by fuel and sector, but these data sets are notorious for revisions and peculiar aberrations. Both the IEA and BP Amoco data are available in spreadsheet format, which is good news for potential number crunchers but bad news for the results when the data limitations are ignored. Unfortunately, this has happened more often than it should!

Table 2 illustrates the patterns of world primary energy consumption since 1925. The key point about energy demand is that it is a derived demand. Energy is not wanted for its own sake but for the light, heat and work which it can provide. The ability of any energy form to do so is a function of its energy content and the efficiency conversion embodied in the energy-using appliance. Table 3 illustrates these efficiency conversions for a selection of energy-using appliances. Hence energy demand is not only a function of the usual variables – own price, income, price of other goods and tastes. It is also, crucially, a function of the stock of relevant energy-using appliances. Thus the energy-consumption decision is the result of three subdecisions. First, to buy or not to buy the appliance. Second, what type of appliance

Table 2 World Primary Energy Consumption 1925–1992

Million tonnes of oil equivalent

	OIL	GAS	COAL	HYDRO	NUCLEAR	TOTAL
1925	295	72	1845	15	0	2227
1938	567	150	1938	34	0	2689
1950	1083	378	2390	65	0	3916
1965	1537	630	1493	80	7	3747
1972	2559	1036	1558	111	38	5302
1982	2788	1316	1878	155	232	6369
1992	3128	1781	2164	187	532	7882
1998	3389	2016	2219	226	627	8477

Percentage

	OIL	GAS	COAL	HYDRO	NUCLEAR
1925	13.2	3.2	82.8	0.7	0.0
1938	21.2	5.6	72.1	1.3	0.0
1950	27.7	9.7	61.0	1.7	0.0
1965	41.0	16.8	39.8	2.1	0.2
1972	48.3	19.5	29.4	2.1	0.7
1982	43.8	20.7	29.5	2.4	3.6
1992	40.1	22.9	27.8	2.4	6.8
1998	40.1	23.8	26.2	7.4	2.7

Note This is based on two time series which are not comparable (see Source). The pre-1965 data tends to overstate consumption compared to the post-1965 data

Source 1925–50 Darmstadter et al , 1971. 1965–1998 BP Amoco, 1999

Table 3 Approximate Conversion Efficiencies of a Selection of Appliances, and Type of Energy Transformation

Appliance	Efficiency	Type of Conversion
Electric generator	95%	Mechanical-electrical
Hydro-electric turbine	90%	Gravity-electrical
Domestic boiler	75%	Chemical-thermal
Florescent light bulb	25%	Electrical-radiant
Car engine	15%	Chemical-thermal-mechanical
Steam locomotive	10%	Chemical-thermal-mechanical
Incandescent light bulb	5%	Electrical-radiant

Source Allen, 1992

to buy in terms of which fuel (where technology allows a choice) and what levels of efficiency.[7] Finally the user has the choice of capacity utilization Despite this fairly obvious point, it was only in the early 1980s that economists working in the empirical area began to incorporate into their demand functions an appliance-stock variable (Hawdon, 1982). Even now empirical work on energy demand frequently appears with no appliance stock variable (Al-Faris, 1997).[8]

Two broad trends emerge from Table 2 which energy economics has tried to explain. The first trend is the inexorable rise in overall demand. This reflects the growth in energy-using appliances as both output and the standard of living have increased. Much of this has been driven by technological change. Table 4 illustrates some of the major technological discoveries which have driven this thirst for energy.

Table 4 Technology and Energy-Using Appliances

Invention	Date
Vacuum filament light bulb	1881
First car	1886
First powered flight	1903
Model T	1908
Household refrigerator	1913
Military tank	1916
Commercial jet airliner	1953
Commercial nuclear station	1956

Note These dates are controversial since it appears to be a matter of national pride when certain things were invented

Source McNeil, 1990

The second trend is the changing structure of energy consumption. Energy consumption is characterized by what energy economists have dubbed transitions. The USA, for example, since 1860 has gone through two transitions (Eden et al., 1981). The first began around 1860, when wood fell from over 80 percent of primary energy consumption to be replaced by coal at over 80 percent by 1900. The second transition occurred this century, when oil (and to a lesser extent gas) pushed out coal as the dominant fuel. Much of the recent interest in energy transitions concerns the experience of developing countries where dependence upon traditional energy is replaced by growing commercial energy consumption (Pearson, 1988).

There are many different definitions of traditional versus commercial energy. A useful one is to characterize them in terms of their economic characteristics (Pearson and Stevens, 1987). Thus commercial energy is energy which moves in corporately controlled markets and involves a foreign exchange input. This includes oil, gas, coal, primary electricity and modern renewables. Traditional energy on the other hand moves in small local markets (or is not traded at all in a subsistence context) and involves no foreign exchange input. Traditional energy includes wood together with vegetable and animal residues. Some refer to a traditional energy as 'non-commercial energy' but this is misleading since much of it is actively traded. Reference to animal and vegetable 'wastes' in the literature is equally misleading, since 'waste' implies a zero-opportunity cost which is not the case.[9] In developing countries, traditional energy is important (Hall, 1991). Indeed the 75 percent of the world's population in developing countries relies on biomass for 35 percent of its energy. In some countries, for example Ethiopia, Tanzania and Nepal, the figure exceeds 95 percent (Pearson and Stevens, 1992). As countries develop, consumers switch away from traditional energy to commercial energy. This transition is one of the strongest signs that economic development is taking place. One consequence is that for a period, demand for commercial energy grows very rapidly. For example, the Newly Industrialized Countries represented by the six 'Asian Tigers' between 1981 and 1991 added 2 million barrels per day (b/d) to world oil demand (BP Amoco, 1991).

Some elements of global energy transitions since 1925 are apparent from Table 2. Between 1925 and 1972, the main trend was increased oil use at the expense of coal. This was due to two factors. The price of oil during this period fell in real terms relative to that of coal and, after 1960, fell in absolute terms. Second, many of the additions to the appliance stock occurred in transportation principally for oil. Oil's seemingly inexorable rise was halted by the oil price shocks of the 1970s. Conventionally, the first price shock is dated from October 1973 to January 1974, the second price shock from October 1978 to October 1981. In 1986, the oil price collapsed. Since this was as much a shock to the producers as the price spikes of the 1970s were to the consumers, this has been entitled the third oil shock.

During the 1970s, gas (which became significantly more viable at higher energy prices) began to make ground. Nuclear also expanded in this period, prompted largely by fears for the security of supply of oil. The expansion in new nuclear capacity was halted following the accidents at Three Mile Island in the US (1979) and Chernobyl in the USSR (1986). However, because of the long lead times, nuclear capacity continued to come onstream and further expansion was considered in developing countries (Mitchell et al., 1996). Coal also began to regain market share, pushing out oil in the static sector (mainly electricity generation and raising heat under boilers). Following the collapse of oil prices in 1986, oil demand began to pick up again. This reflected a fundamental advantage of oil over other energy, namely that transportation cost per unit of energy is extremely low and does not require large-scale specific infrastructure investment.

2.2 *Conservation*

Conservation implies less useable energy input per output of useful energy and has been a constant theme of interest in energy economics. In the 1970s, a world fearing energy shortage prompted concern over investment in energy conservation as an alternative to investment in supply. This led many to describe conservation as 'the fifth fuel' although this is misleading in terms of the economic effects of reducing energy demand rather than increasing supply (Weyman Jones, 1986). More recently, interest in conservation was revived by concern over environmental damage from energy use. Conservation was seen by many as the basis for a 'no regrets strategy' to combat the perceived threat from global warming. Thus even if global warming due to energy emissions proved to be a myth, benefits from the policy would still ensue.

Interest in conservation focused upon two areas. The first was the likely impact of widespread encouragement of energy conservation. The second was the extent to which the conservation decision could be left to market forces or whether, because of potential market failure, governments should intervene.

Claims made for the benefits of conservation have been exremely controversial, often because of a failure to distinguish between the technically feasible, the economically feasible, and the market feasible (Jaffe and Stavins, 1994; Metcalf, 1994; Stanstad and Howarth, 1994; Haugland, 1996). The distinction is well illustrated by use of simple production function analysis. With capital and energy inputs on the two axes, the technically feasible is at the point on the isoquant which uses least energy. However, this ignores the capital cost implications. The economically feasible is given by the iso-cost line reflecting the relative prices of energy and capital, and the optimal level of energy use is given by the point of tangency with the isoquant. However, this analysis too is incomplete since it ignores the possibility of market failure whereby the choice of technique is sub-optimal.

Many of the more extravagant claims (often from green pressure groups) are based upon what would be technically feasible if everyone magically had access to state-of-the-art technology. This ignores the financial viability of providing that access and may also require a cavalier attitude to individual preferences. It is important to distinguish between conservation and deprivation. Sitting in the cold and dark saves energy but cannot be described as 'conservation'. When economic feasibility is considered, much that is technologically possible is ruled out because of the excessive cost of implementation. In the 1970s, some engineers developed the concept of 'energy analysis', arguing that goods and services should be valued on their energy input rather than in money (Webb and Pearce, 1975; Thomas, 1978; Webb and Ricketts, 1980; Slesser, 1994). The obvious error in this approach is that energy is not the only scarce resource.

However, even what is economically feasible may not happen, because of market failure **(Reddy, 1991)**. The basis of the consumer's decision is whether the higher price of the more fuel-efficient appliance is more than offset by the present value of the savings in fuel costs. This raises many of the issues which determine what conservation is feasible in market terms **(Bates, 1991)**. First, the price differential – the first cost hurdle – often covers more than just the cost of improved energy efficiency. Technological improvement embodied in appliances tends to be incorporated in step fashion. The appliance is only redesigned periodically, and several new elements may be incorporated at once. Thus the more expensive, fuel-efficient appliance also tends to be top of the range with respect to other specifications. Second, the calculation of the fuel saving requires a level of information that most members of the public do not have. The difference in fuel consumption measured in litres, kilowatts or whatever must be translated into cash by the application of an expected fuel price, and then discounted to a present value. Appliance labelling can obviously assist consumers, but claims of running-cost savings made in the stores selling white goods often look less convincing when the sales person is asked what discount rate is being used. Some see this information-driven aspect of market failure as a reason to regulate for minimum efficiency standards.

Third, the capital cost of the appliance and its running costs may be borne by different people. Even highly energy-efficient appliances may not be adopted if the person paying the higher capital cost (say the landlord) does not benefit from the lower running costs which accrue to the tenant. Finally, there is the issue of the discount rate used explicitly or implicitly by the consumer to calculate the present value of the stream of savings on energy costs. All the empirical evidence suggests consumers use extremely high discount rates. This significantly reduces the benefits from future energy savings. Some have debated whether such behaviour is illogical and can be regarded as a form of market failure. However, others have pointed out that such high discount rates reflect a combination of poverty and high levels of uncertainty about the future. (**Train, 1985**; **Hassett and Metcalf, 1993**; **Awerbuch and Deehan, 1995**).

Another aspect of conservation of interest to energy economists is the 'rebound effect' (**Greene, 1992**). If improved efficiency lowers running costs, then appliance utilization may increase, thereby increasing energy consumption. The extent to which this occurs depends on the nature of the appliance. For appliances which require continual running, such as fridges, the argument does not apply. The debate, empirically still open, is of importance only if the objective is to reduce energy consumption, as for an environmentally driven programme. Otherwise, a consumer's decision to take the benefit in greater utilization is perfectly legitimate. For example, in some cases of improved woodfuel stove programmes aimed at reducing wood use, consumers have actually increased use, taking the efficiency benefits in the form of more cooked food (Barnes et al., 1994).

An area related to conservation is the experience of demand-side management as a pro-active policy to promote energy conservation (London Economics, 1994; Sioshansi, 1994; Sioshansi, 1995; Wirl, 1997).

2.3 Energy Modelling

Energy modelling involves developing a simplification of a real-world process or system. The model may be physical, graphical, verbal or mathematical. Its purpose is to help explain, predict and control some aspect of energy in the real world. This immediately creates problems since the real world is infinitely complex. The modeller must maintain manageability while endeavouring to capture the essential elements of the energy process or system. Energy models range from the simple analysis of the direct relationship between two or more variables to a highly sophisticated analysis of the whole system (Munasinghe and Meier, 1993; **Bhattacharyya, 1996**).

Energy models differ with respect to their formulation, scope and application. In formulation, energy models tend to be based upon statistical and econometric methodology, optimization where linear or non-linear programming techniques are applied, input-output analysis and scenario building. In scope, energy models can cover sectoral, national, regional and international issues.

As energy rapidly moved up the policy agenda after the first oil shock, so too did energy modelling (**Nordhaus, 1973**). The applications were many and varied, and what follows can give only a very brief and superficial view of a huge and generally technically complex literature.

2.3.1 DEMAND MODELLING

In terms of application, the first area of interest in energy modelling tended towards energy demand, reflecting better availability of data than in other areas of interest. Much of the modelling work done in energy economics has tried to explain the shape and driving forces behind the data in Table 2 (Nordhaus, 1977; Maddala et al., 1978; Hogan, 1988; Hawdon, 1992; **Watkins, 1992b**; Bentzen, 1994; Kaufman, 1994; Han and Lakshmanan, 1994; Dahl, 1994; Jones, 1994 and 1995; Madlener, 1996). Initially it tended to be focused on the industrialized countries, reflecting the availability of data. More recently, greater data availability has enabled such studies to be made for the developing world (**Bhatia, 1987**;

Ibrahim and Hurst, 1990; Dahl, 1992, Al Faris, 1992, Balabanoff, 1994; Pagá and Birol, 1994; Reddy, 1995, Gately and Streifel, 1997)

The determinants of energy demand through the methodologies of demand modelling provided an area of extensive empirical research. Since this interest was triggered by the first oil shock of 1973, much attention has focused on the own-price elasticity of demand for energy. As described earlier, the consumption of energy involves a three stage decision – the decision whether or not to buy the energy-using appliance; the choice of appliance in terms of fuel type and efficiency; and finally, the behavioural decision on capacity utilization. The capacity-utilization decision given the appliance stock determines the short-run price elasticity. Most research suggests this lies between –0.2 and –0.5 (Heal and Chichilnisky, 1991) which makes demand unresponsive. It is important in this context to remember the distinction between conservation and deprivation. Some reduction in capacity utilization can be achieved with acceptable levels of inconvenience. Small reductions in temperature or lighting intensity can be managed. However, beyond some level of capacity reduction, conservation becomes deprivation. The extent to which own-price elasticities are symmetrical when prices rise or fall is also of some interest in the literature (Sweeney, 1987; Dargay and Gately, 1995).

Eventually, it becomes feasible to change the energy-using appliance to one that converts energy inputs into output more efficiently, or one that uses a different fuel. As this changes the stock of energy-using appliances, energy consumption patterns change, reflecting the long-run price elasticity. Estimates must be treated with caution, since the effects of the first oil shock are still feeding into the system. This very long delay reflects the time taken to alter the stock of certain energy-using appliances, for example housing. A convergence of empirical opinion suggests a long-run own price elasticity of –1.25 (Heal and Chichilnisky, 1991), which makes demand eventually responsive.

The crucial determinants of demand are the level and expectations of consumer prices – the market price plus any indirect taxation. Expectations of prices are particularly important since these will drive the decision over what type of energy-using appliance to buy. An area of conjecture for the future is how far consumer governments will raise such prices (Seymour and Mabro, 1994). They might raise prices to reduce consumption, because of concerns over the environment or security of supply. Or they might want to raise revenue. Energy is an ideal source of government revenue, because large and easily identifiable suppliers make for low collection costs. Energy's widespread use gives it a large tax base, and the inelasticity of demand means it can carry a high tax rate. Finally, revenue greed by politicians can be wrapped up in environmental concerns.

Another important variable determining demand is the income elasticity reflecting the impact of income and output. During the 1950s and 1960s, the global average income elasticity was around one (Schipper and Meyers, 1992). This appears to have fallen to around 0.8. However, this global average disguises the possibility that in the industrial countries, the link between income growth and energy-demand growth may have been broken following the two oil shocks of the 1970s. Meanwhile, in the developing countries, the link remains. For example, for electricity, the income elasticity in developing countries is high at 1.4 to 1.8 (Pearson and Stevens, 1992)

Another aspect of substitution which has attracted attention in the literature has been the relationship between energy and capital. Are they substitutes or complements? The theory at both micro and macro levels provides contradictory views. At the level of an individual appliance, less energy can be used by increasing the technical complexity. This makes capital a substitute for energy. However, there may be a complementary relationship if higher running costs reduce capital utilization. At the macro level, increased energy costs reduce the use of energy-using capital appliances (a complementary effect). However, higher energy costs also increase the use of capital in less-energy intensive sectors (a substitution effect). Empirically, the situation is equally ambivalent. European data points to substitution while the US data points to complementarity (Berndt and Wood, 1975 and 1979; Griffin and Gregory, 1976; Heal and Chichilnisky, 1991; Hisnanick and Kyer, 1995).

As the availability of the data has increased, especially with regard to the level of detail, so it has been possible to address ever more complex issues at very disaggregated levels. (**Greening and Tarn Jeng, 1994**; **Ironmonger, Aitken and Erbas, 1995**). This type of approach, although highly dependent upon the availability of detailed data, is certainly where the future of such research lies.

2.3.2 SUPPLY MODELLING

Supply side modelling has a long history in the context of upstream oil and gas. There has also been some focus on other areas of supply modelling. In downstream oil there have been many studies on optimizing behaviour in refineries (Masseron, 1990) and much of the early optimization in linear programming was orientated towards solving the problem of how to match diverse (in terms of both geography and quality) crude supplies with diverse refinery requirements (Bamberg, forthcoming). More recently there have been a great many highly technical studies on electricity dispatch, as newly deregulated/privatized electricity sectors try to come to terms with optimizing supply in a market-driven context (Smeers and Yatchew, 1997).

This section focuses on the oil and gas upstream studies which have been undertaken. Conceptually, production will be determined by the expected after-tax profit of each barrel (or cubic foot) produced. In theory this requires consideration of sales price, production costs and taxation. In practice many of the studies simply examined the links with price (Griffin, 1985; Griffin and Jones, 1986; Kandel, 1990; Jones, 1990; Cleveland and Kaufmann, 1991; Dahl and Yucel, 1991). As a result, the subsequent estimated equations were frequently unsatisfactory with the computed elasticity having the wrong sign. It was clear that focusing on a simple relationship between supply and price missed the complex interplay of depletion, technical change and other economic and political factors. To address these concerns, some studies tried to add past prices, to reflect an element of expectations, and some variable (often reserve changes or past production) to reflect depletion (MacAvoy, 1982; Hogan, 1989; Chermak and Patrick, 1995).

The first attempt to model elasticities in the context of exploration was Fisher, 1964. He incorporated wild cats drilled, the fraction which are successful and the average size of each successful discovery. He then estimated an equation showing total new reserves with respect to price, although each of the three variables had its own elasticities – the drilling elasticity with respect to wild cats, the success elasticity and the discovery elasticity to reflect the size of the find. Following from Fisher, in Erickson, 1968, and Erickson and Spann, 1971, variables were added to try to pick up the effect of exploration by large integrated firms. They also represented crude oil and natural gas as a joint product of petroleum exploration. Khazzoom, 1971, developed the Fisher model further by estimating separate equations for exploratory and development drilling.

MacAvoy and Pindyck, 1974, followed the disaggregation by well type and also split the US into geological regions – onshore versus offshore. The statistical results however were unstable with some very doubtful test results. Pindyck, 1974, estimated the wild cats drilled and success rate equations in a gas model for the US. Eyssel, 1978, extended Fisher's model both in terms of the time series (to 1970) and widened the detailed regional coverage. He also developed the disaggregation of wells by estimating for total drilling activity, new field wildcats, and other exploratory wildcats (new pool, deeper pool, shallower pool and outposts). He also attempted to include structural aspects of the industry, and economic variables other than price. Well depth and lagged discovery size were used as surrogates for the rate of return, and the proportion of wells drilled by independents (as opposed to the majors) was used as a surrogate for industry structure. A similar study was done by Kolb, 1979, again with variations in the districts, with results which confirmed Eyssel's results. Measuring exploration effort against price is another alternative. Scarfe and Rilkoff, 1984, estimated oil and gas exploration expenditure in Canada with respect to a hybrid price of undeveloped reserves.

Many other studies have been completed that do not directly follow Fisher's approach but involve similar elements. A common approach has been to consider only drilling levels (Rice and Smith, 1977; Deacon, DeCanio and Johnson, 1990; Ghouri, 1991). In a variation, Cleveland and Kaufman, 1991, considered the ultimate recovery of domestic oil reserves by utilizing a Hubbert type model – Hubbert, 1962. An excellent survey of many of these studies is provided by Dahl and Duggan, 1996a.

Models were also developed which allow the reserve base to be augmented by new discoveries (Uhler, 1979; Pindyck, 1978a; Barrett, 1986; MacAvoy, 1982). Generally the results were disappointing, almost certainly because of the over-simple specification of the models, with factors such as price controls, depletion and technical change being ignored. Other studies considered the impact of taxation (Griffin and Moroney, 1985; Walls, 1994; Berman and Tuck, 1994; Erikson, Millsap and Spann, 1974), the impact of company size (Iledare, Pulsipher and Baumann, 1995), and production function approaches (Epple, 1975; Dahl and Duggan, 1996b; Deegan, 1979).

An interesting new area in the literature is the development of option theory and its use as a decision tool in choosing to develop or not. Examples of this literature include Ekern, 1988; and Favero, Pesaran and Sharma, 1992.

There are many complications in modelling oil supply but an area that has attracted particular attention in the literature concerns the actual motivation of the owners. In the face of limited domestic absorption capacity in some countries, it has been argued that foreign investment could provide a worse solution than leaving the oil in the ground (Eckhaus, 1973; Ezzati, 1976 and 1978; Moran, 1978; Cremer and Salehi-Isfahani, 1980 and 1989; Scott, 1981). This has given rise to speculation that for some producers the supply curve above a given price will become backward-bending, and that supply was therefore being manipulated, with given prices to achieve a specific revenue. This has been increasingly discussed in the literature – Adelman, 1980 and 1982; Benard, 1980; Reza, 1981; Scott, 1981; Bohi and Montgomery, 1982; Noreng, 1982; Salehi-Isfanahi, 1986; Linderoth, 1992. Such an explanation, it has been argued, could explain the discovery in some empirical studies of negative price elasticities, which have tended to be dismissed by the authors because they were 'not what can be expected'.

Eventually, as more data on upstream oil and gas became available, more sophisticated modelling methods produced a batch of new studies. Pesaran, 1988 and 1990, developed an inter-temporal model of exploration and production policy of price-taking suppliers in the UK North Sea where the optimal decision rules were derived by solving a constrained stochastic inter-temporal profit maximization problem. Favero, 1992, developed the Pesaran model by including taxation. The paper showed that the tax system had been non-neutral and therefore should have been included in the analysis. Inclusion of taxation produced more plausible estimates. Favero and Pesaran, 1994, developed the model further. Specifically, the development process is explicitly included in a model of the investment decision. The paper is also able to take advantage of longer time series than either of the earlier papers, which enables more account to be taken of the lags which exist between price and tax changes and oil supplies. The most interesting aspect of the results is that they identify very long lags between changes in price or tax, and changes to oil supply. These lags appear to be largely determined by the average lengths of the appraisal and technical development periods, which at the time of the study were in the region of five to six years.

Simulation models (or play analysis) are used for relatively undeveloped areas. They use geological information including seismic and other pre-drilling information to assign probabilities to the discovery of commercial oil or gas. Then, by use of the Monte Carlo method, the probabilities are combined to estimate the presence of commercial oil or gas in a play. Examples of such models include Kalter, Tyner and Hughes, 1975; Mansvelt-Beck and Wigg, 1977; Eckbo, Jacoby and Smith, 1978; Kaufman and Barouch, 1978; Davis and Harbaugh, 1981; White, 1981; EIA, 1982; and Bird, 1986. The obvious limitation of such models, of course, is that they only identify oil-in-place and say nothing about whether this will be developed into producing capacity, although the results can be used to examine the effects of different economic scenarios, leasing rules and development constraints on future supply. A shorter description of the EIA model can be found in Farmer et al., 1984 which gives a good flavour of the type of approach.

Discovery process models are essentially based upon statistical sampling techniques. They rely on the observation that discovery rates in a region decline with cumulative exploration. This statistical relationship from other regions can then be extrapolated to a region to forecast future discoveries. The technique was first developed by Arps and Roberts, 1958. More recent examples of the technique include Prato and Miller, 1981; Drew, Schuenemeyer and Bawiec, 1982; Gas Research Institute, 1985; Vidas, Manger and Woods, 1985; and US Geological Survey (USGS), 1980. As with simulation models, the economic dimension is very limited. Although economic components are included, they are not subject to validation against historical data. Also they predict once-and-for-all changes in endogenous variables – for example exploratory drilling – given a change in an economic variable such as price or costs. This makes the generation of time series forecasts impossible without some form of ad hoc constraint.

Out of this earlier work came hybrid models, developed because the simulation and discovery models tended to neglect economic factors, while the econometric models (according to proponents of hybrid models), fail to account properly for non-renewability and other physical properties which influence the supply process. Early versions of hybrid models include Attansi, 1979; Attansi, Drew and Schuenemeyer, 1980; and EIA, 1988. A more recent attempt is Walls, 1994, which develops a model to examine the Gulf of Mexico outer continental shelf. The model is estimated on data for 1971–88 and is used to forecast for 1989–2000. Walls argues that offshore modelling is difficult for two reasons. First there is the important influence of government behaviour, which the model accounts for by including the number of tracts leased in the drilling equation. Second, there are strong lags in the response of drilling to changes in economic variables, and in the response of development and production to successful drilling. The approach seems to offer significant promise for the future. However, the method is extremely demanding of data, which would certainly restrict the areas of the world where it might be used.

In reality, expectations about determining variables are more important than their actual level at any point in time. However, it is difficult to account satisfactorily for such expectations, especially because the modeller is unable to secure a long enough time series. For example, Walls, 1994 regrets being unable to undertake a rational expectations hybrid because data is only available from 1971. In such a model the econometric equations would be derived from a stochastic dynamic optimization framework. 'Laws of motion' govern the stochastic variables such as prices and discoveries, and are used to form rational expectations of their future values. Joint estimation of these 'laws of motion' and the equation that solves the optimization problem are then performed. However such manipulation requires a large number of observations. Walls cites three studies which attempt such an approach for gas – Epple, 1985; Hendricks and Novales, 1987; and Walls, 1989.

Schmidt, 1984, was the first to try and model the distinction between current and future prices by explicitly examining the effect of expected future prices on drilling activity. A more recent attempt to use price expectations in relation to drilling rates can be found in Kaufmann, Gruen and Montesi, 1994. Several important conclusions emerge. The price elasticity is considerably smaller than previous studies suggest, therefore policies to increase exploratory and development wells might be much more expensive than previously believed since large increases in oil prices are required to have much of an impact.

2.3.3 ENERGY MODELLING AND FORECASTING

Because of the capital intensity of much of the energy industries, energy projects tend to have long lead times. Hence the concern with forecasting when and where new capacity is needed, and when and where new market opportunities might arise.[10] During the 1970s, this tendency to forecast at a project level was reinforced by more global concerns about energy supply and price. These concerns led to block-busting studies to evaluate the global situation, driven by an underlying concern either with the availability of oil resources or the role of the Middle East in oil supplies (World Energy Council, 1993; Nakicenovic, Grubler and McDonald, 1998).

Over the years, the level of sophistication used for forecasting has changed. In the 1960s and 1970s there was a tendency to adopt simple extrapolation techniques in which the past trend was used to produce a line of best fit which was then extended (Meadows, 1972). As econometric modelling became more sophisticated so too did the forecasting methodology, although with little evidence of success. Also, as the techniques became increasingly more complex, so the value of the forecast itself diminished. For the real decision makers they became black boxes. The forecast had to be believed or not, which effectively made the forecasters increasingly responsible for decision making. In a reaction to this, many turned increasingly to scenario planning (Beck, 1981; Loasby, 1984) on the grounds that the real decision makers could be involved in the process and understand how the forecast was derived.

In general, these forecasting exercises adopted, at least as a starting point, a 'business as usual' basis for the future. Demand continued to grow with limited allowance for technical change or the impact of expectations of ever higher energy prices. When these forecasts were put against an apparently finite energy-resource base, gaps were inevitably identified. This form of 'gapology' fuelled concern about energy scarcity and in many cases prompted government intervention.

Such 'gapology' was linked to concern over the security of supply of energy. The rise of oil described earlier had significantly increased oil-import dependence by 1973. During the 1970s, it was perceived that Middle East politics was impinging upon energy supplies. The first oil shock had been associated with the Arab Oil Embargo, the second oil shock with the Iranian Revolution. Thus energy had become a strategic commodity, justifying government intervention ranging from intensive conservation campaigns to the subsidization of alternatives to imported oil. The French nuclear programme and the Brazilian alcohol programme are typical examples. Security of supply has spawned a large literature (Treverton, 1980; Howrich and Weimer, 1987; Fried and Blandin, 1988; Fried and Trezise, 1993; Bohi and Toman, 1996; Green, Jones and Leiby, 1998; Lynch, 1998b), and in many countries provoked strong regulation of the energy sector. Under agreements within the International Energy Agency, created as an offshoot of the OECD in 1974, strategic stocks were instituted and emergency sharing schemes agreed (Scott, 1994).

This approach has a number of fundamental flaws. Predictions of energy supply and demand tend not to incorporate discontinuities, though most pay homage in passing. This neglect is entirely understandable. Discontinuities make forecasting difficult. It is impossible to predict when a discontinuity will occur and when it will begin to have an influence, and very difficult to predict how long it will be before the discontinuity significantly bends the trend. However, discontinuities are crucial, because if they occur, the future will be fundamentally different. The history of energy has been strongly influenced by such discontinuities. There have been the three oil price shocks, the oil nationalizations of the first half of the 1970s, the nuclear accidents at Three Mile Island and Chernobyl, the development of combined-cycle gas turbine technology, and possibly the 'discovery' of global warming. Conventional wisdoms, of which there have been many, have failed because such discontinuities were ignored.

Many projections also take a conventional view of behavioural drivers which are seen to be broadly the same as in the past. Again, this is understandable. Dramatically different patterns of behaviour, like discontinuities, are difficult to predict in terms of both form and impact. They are not amenable to empirical justification. Too much deviance in behaviour invalidates the past, so there is no data for empirical validation. Yet behaviour can quickly change dramatically. History is littered with fundamental behavioural changes by consumers, governments and companies central to the direction of the industry. These range from the significant changes in the income elasticity of demand for oil experienced since 1970, to the dramatic reduction in offshore operating costs experienced since 1986.

Equally crucially the forecasts neglect their own impact on such behavioural drivers, through information feedback loops. If the forecasts are believed, they will inevitably influence both producer and consumer behaviour, which in turn is likely to undermine the forecasts. The classic example (Stevens, 1997) occurred in the late 1970s and early 1980s, when expectations of supply scarcity and ever-rising oil prices encouraged companies and governments to develop new supplies of oil and at the same time reduce oil intensities in the economies. The result was rising supply and falling demand, culminating in the 1986 oil price collapse which discredited the original forecasts.

The further into the future the forecasts are made, the larger are the potential discontinuities and the greater the potential for change in behavioural drivers which singly or collectively threaten the viability of any forecast.

3 Supply Side Issues

Certain energy-supply issues have dominated the literature, and they have tended to have a strong orientation towards oil and gas supply. This reflects the recurrent fear of energy scarcity arising from the alleged depletability of oil and gas reserves. In recent years, the other main supply side consideration in energy economics has concerned the deregulation, liberalization and privatization of the supply industries. The literature on this subject has been enormous and has grown exponentially since the early 1980s. As with the literature on energy and the environment alluded to earlier, inclusion in these volumes would either be superficial or swamp the volume and therefore a decision has been made to omit literature samples in these volumes. For those interested a good introduction to the issues can be found in Vickers and Yarrow, 1988; Foster, 1992; Jackson and Price, 1994; Cook and Kirkpatrick, 1995; and Mackerron and Pearson, 1996. Another useful set of publications which cover this area comes from the Institute of Economic Affairs in London (with special emphasis on regulation) and the Surrey Energy Economic Discussion Papers (SEEDS) from the University of Surrey in Guildford (with special emphasis on the UK experience). Electricity supply issues have also been important (Turvey and Anderson, 1977) but these will be covered in Section 5, on developing countries.

3.1 Depletion

Many writings on energy economics begin with an exposition of the first two laws of thermodynamics – energy can be neither created nor destroyed; and although energy can be transformed, such transformations can never be completely efficient. The implication is that useful energy is finite and must reduce as a consequence of its use in economic processes.

Within this context, much of energy economics has been concerned with the depletability of hydrocarbon energy forms – oil, gas and coal. What is produced today cannot be produced tomorrow, and postponing production today enables production tomorrow. Losing the future use is referred to as the 'user cost'. Much theoretical work has been undertaken on this issue following the original work of **Hotelling, 1931**, although there was an earlier version – Gray, 1914 – which concerned optimization by an individual mine but was much less complete than Hotelling. Also Fisher, 1930, examined a case where shipwrecked sailors allocated a fixed stock of hard-tack and concluded price should rise in line with the rate of interest.[11]

More than any other piece of economic theory used in the field of energy economics, the work of Hotelling and its successors have had a major impact at a practical level. The main thrust of the theory has been to assess the optimal production profile for a depletable resource. Two key implications of Hotelling's work have had a popular impact. The first is the view that there exists a fixed quantity of oil (or gas or coal) which could therefore 'run out'. The second is that over time, the price of the depletable resource rises in line with interest rates to reflect the difference between future and present values and ever-impending shortage.

These views became extremely powerful in the 1970s among the 'resource pessimists'. Following the first oil shock of 1973, many turned to economists for an explanation of what had happened. In the best traditions of the subject, the economists looked into their tool bags to see what theories or concepts might assist their analysis and rediscovered the ideas of Hotelling. Unfortunately, many of those who did so knew nothing about the facts of the international oil industry[12] and proceeded to ignore them, developing instead their own variations on the Hotelling theme (**Nordhaus, 1973**; Herfindahl and Kneese, 1974; Schmalensee, 1976; Stiglitz, 1976; Peterson and Fisher, 1977; Sweeney, 1977; Hoel, 1978; Lewis, Matthews and Burness, 1979; Tullock, 1979; Lewis and Schmalensee, 1980; Pindyck, 1978b and 1980; Devarajan and Fisher, 1982; Hansen, Epple and Roberds, 1985; Loury, 1986; Starrett, 1987; Hogan and Leiby, 1985). Hotelling's ideas were also incorporated into much of the analysis associated with cartel behaviour discussed below in Section 4.1.2.

The basic hypothesis to emerge was that the resource owner would plan a production profile to maximize the net present value of the fixed stock of the resource. DasGupta and Heal, 1979, stated the Hotelling Rule as $g=i$ where g is the annual growth rate of the price net of costs and i is the interest rate. Attempts to prove such a relationship failed. Miller and Upton, 1985a, restated the Rule as the Hotelling Valuation Principle which emphasized the in-ground unit value equal to the net price.

Very rapidly these ideas took hold and became standard fare in any text on energy economics. This academic basis for impending energy scarcity coincided with two other sets of forces also at work in the 1970s. Firstly, there was growing concern about the finite nature of resources, associated with the Club of Rome's report on impending natural resource shortages (Meadows et al., 1972). Secondly, oil companies operating in the USA wanted to force a deregulation of oil and gas to restore flagging profitability. To persuade Congress of the virtue of such a move, they also began an effective campaign to spread the idea that an energy crisis was just around the corner.

All these factors meant that the concept of depletability captured the popular imagination. Such views seriously influenced actual behaviour in energy markets. Oil producers restrained production from existing capacity to benefit from the expected higher future prices. They believed that 'oil in the ground was worth more than money in the bank'. Oil consumers moved away from the purchase of oil-burning appliances. While restraining production happens quickly, reducing demand takes longer. Thus, initially prices would tend to rise reinforcing belief in the hypothesis. The fact that oil prices fell in real terms between 1974 and 1978 was simply ignored. The second oil shock was perceived by many observers of the industry as undoubted evidence that oil prices would rise forever. This was despite the fact that the second oil shock had little to do with real forces of supply and demand. The mechanisms which increased the price from $13 to $34 a barrel between 1979 and 1981 had no place in the theory. Rather, it was the effect of buyers' panic in a very imperfect market.

However, the contribution to real understanding of what drives supply was at best marginal and in many cases downright misleading, for two reasons. First the theory itself was fundamentally flawed. This was pointed out by a number of economists: McKie, 1978; Houthakker, 1983; Watkins, 1992a; Adelman, 1986a, 1992a, 1993; Gordon, 1994; and most elegantly **Adelman, 1990a**. The basis of his attack was simple. The Hotelling hypothesis rested upon the assumption that there existed a fixed stock of oil which could not be reproduced, that is, oil is an exhaustible resource. Adelman argued that this quite simply is false. There exists an unknown stock of oil from which investment in producing capacity extracts a production level. At some point, possibly, the cost of this extraction will exceed the price. This could be because of growing shortage or simply because no demand for oil exists. At that point investment in the oil industry stops and the industry dies. Whatever oil 'reserves' are left is unknown, unknowable and quite uninteresting. There is not a fixed stock of the resource.[13]

The second weakness of the Hotelling literature was that those using his hypothesis ignored the facts in two areas. First, the oil shocks themselves had nothing whatever to do with the depletion of oil resources. At most, delays in investing in producing capacity may have contributed to the first oil shock, when events did take place in a tight market. Second, none of the empirical studies showed signs of the resource depletion which should have been signalled either by rising costs of production or, as a proxy, rising prices over time – Manthy, 1978; Slade, 1978; Stollery, 1983; Halvorsen and Smith, 1984; Neff, 1984; World Bank, 1984; Farrow, 1985. Even Miller and Upton, 1985b, who had created the Hotelling Valuation principle, failed to establish that the in-ground valuation equalled price and discovered instead a relationship such that the in-ground valuation was half the price. Similar studies (Bradley and Watkins, 1994; **Adelman and Watkins, 1995** and 1996) confirmed the inequality. There is simply no empirical support for the hypothesis, although one recent study (McDonald, 1994) has suggested that the results of in-ground values are distorted by regulation with respect to well-spacing and extraction rates in the US and Canada.

Despite such empirical work, those predicting imminent oil shortages periodically emerge and if nothing else provide for a good and lively re-discussion of the issues. Most recently this debate has been triggered by Campbell, 1997, and Campbell and Laherre, 1998, leading to a powerful rebuttal by Lynch, 1998a, 1999. At the risk of oversimplification, the debate is between the geologists and the economists, although many theoretical economists following Hotelling unwittingly find themselves in the camp of the geologists, and an extremely sensible discussion on the subject recently has come from a geologist (**McCabe, 1998**). In essence the geologists perceive a finite resource base while economists see the availability of resources as a function of technology and investment.

The main value of Hotelling's contribution was the observation that resource scarcity would be signalled by rising prices. The subsequent outpouring of theory adds little to this simple rule. There does exist an alternative view to the Hotelling notion of impeding scarcity which can be characterized as 'resource optimism'. For example Barnett and Morse, 1963, argued that ingenuity would more than offset any tendency to depletion, and established that the record of both price and profitability for exhaustible resources was that they had fallen over time. Koopmans, 1974, examined the applied as well as the theoretical literature, and concluded that exhaustion was not a practically relevant problem. Others using Hotelling's own model argued that exhaustion was not a pressing problem. For example, Herfindahl, 1967, Gordon, 1966, and **Gordon, 1967**, suggested that the exhaustion costs (that is, user costs) were so low as to be unimportant. Gordon also pointed out that delaying production also delayed cost increases. Others developed these ideas in greater detail (Cummings, 1969; Baumol and Oates, 1988; Modiano and Shapiro, 1980). **Gordon, 1994** also points out a number of other 'misinterpretations' associated with Hotelling. The result, he suggests, is that 'Later writers lost touch with reality. This produced sterile, empirically questionable discussions ... whatever the practical relevance of the Hotelling analysis, its policy implications are probably nil ...' (page 3).

This conclusion, however, does not necessarily preclude the importance of reserve depletion as a driver of supply, especially in mature producing areas such as the US. A recent major study of international oil reserve additions was completed by Watkins and Streifel, 1997. The main purpose of the study was to distinguish between movements along a supply function and a shift in the function itself. The inference is that a shift outwards reflects growing availability of oil which implies technology is overwhelming depletion. A shift inwards reflects growing scarcity and implies the classic outcome from the depletion of an exhaustible resource. Arguably, movements along the curve could be viewed as reflecting short-run price elasticity while shifts reflect longer-run elasticity. The study concludes that 'a gloomy outlook for world oil supply in general and for non-OPEC producers in particular is not warranted' (Watkins and Streifel, 1997, page 3).

The legacy of Hotelling's original article remains alive and well today. It pervades much of the theoretical literature on energy economics and encourages periodic forays by those who believe the oil is about to 'run out'. Seldom in economic analysis has so much been influenced by so many from so little.

3.2 Costs

Energy industries are generally characterized by large scale of operation and high levels of capital intensity.[14] In part, this is because energy activities attract economies of scale. This is especially relevant for oil and gas since they flow in three-dimensional space. For any carrying units such as tanks, pipes or processing units such as distillation towers, the capital cost is a function of the surface area while the capacity is a function of volume, and an exponential relationship exists between surface area and volume. Hence the larger the operation, the lower the average cost.

For the economics of energy, large scale and capital intensity carry crucial implications. The size of projects means long lead times which make it difficult to match capacity to needs on an industry-wide basis. Offshore oil fields have lead times of five to ten years,[15] refinery building takes four to five years and large power stations can take up to ten years, especially if planning permissions are strict. The classic example was tanker and refinery building in the period 1967–1975. Massive investments made in capacity remained idle following the fall in oil demand after 1973. Many tankers rolling down the slipway in 1974–75 went straight into 'mothballs'. They were still being built despite falling demand because the orders had been placed two to three years before. Such lead times mean the energy sector tends to experience either feast or famine. Price signals, unless correctly anticipated, will only produce a supply or demand response well into the future. The result is that the energy sector, especially oil, tends to be relatively volatile.

High capital intensity creates very low variable costs (short-run marginal cost). If the production decision is based upon equating marginal revenue with marginal cost as economists believe, there is always likely to be pressure to produce. Finally, large-scale operations with high capital intensity provide significant barriers to entry which would tend to inhibit competition.

Several aspects of the cost of producing oil have been examined in the literature.[16] One of the earlier debates was between P. Frankel and M.A. Adelman (Adelman, 1972; Frankel, 1973;

Skeet, 1988). Frankel argued that oil was an industry characterized by decreasing costs and a natural tendency to monopoly. Adelman by contrast argued that at any point in time, costs generally would rise although over time technical progress may shift the marginal cost curve down. Thus there was little that was 'natural' about the tendency of the industry to restrict competition.

Another issue was the availability of a backstop technology to oil as an important element of costs and energy supply (Heal, 1976; Hoel, 1978; Houthakker and Kennedy, 1978; Dasgupta and Heal, 1979). A backstop technology is one which provides unlimited supplies of an alternative energy source to oil. The cost of producing such an energy source theoretically places an upper limit on oil prices. Following the depletability argument of rising oil prices, the price of the backstop technology limits that rise. If, because of technical change, the price of the backstop technology were to fall, this would force down the whole oil price path. This provided an argument for subsidizing research and development into the development of such a backstop technology. Establishing a point at which the backstop became available would bring down oil prices. This would benefit oil consumers but not the developer of the backstop, who would have to wait until that point to gain a return. In the event of a monopoly-controlled market, the monopoly oil supplier might pursue 'limit pricing', keeping prices just below the backstop price to deter entry.

There has also been an active debate in the literature over determining finding costs, in-ground, value and development costs as part of the depletion-versus-ingenuity debate referred to in Section 3.1 (Adelman, 1972, 1990a, 1992b; Adelman, DeSilva and Koehn, 1989; Adelman and Shahi, 1989). Since 1986, the efforts of the oil industry to reduce costs to accommodate much lower prices by the application of technolgy have also received attention (Kemp and MacDonald, 1995).

Another area which might have been expected to receive significant attention in the literature concerns the relative cost of supplying different fuels, especially renewables (Hill et al., 1995; Boyle, 1996; EUREC Agency, 1996; Cassedy and Grossman, 1998). Concern about the finite nature of fossil fuels in the 1970s did produce much discussion (Penner and Icerman, 1975; National Research Council, 1980) as did the search for more environmentally friendly energy sources in the 1990s. However, the debate has been rather disappointing and has tended to be dominated by engineers rather than economists. Several factors explain. Much of the initial discussion revolved around the economics of nuclear (NPC, 1973; NEPSG, 1977; MacKerron, 1982; Chesshire, 1992; MacKerron, 1992 and 1996; Jones, 1996). A key problem was the veil of secrecy which surrounded nuclear because of its links with the weapons programme. Since data was scarce and extremely debatable, eventually the discussion descended into what amounted to a theological debate.[17] As for other renewables, such as wind, solar, tidal and geothermal, since these tended to be small-scale applications, there was always the problem of how to account for scaling up. Like the proponents of nuclear, proponents of renewables believed passionately in their benefits which at times clouded judgement. In an area where assumptions are key, as with conservation, discussed in Section 2.2, this often resulted in exaggerated claims and doubtful arithmetic.

3.3 Fiscal Issues

Fiscal issues are mainly relevant in the context of oil and gas. Both have an inherent value to their owners, who understandably wish to realize that value in its production. This was the original logic behind a royalty. For oil, the situation is complicated by two factors: its concealment and the presence of significant amounts of rent.

Oil and gas are concealed underground,[18] and discovery requires a phase of exploration. This carries several implications. Exploration is an extremely high-risk activity. While modern technology has helped reduce the risk of dry wells, prospect risk remains. Drilling exploration wells (so-called 'wild cats') tends to be increasingly expensive as the explorers move into more hostile environments. This means exploration requires a high return, which creates sensitivity to price expectations and fiscal terms since it is profit after tax which is relevant. Because fiscal regimes are country specific, studies have tended either to examine the impact of tax systems in a given country or make inter-country comparisons. Frequently, such studies are aimed at establishing the extent to which fiscal systems may inhibit exploration, development and production of oil. Early examples were Erikson, Millsaps and Spann, 1974; Erikson, 1968; Erikson and Spann, 1971; and Spann and Erikson, 1973.

Subsequent studies have examined the impact of fiscal systems on specific countries: for a general survey, Kemp, 1992 and 1994; for the USA, Mead and Deacon, 1979, Verleger, 1980, Jacoby and Smith, 1985, Mead and Muraoka, 1985; for Canada, Watkins and Scarfe, 1985; for Australia, Bradley, 1985; for developing countries, Blitzer et al., 1985; for the North Sea, Kemp and Rose, 1985, and Kemp, Reading and Macdonald, 1992. Work has also been done to determine optimal tax regimes by means of simulations (Stauffer and Gault, 1985). However, such studies are provided increasingly by consultants and academic access is limited. For a summary of some of the theoretical issues see **Blitzer et al., 1984**. Another thread in the fiscal literature is work based upon the theory of optimal lease payments in a context where acreage is effectively auctioned (Hughart, 1975; Leland, 1978; Reece, 1978 and 1979; Hyde and Markusen, 1982; Teisberg, 1980; Ahmadian, 1997). Several empirical studies have also attempted to analyse bidding behaviour in such a context (Smith, 1982; Dworin and Deakin, 1982; Mead, Moseidjord and Sorensen, 1983 and 1986).

Another consequence of the concealment of oil deposits is that the resource base is unknown during the negotiation over fiscal terms. The division is made in ignorance of the size of the spoils, which means the agreement is inherently unstable. A large discovery can make an apparently fair division into one which is excessively generous to the company (Adelman, 1972). This inherent instability is reinforced by the nature of bargaining power. Before commercial discovery and development, bargaining power resides with the companies. However, once the resouce base has been developed and much of the costs sunk, bargaining power swings dramatically in favour of the government, which can squeeze the company through the 'obsolescing bargain' (Vernon, 1971). This area of energy economics has developed as part of the growing interest in political risk analysis, but much of this literature is also in the realms of consultancy.

Finally, there is the fact that especially for oil there are significant amounts of rent to be captured. This rent stems from two sources. If oil were a competitive market rather than a managed market (as is the case in reality) then owners of lower-cost geology would secure for themselves producer's surplus – the difference between the supply curve and the ruling price. For countries with access to very large oil reserves in large onshore locations, this amounts to potentially relatively large sums (Adelman, 1972). However, given that the oil market is managed with a core group of suppliers controlling supply to defend a higher price, there exists a further source of rent which could be termed super-normal profit. Again, given the relative success of these market controllers this source of rent is also significant, and following the two oil shocks of the 1970s was for a time nothing short of spectacular. The existence of rent from both sources has led to a considerable literature on how it should be divided between the owner – normally the government – and the company undertaking to explore, develop and produce the resource (Penrose, 1959; **Garnaut and Clunies Ross, 1975**; Dam, 1976). An example in the context of minerals more generally is provided by **Palmer, 1980**.

3.4 Transport

Geology and the economies of scale and scope determine the need for energy transportation. As can be seen from Table 5, there is a mismatch between where the oil and gas are found and produced and where they are consumed. Also, economies of scale and scope encourage large single producers located in one point which must then supply a dispersed area.

Table 5 Production and Consumption of Oil by Region 1998 (Million barrels per day)

	Production	Consumption
North America	14165	21405
South and Central America	6730	4635
Europe	6885	16065
Former Soviet Union	7360	3700
Middle East	22795	4230
Africa	7525	2370
Asia Pacific	7645	19125

Source BP Amoco, 1999

The transport dimension creates a number of areas of interest for energy economists (Poirier and Zaccour, 1991). The way in which problems with transport have inhibited the spread of gas is of particular interest. The history of gas has tended to be that of a constrained industry. Given gas's enormous advantages in terms of conversion efficiencies and cleanliness, demand growth has been surprisingly slow in many cases (Jensen, 1994; Radetski, 1994).

Several factors explain. In the OECD countries, starting with the first oil shock, gas became regarded as special, indeed so special that many institutional barriers were put in place to prevent its being burnt. In the US there were restrictions on the use of gas in electricity generation. Similar restrictions in the European Union were abolished only in 1990. In the European context the monopoly (and in many cases the monopsony) enjoyed by the state-owned gas utilities meant that the price of gas was very much higher than would have been the case under competition. This also inhibited gas's spread into energy markets. Arguably, the institutional changes and reform associated with deregulation, coupled with gas's undoubted environmental advantages over other fossil fuels, may well lead to a growing role for gas in the OECD.

Inhibitions on gas use in developing countries came from a rather different source (Davison et al., 1988). During the 1970s, prompted by higher energy prices and the loss via nationalizations of their traditional sources of equity crude, oil companies went exploring. However, as often as they found oil they also found gas. Where the initial development was for domestic use and the local currency was not convertible, this created a severe disincentive for the oil company to develop the reserves. Where exports were a possibility, problems associated with transport (discussed below) meant that development of the gas in many cases was delayed. Demand for gas is a function of the supply of gas, so delayed development meant that gas failed to enter the energy balances of many developing countries.[19]

Problems associated with transporting gas were a main constraint upon the development of a global gas market. Unlike oil, gas has a very low energy content per unit of volume. Liquefied Natural Gas (LNG) projects were extremely high cost, very inflexible and often provided very little revenue to the producing governments (Pezeshki and Fesharaki, 1995; Stauffer, 1997; Bartsch, 1998). Pipelines were expensive over long distances because of the amount of energy required to move the gas, and transit pipelines also had serious problems, both economic and political (Stevens, 1998a).

Oil faced no such problems over transportation. Its ability to flow in three dimensional space plus its high energy content made it ideal as a form of transportable energy. Above all, the technical economies of scale discussed earlier transformed the international oil market. The history of the tanker industry is a fascinating story of feast and famine driven by 'forecasters' cluster' [20] and competitive forces, yet surprisingly the industry has attracted only limited attention in the literature (Hawdon, 1978; **Zannetos, 1987**; Beenstock and Vergottis, 1989; Stopford, 1995).

Another transport issue which has attracted the attention of energy economists concerns the fact that pipelines and supply grids for gas and electricity are often natural monopolies. While the state dominated energy-supply industries, the presence of natural monopoly as such caused little concern. However, once the energy industries began to undergo deregulation and liberalization the issue came to the forefront of policy issues. Control of any grid system effectively gives monopoly power to the owner, and unless checked will lead to vertical integration which will act as a severe constraint upon the development of competition. A much quoted example is British Gas following its privatization in 1986. The result has been a number of studies relating to issues of access and industry structure (Foster, 1992; Beesley, 1993, 1994 and 1995; Stevens, 1996).

4 Markets

Energy markets differ enormously by fuel in terms of their basic characteristics. Oil is a truly international market, and this is reflected in the tendency of international crude prices to converge, a subject of considerable research in its own right (Weiner, 1991; Sauer, 1994; Ripple and Wilamoski, 1995; Gulen, 1997). This convergence arises because of the relatively small proportion of transport costs in the delivered price, which in turn arises from the ease of handling oil and the economies of scale associated with pipeline and seagoing transportation (discussed in Section 3.4). Relatively small differentials in regional prices can provoke a rapid arbitrage response whereby oil physically moves between regional markets (Horsnell, 1997). This process has recently been reinforced, since the development of paper markets has meant that the differential can now be locked in, significantly reducing risk.

By contrast, traded traditional energy – wood-based fuel plus vegetable and animal residues – experiences limited movement in very small local markets. This reflects the severe constraints imposed by the extremely high transport costs relative to a very low unit value. Other commercial fuels also face constraints which make their markets either national or at best regional. The problems faced by gas have been discussed in Section 3.4. Coal has a relatively low energy content per tonne compared to oil, and its lack of fluidity makes handling difficult. This tends to restrict trade to consumers with mainly seagoing access, although rail transport can also be important, especially in the very large coal consumers such as India and China. Electricity faces the limitations imposed by transmission losses as it is pushed down a wire. These losses rise exponentially with distance and await technical change in the area of super conductivity before extensive trade becomes possible.

This section of the introduction on markets has a strong oil orientation, which reflects the reality in the literature. Issues connected with deregulation and liberalization of energy markets have recently led to much greater emphasis on gas and especially electricity markets, but much of this material is in the form of conference papers and is only just beginning to emerge into the mainstream academic literature.

4.1 Market Players

4.1.1 THE COMPANIES

The literature on oil companies has focused on three areas: horizontal integration of companies, vertical integration and company strategy.

The oil industry has always been characterized by excess capacity to produce crude oil (**Penrose, 1965**; Blair, 1976; Stevens, 1985). This excess existed for three reasons. The price of oil always exceeded the cost of replacing the barrel, generating an incentive for the owner of the oil-in-place to develop the resource. Unexpected supply loss generated rapid capacity development to offset the loss which, when the loss was restored, became surplus to requirements. Finally, the industry, as discussed below, tended to have a strong herd instinct. If one area's capacity was being developed in response to expectations of rising price, the same was true elsewhere. Had this excess capacity come to market, it would have generated severe price wars, in theory forcing down prices to their short-run marginal costs. This was the experience of the industry before 1928 (Penrose, 1968; Yergin, 1991). To end these interminable price wars, since 1928 the industry has been characterized by high levels of horizontal integration (with varying degrees of effectiveness) aimed at controlling supply and managing markets.

The control began with the Achnacarry Agreement of 1928 (also known as the 'As Is' agreement). This was effectively an explicit form of cartel (Blair, 1976) although it did not emerge into the public domain until 1952 (US Federal Trade Commission, 1952). After the 1940s, when such behaviour became unacceptable, and up to the early 1970s, horizontal integration was maintained by the links between the major oil companies through their joint ownership of the Middle Eastern operating companies (**Penrose, 1965**). These joint ventures solved the problem of uncertainty over rival behaviour in an oligopolistic world. At the same time, it successfully controlled the excess capacity to produce crude oil present in the Middle East. Since this arose because of shared information (effectively interlocking directorships), no anti-trust legislation was being breached. The result was an orchestration of supply which protected price.

When the host governments nationalized the operating companies in the first half of the 1970s, horizontal integration through OPEC replaced that of the companies, an issue discussed below in Section 4.1.2. The contrast between the oil majors of the 1950s and 1960s and OPEC provides an interesting case study of cartel-type behaviour. The joint ownership of the Gulf producing companies by the majors provided excellent information about each company's behaviour and intentions. In addition, each of the producing joint ventures had lifting agreements which enabled partners to penalize those wishing to expand their market share. The mechanisms effectively could both detect and deter cheating – the classic problem of the cartel literature. OPEC by contrast could do neither for any length of time.

Much of the literature on reform of gas and electricity markets has addressed issues of cartelization and market control.[21] This inevitably follows from the obvious point that most gas or electric utilities were state monopolies. Hence the incumbent or bits of a divested incumbent had a strong interest in restricting competition. The purpose of regulation and competition was among other things to try and prevent the use of horizontal integration to limit competition.

Energy industries in general have often been characterized by vertical integration. The economics literature does not make this explicit, but vertical integration can take two forms. Financial vertical integration is when different stages in the same value chain are owned by one holding company. The crude-producing affiliate, the refinery and the marketing network are all owned by the same company or, alternatively, the electricity generator owns the transmission network and the local distributor. Hence the holding company effectively controls the cash flows of the affiliates. Operational vertical integration by contrast is when the owned products (oil, kilowatts or whatever) are moved between these affiliates. Operational vertical integration obviously requires the presence of financial vertical integration but the reverse is not true and markets can replace operational vertical integration. Much of the recent work on gas and electricity reform concerns how those elements of a vertically integrated chain with the characteristics of natural monopoly – the transmission and grid systems – are managed in a deregulated market.[22]

The difference between financial and operational vertical integration for the oil industry has been significant. The major private oil companies, before the second oil shock of 1978–81, were both financially and operationally vertically integrated (De Chazeau and Khan, 1959).[23] Several factors explain. Transactions and information costs made operational vertical integration superior to markets which were non-existent or highly imperfect. Operational vertical integration also had the benefit for the companies of inhibiting competition. In theory at least, an operationally vertically integrated oil company provides significant barriers to entry. If the companies only exchange crude between their affiliates, there is no access to crude for third parties. Also, it is possible to practise price discrimination by integrating into the low-priced market preventing arbitrage. Finally, operational vertical integration enabled the companies to play tax games through the use of transfer prices to minimize their global tax bill (Turner, 1978).

After the second oil shock, the private companies moved away from operational vertical integration, preferring instead to use markets. However, the national oil companies which had developed a financial vertically integated capability used operational vertical integration (Stevens, 1998b). The increasing use of markets reflected several factors. The nationalizations of the 1970s plus the discrediting of long-term contracts (Hartshorn, 1980) increased the number of arm's-length transactions, which meant both a greater number of buyers and sellers plus greater market transparency. The consequent lowering of transactions costs encouraged further use of markets, which created a self-feeding process of more players and transparency. Barriers to entry weakened as new non-integrated crude producers entered the market and as the majors began to sell off refineries to smaller petropreneurs (Bleakley et al., 1997). In such a world, constraint of competition became increasingly less relevant. Finally, the tax authorities began to constrain transfer pricing games. Hence operational vertical integration among the private companies, except in certain specific cases, disappeared. Most recently the issue under discussion (although too recent for the academic literature, which tends to lag behind) is the value of financial integration for private oil companies given the continued poor profitability of the refinery sector (Horsnell, 1997; Stevens, 1999b).

The national oil companies which invested downstream preferred to use operational vertical integration. The official reason for this was to lock-in market share but a more plausible explanation was to deepen the information asymmetries at the heart of the principal-agent relationship, thereby enabling greater rent capture. It is this which has prompted an increasing number of producer governments to begin the process of scrutinizing the behaviour of their national oil companies (Stevens, 1999a).

As with horizontal integration, the reform of gas and electricity utilities in many countries has also raised issues of vertical integration, following the ridiculous way in which British Gas was privatized as a vertically integrated monopolist and monopsonist. As indicated this literature is excluded from these volumes.

A third area of the literature on the oil companies examines their general strategy and behaviour. There is quite a large literature on the history of oil companies (for example for BP see Ferrier, 1982, and Bamberg, 1994; for Shell see Howarth, 1997; for a more general history plus a good bibliography see Yergin, 1991). While some of these are 'official' histories which lack critical analysis and tend to apologia, some do tell it 'warts and all' (Bamberg, forthcoming) and as a result become an excellent source for research. Of especial interest has been the tendency of companies to follow a strong herd instinct (Ollinger, 1994; **Lynch, 1995**; Mitchell, 1996), an issue which is beginning to attract attention in more mainstream economics (Lux, 1995). This herd behaviour has been a noticeable feature of the industry as the major private companies continue to dominate much of the industry. It has also been an important factor behind the tendency of the oil industry to experience feast or famine referred to earlier. Again, too recent for the standard academic literature, work is emerging on the behaviour of national oil companies, especially in the context of theories of public choice (Mommer, 1998; Stevens, 1999a).

4.1.2 OPEC

The first oil shock triggered a huge literature aimed at understanding oil markets, much of which focused on the role of OPEC. An excellent summary of the subject can be found in **Gately, 1984**; Cremer and Salehi-Isfahani, 1991; and **Salehi-Isfahani, 1995**.

One set of studies saw OPEC as a monolithic cartel which sets price with no competition among its members acting as the residual supplier to balance the market. Gilbert, 1978, identified OPEC as a Stackelberg leader which maximizes profit by choosing an optimal production path, taking into consideration the reaction of the fringe to its policies. Pindyck, 1978a, used an inter-temporal model to derive the price that would maximize the sum of the discounted profits of the cartel, taking into account depletion, reserve levels and production costs. Salant, 1976, assumed that OPEC takes the sales path of the fringe as a given and maximizes its joint discounted profits. Hence the cartel takes account only of the response of consumers to its policies.

Other studies moved away from the idea of OPEC as a monolithic entity, as the obvious divisions within the organization undermined the credibility of the assumption. Hnyilicza and Pindyck, 1976, and Pindyck, 1977, 1978a and 1979, considered OPEC as a two-part cartel divided into 'savers' and 'spenders'. The key difference between the two lay in their respective discount rates. The model implied that the 'spenders' would supply first because of their higher discount rate, leaving the 'savers' to supply last. They use Nash's 'theory of cooperative games' to simulate interactions between the two groups. A variant on examining internal OPEC behaviour is provided by Eckbo, 1976. He divides OPEC into three groups. The first – the hard core – could expand production substantially. The second – the price pushers – produce close to potential and have a strong need for current income. The last group – the expansionist fringe – have smaller reserves than the core and a strong need for current income, but produces more slowly than the price pushers. He then tries to construct an optimal price path for OPEC acting as a rational monopolist. His conclusion is that there is substantial divergence between preferred prices.

A number of other studies tried to develop this approach further (Kalymon, 1975; Ben-Shahar, 1976; Singer, 1978; Willett, 1979; Marshalla and Nesbitt, 1986). For example, Geroski, Ulph and Ulph, 1987, described OPEC as a cartel where members' supply varied over time in response to previous data and the cooperation of other producers. Supply also varies according to their attitude to other's cheating and the relative weight on long-run versus short-run profits as determined by their financial needs. OPEC was divided into four groups – fringe, high absorbers, low absorbers and Saudi Arabia – and the model was estimated by using quarterly data for 1966–81. Al-Roomy, 1987, tried to extend this model to the whole world oil market using monthly data from 1974 to 1984, and concluded that price movements cannot be explained purely by conventional supply and demand analysis.

Griffen and Teece, 1982, emphasized the key role of Saudi Arabia as the swing producer within the cartel. In similar vein Adelman, 1976, 1978, 1982, 1986b and 1995, described OPEC as a loosely cooperating oligopoly-cartel or a residual-firm monopolist within which Saudi Arabia acted as a swing producer. Mabro, 1992, adopted a Stackelberg approach but with Saudi Arabia as the price leader rather than OPEC. Many other studies have also adopted a similar approach. Cremer and Weitzman, 1976, used it to develop a simulation model, while others developed the theory (Gilbert, 1978; Newbery, 1981; Ulph and Folie, 1981; Hansen, Epple and William, 1985).

Other studies of the oil market in response to the outpourings on OPEC were based upon a competitive model rather than the cartel version. MacAvoy, 1982, focused on supply and demand interaction, with price increases being attributable mainly to supply disruptions. In his model he simulated equilibrium prices under a number of assumptions using actual reserves, income and some assumed elasticities. Also in a competitive context, some (Johany, 1980; Mead, 1979) used the change in property rights and its subsequent impact on producers' discount rates to explain the supply responses which allegedly triggered the first oil shock. Hence the companies threatened with takeover had high discount rates prompting high production levels. When the governments took over control of the production decision, discount rates fell and so too did production as a consequence.

Roumasset, Isaak and Fesharaki, 1983, attempted to explain the price path since 1973 in terms of deviations from efficient prices which would have ruled absent monopoly profits. As a result they argue that the first two oil shocks can be explained in terms of a response to supply-side changes, especially changes in the perceived cost of the backstop technology and the fall in real interest rates in the mid to late 1970s.

Another thread in the oil market literature explains OPEC production behaviour in terms of a capacity-utilization model. In this literature (Gately, 1984; Powell, 1990; Porter, 1992; Suranovic, 1993; Gately, 1995) the basic assumption is that OPEC attempts to maintain capacity utilization near a specific target level. If production rises, causing the target to be over-fulfilled due to higher demand, then prices rise which will reduce demand and restore capacity utilization to its target level. If capacity utilization is below target, OPEC reduces price to stimulate demand. However, while some empirical support for the hypothesis was found before 1985, thereafter the relationship no longer appeared to hold empirically.

In addition to studies based upon explicit economic theory, there are also a number of studies more in the tradition of political economy which try to explain OPEC supply, especially from Saudi Arabia. Doran, 1977, examined differences in pricing strategies between OPEC members based upon the size of their reserves and hence differences in terms of time horizons as oil producers. Stevens, 1982a and 1982b, examined the specific role of Saudi Arabia and concluded that its supply function was the result of its role as residual supplier driven by a number of complex economic and political considerations. Moran, 1982, examined Saudi Arabian policy and concluded that it was driven by politics rather than economic optimization. Quandt, 1982, examined production decisions in Saudi Arabia, emphasizing the nature of the decision-making process, which, he argued, prevented the Kingdom from pursuing a consistent long-term policy. Golub, 1985, tried to differentiate in Saudi policy between crisis and routine periods. In the former he argued political drivers dominated the decision process but in the latter the profit motive began to play a role. Askari, 1991, also analysed the role of Saudi Arabia between 1973 and 1981, describing its supply decisions as being based upon trying to reconcile twin desires, to support OPEC and at the same time prevent oil prices rising to a level which would do damage to the world economy.

There have been attempts to test empirically many of the hypotheses associated with OPEC's supply behaviour. Griffin, 1985, tested several hypotheses. He tested a competitive model and concluded that price – exogenously determined – did influence the production decision of Saudi Arabia, implying a positive and significant price elasticity of supply, but the model, based upon ordinary least squares, had no dynamic properties. He also tested for the property-rights thesis to see if production was influenced by how much it was controlled by government, but failed to find empirical support during the period 1971–81. The target-revenue model was also tested using a variable for target investment. The results were ambivalent. He also tested a cartel model and found empirical support for the view of dominant firm models with Saudi Arabia acting as the market leader and varying production inversely to the competitive output including the rest of OPEC.

Griffin's work attracted considerable criticism. Salehi-Isfahani, 1987, argued the equations were mis-specified for the target revenue model. In particular, he was critical of the absence of price expectations and included a variable for expectations. Using the same data as Griffin, he established support for the hypothesis that production was restrained when prices reached certain levels. Cremer and Salehi-Isfahani, 1991, also criticized Griffin for neglecting dynamic forces apparent in the presence of acute serial correlation and for neglecting the depletable nature of oil.

Dahl and Yucel, 1991, tested variants of the target-revenue model using quarterly data for Saudi Arabia from 1971 to 1987 and rejected the hypothesis. They also tested for the swing-producer model based upon output coordination between individual members of OPEC and total OPEC production and rejected the hypothesis. The swing model was also tested by Griffin and Nielson, 1994, focusing on strategies used by OPEC to maximize cartel profits covering the period 1983–90. Support was found only for the period 1983–85. Thereafter they concluded Saudi Arabia adopted a 'tit-for-tat' policy aimed at punishing excessive cheating by other OPEC members. Gulen, 1996, used monthly data for OPEC from 1965 to 1993 to assess the success of OPEC in coordinating output after formally adopting a quota system in 1982. His tests supported the existence of coordination and of Saudi Arabia's swing role between 1982 and 1985. Other recent empirical tests can be found in Al-Turki, 1994, and Al Yousef, 1998.

In recent years, work of the type described above has been superseded by work which tries to account for the changes in the role of OPEC as a result of the third oil shock of 1986. In particular, a number of studies have attempted to consider the future of oil supplies by use of what might loosely be described as scenario analysis incorporating political as well as technical considerations. These studies have been produced by consultancy firms such as CERA and DRI. They have also been produced by various international and government bodies including the IEA's World Energy Outlook in various years, estimates from the OPEC Secretariat, and forecasts produced by the US Department of Energy. There have also been innumerable such studies from individuals (Curcio, 1989; Lynch, 1990; Porter, 1990; Seymour, 1990; Odell, 1994; Mitchell, 1994a; O'Dell, 1994; Miremadi and Ismail, 1994; Ismail, 1995; Knapp, 1995; Stevens, 1996; Pauwels, 1997; Campbell, 1996). In general, the studies tend to fall into two camps: those that see non-OPEC facing growing constraints, forcing ever-increasing dependence on the Gulf, and those that argue non-OPEC still has much more to produce.

4.1.3 OIL COMPANIES AND PRODUCER GOVERNMENTS

Several areas in the energy economics literature have been concerned with the relationship between oil companies and governments. The first area concerns fiscal issues, already dealt with in Section 3.3.

The second area concerns the political economy of interactions between oil companies and governments. The size of the major private oil companies is so large that it raises many issues. At a national level, probably the relationship within the US has attracted most attention in the literature (Sampson, 1975; Blair, 1976; Goodwin, 1981; Yergin, 1991; Bradley, 1996).[24] This domination reflects the US system of Congressional hearings and freedom of information which has forced into the open a great deal of material, making research feasible. Furthermore, the oil companies in the US have always attracted a great deal of public concern and suspicion ever since the days of Rockefeller and the Standard Oil Trust. The activities of the oil companies remain at the forefront of policy debate. For example in the 1970s there was a concerted move to force divestiture on vertically integrated US oil companies (US Treasury Department, 1976).

The oil industry has also provided many case studies for the general investigation of multinational corporations, especially in developing countries (Penrose, 1968; Tanzer, 1969; Das Gupta, 1971; Mikesell, 1971). More specifically, the issue of company relations with producer governments has attracted a fair amount of attention in the literature over the years (Hartshorn, 1967; Mughraby, 1966; Stocking, 1970; Stevens, 1975; Khan, 1987). Much of this literature has been concerned with the relative control of companies by governments. Another highly specialized sub-set of the literature on company–government relations has dealt with the aftermath of a breakdown in such relations resulting in nationalization, and how compensation might be determined. Most of this literature tends to be written from a legal perspective but a useful summary of the economic issues, albeit eventually favouring a particular viewpoint, can be found in Penrose et al., 1992.

The whole debate about liberalization and deregulation which began in the 1980s has also opened up a huge literature on company–government relations with respect to both gas and electricity. However, as indicated earlier, this topic is too large to be effectively included in these volumes.

4.2 *Markets and Prices*

Two areas have dominated the literature on energy pricing. The first concerns oil pricing. This is hardly surprising given that the oil price shocks of the 1970s moved energy economics up the economists' agenda. The second concerns the domestic pricing of energy in a more general context of energy policy.

4.2.1 OIL PRICES

The literature on oil prices contains a number of strands. The principal one concerns the determination of oil prices, and in turn has three sub-strands. The first concerns narratives of what happened. This literature (Sampson, 1975; Seymour, 1980; **Mitchell, 1982**; El-Mokadem et al., 1984; **Adelman, 1986b**; Mabro, 1988; Yergin, 1991; Adelman, 1995) tries to describe both the economic and the political context which gave rise to the oil shocks. It is essentially story telling with a beginning, a middle and an end, describing who did what to whom. However, this description does not do credit to the importance of such literature. As indicated earlier in the discussion of Hotelling's work, the facts of what happened are absolutely crucial. Without the facts, economists are likely to produce most unhelpful analysis.

The second sub-strand of this explanatory literature attempts to provide explanations based upon economic theory, often involving modelling. Much of this literature has already been discussed in Section 4.1.2 and revolves around discussions of OPEC's role or lack of role. The final sub-strand is an attempt to combine both approaches. Although the framework for analysis is taken from economics, it also supports a narrative of what actually happens or happened with respect to oil prices (**Mabro, 1992**; **Stevens, 1995a**; Stevens, 1996; De Oliveira and Guimaraes Lodi, 1994; Giraud, 1995; Horsnell and Mabro, 1993).

The second main strand of the literature on oil prices has concerned the development of commodity trading of oil, the development of paper barrel markets and the consequent price volatility. Until the second oil shock of 1978–81, there were virtually no paper markets for oil of the type which had been present in other commodity markets for decades or even centuries.[35] The explanation for the lack of such markets was simple. Their development had been triggered by the need to hedge price risk by those interested in supplying or using the physical commodity. Once started, such markets inevitably attracted speculators interested only in paper rather than physical commodities. Following the 'As Is' agreement of 1928 and up to the second oil shock, oil price risk was minimal. Before 1949, the price was set by the Gulf Basing Point System (Penrose, 1968). All products were sold internationally at a price set by the US domestic price in the Gulf of Mexico with the addition of a phantom freight rate to equalize the landed price irrespective of the origin of the oil. This was modified in 1944 to include a second basing point at the Abadan refinery in the Persian Gulf, and effectively opened up the world oil market to low-cost Middle East oil. After 1949, the oil price was based upon posted prices which were administered prices. The oil companies set the posted price for tax purposes and, provided the buyers of crude found the price plausible, that became the price. During the 1950s and 1960s, operational vertical integration meant those setting the price and those required to believe in its plausibility were the same. Not surprisingly this created significant price stability. Hence the very significant economics literature on commodity markets (Greenaway and Morgan, 1999) effectively ignored oil.[26]

All this changed with the second oil shock. Long term contracts had been discredited – destroyed by the invoking of *force majeure* clauses. Oil markets now relied increasingly on short term (i.e. quarterly) contracts or spot transactions. The spot price became the benchmark by which prices were determined. OPEC tried to maintain an administered price system but this became untenable and the oil price collapse of 1986 led them to move to spot-related prices. Hence prices became more volatile (Plourde and Watkins, 1994). Initially informal forward markets appeared to hedge price risk. These were gradually supplemented by formal futures markets carrying all the paraphernalia of derivative paper markets.

The literature followed these trends with great interest.[27] Some tried simply to describe the various mechanisms of paper trading in the context of oil (Horsnell, 1997; Roeber, 1993; Verleger, 1987). After all, this was completely new territory. Others attempted to assess the relationship between futures prices and forecasting future prices (Moosa and Loughani, 1994; Serletis, 1994; Balabanoff, 1995a) asking if they were good guides. Finally, as concern grew about increasing price volatility, research began to consider the extent of this volatility (Plourde and Watkins, 1994).

The third strand of the literature on oil pricing is an attempt to analyse price forecasting. The record of oil-price forecasters, even by the very poor record of economic forecasting generally, has been absolutely abysmal. The classic example comes from the IASA annual forecasts reproduced in Hartshorn, 1993. These forecasts year on year are forced ever lower as the actual crude price declines, although in honour of Hotelling they retain an unbelievably stubborn upward trend! This abysmal performance has led to much navel gazing in an attempt to understand what went wrong (Huntington, 1994). These post mortem exercises should be required reading for all students of economics as they provide stunning examples of how not to do economics.

A fourth strand in the oil price literature considers the relationship between fluctuations in crude prices and changes to product prices. Much of this concern was triggered by the perception amongst consumers that oil companies were price gouging. This is an accusation which the literature seems reluctant to confirm (Hagen, 1994; Shin, 1994). The final strand has been to evaluate the impact of extreme price fluctuations on both oil importers and oil exporters. This literature has already been discussed in Section 1.

4.2.2 DOMESTIC ENERGY PRICES

The literature on the economics of domestic energy prices addresses the key issues of how prices should be set and at what level. Given energy's pervasive importance in the economy, its pricing is an issue of no small importance. In particular, there is the constant underlying debate of micro economics between the needs for economic efficiency and for equity, to say nothing of other perfectly legitimate objectives of energy policy.

Energy prices differ in type, depending upon who is paying the price and for what reason. The following classification (Bhatia, 1985) illustrates the point:

- Resource prices accrue to the supplier for supplying the energy. Ultimately, it is the net resource price (after taxes and costs have been met) which is important in determining the supplier's willingness and ability to supply.

- Transfer prices are prices charged between affiliates in the value chain. These are most obviously associated with the energy transformation process from primary to useable energy where the primary energy supplier represents the same parent company as the energy transformer.

- Output prices (sometimes called producer prices) are the prices charged to those who use energy as an intermediate good in some form of productive process. Thus these energy prices will direct the choice of techniques and possibly the amount of the good produced depending upon the relative importance of energy costs in total costs.

- Lastly come the energy prices paid by the end-use consumers (consumer prices), defined as those who use energy as a final input to provide energy services. The distinction from producers can best be understood by assuming that the consumers are in fact households and therefore require the sorts of energy services required by households – transport, lighting, space and water heating, cooking, and the use of other electrical appliances such as TV.

The distinction between these prices can be extremely important. For example, efficiency considerations should be paramount in setting output (producer) prices to minimize distortions in the choice of technique and the choice of output. However, in the setting of consumer prices it may be that equity considerations require greater emphasis.

There are many different views on how energy prices should be set (**Munasinghe, 1980a**; Brown et al., 1988; Siddayao, 1985; Bhattacharyya, 1995a and 1995b; Pagá and Birol, 1995) but the following approach has merit. It involves a two-stage process. In the first stage, the price required to achieve economic efficiency is established. Once established, this price can then be adjusted if required to take account of the other objectives of energy policy. The main virtue of this approach is that it makes explicit the cost of deviating from the efficiency objective.

However, the literature examines a great many problems in the setting of efficient prices at both a theoretical and a practical level. Energy supply systems involve large capital-intensive projects which are characterized by long lead times and long lives. The consequence of this is that in many cases a very large difference exists between short- and long-run marginal costs. The question therefore arises, which should be used as the benchmark for efficient prices? (Schramm, 1985). There are also problems of second-best pricing, especially relevant because of the central role of energy prices in an economic system (Newbery, 1985; Dixit and Newbery, 1985).

Efficient prices may require different prices for the same fuel in different uses and between different consumers. The most obvious distinction is between producer and consumer prices outlined above. As indicated, efficiency may require producer prices to be set at marginal costs to avoid distortions, while equity may require lower prices for consumers. Different uses of the same fuel may also require different efficient prices. For example, diesel used in industry and agriculture might be priced at marginal cost for efficiency, while diesel used in transport might be priced higher as a natural way to recover the cost of road usage. The problem is how to avoid arbitrage. If diesel is sold to farmers at a lower price than to hauliers, there is a strong temptation for farmers to resell to hauliers. Kerosene provides the classic example. It can be used by households for lighting and cooking, but up to 30 percent of diesel in transport can be replaced by kerosene and the vehicles will run quite satisfactorily.[28] Thus subsidies on kerosene invariably lead to diesel adulteration.

Where one product is produced jointly with other products, allocating costs to determine marginal costs can be controversial. In energy, the most obvious case where this creates problems is in pricing oil products which are the product of a refining process. Another example is how to price hydroelectricity when the dam also provides for irrigation and flood control.

Finally, a huge literature has recently grown up around the problem of the extent to which energy prices should be adjusted to account for environmental externalities. However, as explained earlier this is outside the remit of these volumes.

Once it has been established what the efficient price for the energy should be, the next stage is to decide what changes, if any, should be made to accommodate other energy policy objectives, and, by implication, to determine the economic cost of seeking to pursue these other objectives. This has received much attention in the literature.

One reason for adjustment is to improve equity. It is well established that the poor spend proportionately more on energy than the rich. In India, for example, in a survey on rural spending in 1973–74, it was established that the lowest three income groups spent 9 percent of their income on fuel and lighting compared to 4 percent for the three highest income groups (Bhatia, 1985). In Niger, it has been estimated that the average manual labourer's family spends a quarter of its income on firewood, while in Upper Volta (now Burkina Faso) the percentage is 20–30 percent (Eckholm in Smil and Knowland, 1980). Unfortunately, the literature also suggests that using subsidized energy prices to help the poor in the Third World has been generally disappointing and often counter-productive (Bhatia, 1985; Kosmo, 1987; Thukral and Bhandari, 1994). Frequently, the poor have to pay much higher prices than the official price as black markets develop. Consumers may be ignorant about what prices should be, and often buy smaller amounts of the fuel than those specified for the official price. Over-pricing can also occur where the official price is set leaving distributors to add their own margins.

Evidence is growing that the benefits from subsidized energy prices accrue to the rich. The first point is that the really poor tend not to consume commercial energy (Pitt, 1985). Thus subsidizing kerosene assists only those who consume kerosene, who in many cases are not the really poor. Secondly, there is the problem of substitution mentioned earlier. Thus kerosene subsidies may well secure benefits for the relatively rich hauliers. Finally, there is the issue of how any subsidy is paid for. If it is met by taxes this can be regressive. The poor may well be forced to pay taxes but actually consume relatively little if any commercial energy. This is especially relevant in the Third World where problems with income tax collection encourage a greater dependence upon sales taxes which are regressive in nature. If the subsidy is met by means which generate inflation, the evidence suggests that inflation is generally regressive.

Another reason to adjust the efficient price is to use energy prices to raise government revenue (Gupta and Mahler, 1995). As indicated earlier, energy's widespread use gives it a large tax base, and the inelasticity of demand means it can carry a high tax rate. Coupled with low collection costs, this makes taxing energy extremely attractive. Furthermore, revenue greed by politicians can be wrapped up in environmental concerns. The practice of imposing very high taxes on certain oil products in the OECD has attracted considerable criticism especially from OPEC. Some of this criticism is just bad economics. For example, claiming unfairness because the value added on oil sales accruing to producers' governments is much less than the sales taxes (i.e. transfer payments) accruing to consumer governments, is comparing apples and oranges. However, there are more legitimate concerns, not least those related to effective trade restrictions (Seymour and Mabro, 1994).

A separate strand in the literature on domestic energy prices has emerged in the 1990s, as governments and their regulators have tried to come to terms with the behaviour of deregulated and privatized utilities in gas and electricity. For example the Electricity Pool Pricing System developed in the UK has received considerable attention (Rudkevich et al., 1998; Tilley and Weyman Jones, 1997).

4.3 Markets Versus Governments

4.3.1 ENERGY POLICY VERSUS THE MARKETS

Much of the study of energy economics has examined energy policy in terms of objectives and policy instruments, and their effectiveness. A central issue in the literature has been the merits of energy policy versus the market (**Helm, Kay and Thompson, 1988; Robinson, 1993; Newbery, 1996**). One school of thought, popular in recent years, argues that government intervention in the energy sector has generally been counter-productive and has inhibited efficient operation. This view is based on two main arguments. First there is the negative argument, based upon theories of public choice and economic theories of politics, that government intervention in any sector is bound to produce sub-optimal results. Second, there is the more positive argument that market forces through a process of discovery produce significantly better results than regulation. Both arguments are essentially unrelated to energy. It is simply the traditionally heavy state involvement in the energy sector which has made it an attractive target for debate.

The counter-arguments for government intervention do have a more specific energy dimension. The traditional argument for government intervention was based upon the existence of market failure discussed in Section 1. This occurred in the presence of imperfect competition, externalities and public goods. Arguably, elements of these aspects of market failure are more likely to be present in the energy sector than other sectors of the economy.

In part as a spin-off from the privatization debate, competition in energy markets has received much attention. Several reasons suggest why competition in energy may be restrained. Many existing energy companies were or are state-owned monopolies. As such, they can prevent entry by control over the licence to operate. Even if this artificial entry barrier is removed, the existing incumbents retain market power which translates into real entry barriers. It is possible by means of regulation to try to prevent such barriers from being used and to dismantle them. However, this takes time. The scale of operation in many cases can inhibit entry. An obvious example is oil refining, in which the minimum effective plant is extremely large, both in physical terms for primary distillation, and also in capital terms given the huge costs involved in upgrading plant. This can change with technology. For example, the development of combined-cycle gas generation has significantly reduced the optimum scale for a generating plant together with its capital cost. Also, for tradeable energy, providing free trade is allowed, scale ceases to be a barrier.

Any grid-based system has characteristics of natural monopoly. Gas and electricity are delivered by means of grid networks. Obviously, minimizing cost requires that only one pipe or wire delivers to the consumer. This means that there should be only one supplier. To have competing grids would be a waste of resources, and the incumbent grid owner would in any case be able to exclude new competition unless technology provided some alternative format for delivery. The example of cellular phones versus fixed wires is often quoted. However, it is difficult to envisage an alternative transmission system for gas or electricity, and for electricity at least there are many uses for which no other energy sources will do. Either the state takes control of the grid or, if it is privately owned, the grid requires regulation in some shape or form.

The existence of externalities also justifies intervention although arguably this is an issue of environmental rather than energy policy. Security of energy supply is sometimes used to justify intervention. If national benefits from a secure supply are understated, the market would under-provide. In terms of stockholding the argument loses some validity since it can be claimed that it is government intervention to prevent stockholders gaining windfall profits which inhibits stock holding. The use of diversity of sources to provide security has more validity, since promoting such diversity may require some form of government incentive, if only because of the loss of economies of scale available if a single energy source is used.

Another argument for intervention is that markets have too short a time horizon because their discount rate is too high. Thus governments should control the rate of depletion of oil, gas and coal enforcing a 'better' discount rate. This is a difficult argument. The empirical evidence on short-termism in financial markets remains ambivalent. However, the argument's main weakness is the assumption that governments do not also suffer from short-termism driven by the need to be re-elected. Other arguments for government intervention over conservation have already been discussed.

4.3.2 THE CHRONOLOGY OF ENERGY POLICY

The history of energy policy has a clearly defined chronology. Over time objectives and instruments have changed. Objectives have been very much subject to whim and are driven by popular perceptions current at the time. Equally, the availability and effectiveness of policy instruments has changed as what is technically and politically feasible alters. In general there is an extensive literature on energy policy for specific countries. For the US see Goodwin, 1981; Bradley, 1996; Kirby, 1995. For the EU see De Bauw et al., undated; Surrey, 1992; Lyons, 1994; Mitchell, 1994b; McGowan, 1996; Matlary, 1997. For the UK see Chesshire, 1982; Robinson, 1992 and 1993; Mackerron and Pearson, 1996. For Japan see Perkins, 1994. For China see Zhang, 1995; Wu and Binsheng, 1995; Wang, 1995. For Pakistan see Looney, 1995. For the SADCC region of Africa see O'Keefe and Munslow, 1984.

In the 1950s and 1960s, there was no such animal as energy policy (Posner, 1973, Robinson, 1993). At best, some countries had policies directed at energy sub-sectors. For example in the UK considerable attention was directed to both coal and nuclear. However, there were no attempts at coordination. Given the ideological philosophies of the time this is rather surprising. Government intervention was the norm in an economy driven by market failure concerns, the legacy of Keynes and the apparent success of Soviet-style planning. The explanation for this *laissez faire* approach is simple. During the period, as discussed in Section 1, energy was perceived to be plentiful. The post-war OECD economic miracle was fuelled largely by growing consumption of oil, which over the period experienced falling prices in both nominal and real terms. Hence in the OECD, the general view was that the oil companies could be left to continue to supply low-cost oil. The only issue which attracted some attention was the use of transfer pricing to minimize taxes paid (Turner, 1978). Developing countries were rather more pro-active with respect to foreign oil companies (Tanzer, 1969; Das Gupta, 1971) both as producing countries, as discussed in Section 4.1.3, and also as oil importers. However, this was an energy sub-sector issue rather than an energy policy.

All this changed with the oil price shocks of the 1970s. Public concern was triggered initially by gasoline queues and power shortages.[29] Concern deepened as ideas about energy 'running out' discussed in Section 3.1 began to take hold. Such public concern obliged governments to 'do something about energy'. This created a problem in most countries since the complacency of the previous twenty years meant that there were few within government who understood energy issues. Knowledgeable people had to be drafted in from the energy industries, with the result that energy policy had a very strong supply-side orientation, a characteristic of policy which in most countries has endured. While this enabled the OECD countries to begin the process of creating and implementing an energy policy (Perkins, 1994), most developing countries lacked even that option. Generally the prime objective of energy policy revolved around security of supply and reducing import dependence as discussed in Section 2.3.3. This meant reducing consumption[30] and encouraging the development of indigenous fuel supplies through processes of direct investment or subsidy. However, concern about export competitiveness coupled with a Keynesian imperative to counter the recession which followed the first oil shock meant energy pricing as a policy instrument was not fully used. In many cases consumers were actually cushioned from the impact of the higher energy prices.[31] Hence most of the policy instruments were linked to direct investments or regulation.

During the 1980s, after the first impacts of the second oil shock began to be muted, policy objectives began to shift. In the OECD countries, supply security began to be less of an imperative. Solutions to the 'energy problem' were seen to lie more in the diversification of energy sources and the use of pricing mechanisms. Furthermore the failure of many of the 'alternative' energy programmes[32] coupled with the general ideological shift away from government intervention embodied in Reaganomics and Thatcherism began, at least in some quarters, to discredit the notion of a strongly pro-active energy policy. In the developing countries, by contrast, supply security remained high on the agenda. In some cases problems

جدول ۳-۱ مولفه، زیر مولفه و شاخص های پرسشنامه

ردیف	مولفه	زیر مولفه	شاخص	منابع
۱	عدم قطعیت ادراکی محیط بازار	آشفتگی بازار	تا چه میزان شرایط بازار و صنعتی که در آن فعالیت می کنید را پیچیده و پویا می بینید	Narver& Slater (1990); Kohli&Jaworski 1990); Jaworski&Kohli 1993); Rose &Shoham (2002); Voss & Voss (2000); Slater&Narver (2000); Matsuno ,Mentzer&Rentz (2005)
۲			تا چه میزان شرایط بازار و صعنت را قابل پیش بینی و برنامه ریزی می دانید؟	
۳		شدت رقابت	به نظر شما تغییرات محیطی بازار دارای روند خاصی بوده و بر اساس اطلاعات گذشته می توان پیش بینی انجام داد؟	
٤			تا چه میزان رقبا می توانند به سادگی وارد صنعتی که شرکت شما در آن فعال است، شوند؟	
٥			آیا در صنعتی که شرکت شما در آن فعال است روند رشد عرضه از تقاضا بیشتر بوده و آیا بازار آن رقابتی است؟	
٦	مدیریت ارشد	ریسک گریزی مدیریت ارشد	تا چه میزان هیئت مدیره شرکت تمایل دارند که شرایط مهم و پر ریسک بازار و صنعت را بپذیرند؟	Kohli&Jaworski (1990); Ruekert, R. (1992); Pelham, A. M., & Wilson, D. T. (1996); O'Cass A, Weerwardena (2010)
۷			تا چه میزان تصمیمات و مصوبات مدیران شرکت در راستای رودررویی با ریسک های ذاتی بازار است؟	
۸			تا چه میزان شرایط پیچیده و غیرقابل پیش بینی محیط بازار برای مدیران شرکت قابل درک است؟	
۹	بازارگرایی	ایجاد هوشمندی	تا چه میزان شرکتدیدگاه‌ها و نظرات مشتریان را درارتباط بامحصول وخدمات بررسی می کند؟	Kohli&Jaworski (1990) Narver& Slater (1990); Kohli,Jaworski& Kumar (1993);Slater &Narver(1993); Slater&Narver(1994); Slater&Narver(1995); Matsuno&Mentzer
۱۰			تا چه میزان از سیستم های دریافت و تحلیل اطلاعات اولیه مشتریان استفاده می شود؟	
۱۱			تا چه اندازه در شرکت ترجیحات و تغییرات رفتاری	

argue strongly that biomass has a significant role to play in a modern energy context (**Hall, 1991**). Many studies have confirmed this link with poverty (Leach, 1992; Jones, 1991). Leach for example points out the very strong inverse relationship between the proportion of biomass in total primary energy use with a country's status in the league tables of per capita income.

Numerous issues are associated with traditional energy. There are the problems associated with fuel poverty and their possible solutions (Douglas, 1982; Foley and Van Buren, 1982; Agarwal, 1986; Leach and Mearns, 1988; Stevens, 1988; Pearson and Stevens, 1989; Soussan, J. et al., 1990; Hosier and Kipondya, 1993; World Bank, 1996). Linked to this has been debate about energy transitions toward greater use of commercial fuels (Pearson, 1988; Leach, 1992; Murray and de Montalembert, 1992). Local environmental consequences have also attracted considerable attention in the literature (Anderson, 1987; Smith, 1987; Pearce et al., 1989; Pearce and Turner, 1990; Deshingkar, 1990; Ellegard and Egneus, 1993). Implications for women and children in a context of growing fuel poverty have also attracted interest (Tinker, 1987; Munslow et al., 1988; Dankelman and Davidson, 1988; Mercer and Soussan, 1991).

Given the tendency of energy studies to 'get things wrong', the 'woodfuel crisis' of the early 1980s is particularly notable. It was a classic example of gapology (discussed in Section 2.3.3). The origins of the crisis (elegantly explained by Leach and Mearns, 1988) lay in the aggregation of a series of assumptions. Per capita wood consumption in a country was extrapolated on the basis of population growth. The resulting rising trend line was linked to assumptions about the stock of trees on the same graph and, hardly surprisingly, a gap emerged. This in turn prompted all sorts of policy recommendations to 'close the gap'. Top of the policy solutions was simply to grow more trees, a solution described somewhat facetiously by Munslow et al., 1988, as the 'fuelwood trap'. Eventually, the energy economists mounted a convincing attack upon 'woodfuel crisis' thinking. First, the actual assumptions were challenged (Leach and Mearns, 1988). Much deforestation had little to do with fuel use but arose from land clearance for agriculture and commercial logging (Allen and Barnes, 1985). Projecting woodfuel demand ignoring relative fuel prices was virtually meaningless. Finally, remote sensing technology began to reveal the extent to which traditional views about tree stocks were extremely wide of the mark. The second challenge related to simplistic solutions (O'Keefe and Munslow, 1989; Leach and Gowen, 1987; Leach and Mearns, 1988; Howes, 1991; Soussan et al., 1992). Tree growing programmes and woodstove programmes came under scrutiny and were generally found wanting (Foley et al., 1984; Gill, 1987). Concern about land management issues came to the fore, starting a debate about the possible benefits of the 'commodification' of fuelwood sources (Runge, 1986; Clarke and Shrestha, 1989a and 1989b; Teplitz-Sembitzky and Schramm, 1989; Pearson and Stevens, 1989; Pearson, 1991).

Another energy area which has attracted interest in developing countries concerns the electricity sector. There has been much discussion about the problems facing the sector and their possible solutions (Barnett, 1992 and 1993; Bhattacharyya, 1995a; De Oliveira and MacKerron, 1992; Munasinghe, 1992; Schramm, 1993a). In this debate, lines have been clearly drawn between a view that the problems have been driven from outside the countries, most obviously by the debt crisis, and the view that the fault lies in excessive government intervention in the sector.

An area which has attracted much attention within this context is rural electrification (**Pearce and Webb, 1987**; Munasinghe, 1987; **Munasinghe, 1988**; Barnes, 1988; Foley, 1989;

Schramm, 1993b). In the 1960s and 1970s, rural electrification carried a very high priority in many developing countries where the proportion of rural population with access to electricity was extremely small. For example in India in 1981 only 17 percent had access and in some countries of sub-Saharan Africa the figure was as low as 2 percent (Schramm, 1993b). The commonly held view was that access to electricity would actually create a boost for rural development and act as a leading development sector. However, during the 1980s disillusion began to set in as the costs of such programmes began to move outside the reach of governments and as their effectiveness began to be questioned. In particular, it was argued (Foley, 1989) that the provision of electricity in isolation would have little effect until rural income rose to a level sufficient to afford the appliances. In other words, electricity supply alone would not promote rural development. As Schramm expressed it '... even the far more modest net benefits identified ex post, compared with ex ante expectations, were subject to doubt and, in several important cases, strongly negative' (Schramm, 1993b, page 504).

Notes

1. There are relatively few general text books on energy economics and most appeared in the 1970s and early 1980s (Webb and Ricketts, 1980; Eden et al., 1981; Weyman Jones, 1986; Gordon et al. (eds), 1987). A good general collection – ten volumes – can also be found in Moroney, 1979–97. There are also some chapters on energy economics in general textbooks on natural resource economics, and, by contrast, a great many general books on oil. Some of the best over the years include Frankel, 1946; Penrose, 1968; Adelman, 1972; Odell, 1986; Masseron, 1990; Yergin, 1991; Hartshorn, 1993; Frank, 1994.
2. References in **bold** type are those contained in the two volumes.
3. This is not true with respect to combined-cycle gas turbine technology, which is transforming the electricity sector worldwide (Watson, 1997).
4. Anyone seriously interested in energy economics should acquaint themselves with some of these technicalities. Quite good simple introductions to the physics and engineering of energy can be found in Slesser, 1982, Allen, 1992, and Hill et al., 1995, although the latter is not recommended for its economics content. For an application of economic analysis in an explicitly scientific context see also Georgescu-Roegen, 1972. For oil and gas technicalities, see Stevens, 1988.
5. Most obviously hydro electricity and nuclear but increasingly other 'renewable' means of power generation – solar, wind, tidal and geothermal. Electricity generated by hydrocarbons or biomass is excluded to avoid double counting.
6. It is an interesting observation based upon casual questioning of energy consumers by the editor over many years that most consumers have no idea what is 'the price' of energy. In other words they do not know what it costs to pay for the energy to run a fridge, a television, a car, etc. Thus one of the most fundamental relationships in economics – that quantity demanded is a function of price – in the field of energy economics arguably has little or no practical meaning!
7. Why anyone should wish to buy an inefficient appliance will be discussed below under the heading of conservation.
8. The reason often given is that the inclusion of an appliance stock variable 'messes up the equation'. It does not seem to occur to the authors that this suggests that either the data are wrong or the equation is mis-specified.
9. For example in India the very widespread use of cattle dung as a fuel has serious implications for soil fertility and the need for artificial fertilizer.
10. Because much of this sort of forecasting was driven purely by commercial concerns, relatively little emerged into the academic literature.
11. This is reminiscent of the old economists' joke whereby a physicist, a chemist and an economist on a desert island grapple with the problem of how to survive faced with an enormous store of canned

40

[9]

Energy Demand Analysis in Developing Countries: A Review

*Ramesh Bhatia**

1. INTRODUCTION

Energy demand analysis is an important component of integrated energy planning and policy in developing countries. Planners and policy makers need to have a good understanding of the factors affecting growth and pattern of energy demand before they can proceed to make demand projections for the future. Given the capital intensity and long gestation periods of energy investments, supply bottlenecks, and adverse effects of energy shortages, detailed demand studies need to be undertaken at the aggregate and sectoral levels. An analysis of the influence of price and nonprice variables on energy demand would also be necessary for designing policy measures relating to energy conservation.

Energy demand analysis (EDA) involves (a) assembling and presenting a consistent set of data on consumption of various forms of energy, (b) estimating the level of shortage or unfulfilled demand at relevant price ranges, and (c) quantifying the relationship of energy demand with relevant economic and noneconomic variables such as income, population, prices of different energy sources and their substitutes, changes in technology, etc. In the context of developing countries, EDA becomes complex on account of the following factors:

(i) Apart from the compilation of data on commercial fuels, there is a need to estimate the consumption of traditional (or non-commercial) forms of energy (biomass as well as animate sources)

The Energy Journal, Vol. 8, Special L.D.C. Issue.

*Professor of Economics, Institute of Economic Growth, Delhi, India.

which account for a major share of total energy consumption, particularly in rural areas;

(ii) There are large segments of the population, particularly the rural poor, who are not in a position to cover their 'basic needs' by purchasing power and depend on 'collected fuels'; but over time, there is a shift from noncommercial to commercial forms of energy;

(iii) To the extent that consumption of energy in developing countries is constrained by available supplies, actual consumption does not reflect actual demand. This situation is particularly serious in cases in which supplies are significantly restricted due to balance-of-payments difficulties or government policies such as restrictions on import of fuels or equipment, underutilization of capacity and infrastructural bottlenecks that affect distribution;

(iv) Usually, a major share of the consumption of commercial fuels and electricity is accounted for by a few industrial enterprises, railways and public transport undertakings. This means that individual investment decisions by one or a few of these large-scale users can have significant impacts on energy availability and total consumption;

(v) It is difficult to assess the response to changes in fuel prices because user choices are also determined by a number of other factors, such as difficulties of obtaining complementary non-energy inputs, the absolute shortages of certain fuels, imperfect product and capital markets.

Thus, unlike the industrialized countries, analysis of energy demand in developing countries not only requires the assembly of a set of consistent data but gets complicated by the set of factors outlined above. In view of this, it becomes even more necessary for developing than industrial countries to study variations at a more disaggregated level, e.g., for major user sectors, for selected energy-intensive industries or for major end-uses, such as irrigation pumping, lighting, cooking, etc.

This paper provides an overview of energy demand analyses at the aggregate, sectoral and end-use levels. Section 2 is devoted to a discussion of the methodological aspects of acquisition, analysis and presentation of data on energy demand/consumption. Section 3 presents a review of approaches to analyzing changes in energy consumption in terms of demographic, economic and ecological variables. Section 4 provides illustrations from empirical work done in a few developing countries. The last section outlines a few research issues in this important area.

2. METHODS OF ESTIMATING CONSUMPTION/DEMAND

Estimates for Commercial Sources

Data on commercial sources of energy such as coal, oil, natural gas, and electricity are generally available in every country. A lot of effort may be required to assess the quality of information and prepare a consistent set of data showing production, trade and final consumption. Usually, new data are available only in aggregate terms or by product, and the pattern of uses does not distinguish between transformation uses/losses and final consumption. For example, coal output data may distinguish coal by grade but may not give details of collieries' own consumption of coal. Use of coal in power generation may include purchases of coal for use in thermal plants but may not include coal (middlings) received from washeries preparing coal for steel plants. Usually, data on coal consumption in the steel industry include a component (of washed coal) which is actually used for power generation. Similar problems may be faced in evaluating data on the use of furnace oil, LSHS (Low Sulphur Heavy Stock) and natural gas since information may only be available by region or by producing company rather than by end-use. Disaggregated data for direct exports of oil products may be available, but information may not be available on indirect exports, such as sales to international airlines, bunkers and exchanges between neighbouring countries. Further, disaggregated data on consumption by sector or end-use may not be available since suppliers keep records in a particular way. For example, details of use of high speed diesel oil (HSD) for trucks, buses, tractors, and agricultural pumpsets usually are not available since retail outlets do not keep records of the purposes for which consumers buy diesel oil. This creates problems of disaggregating consumption figures into transport and agricultural sectors. Sometimes,[1] there may be incentives to substitute one oil product for another due to price differentials, and it may be necessary to estimate the level of such substitution. Hence, detailed and carefully designed surveys are needed to ascertain the shares of diesel going to agriculture and transport or the extent of kerosene diverted to transport or industrial sector.

In countries where production of nitrogeneous fertilizers is important, there may be a problem distinguishing between the use of energy

1. In India, such substitution took place when kerosene (which was subsidised) used to be mixed with diesel. The extent of mixing was estimated at 20 to 25 percent. When the prices of kerosene and diesel were equalized, consumption of kerosene declined. Such factors have to be taken into account while interpreting changes in consumption of different energy products.

products (naphtha, gas, fuel oil, coal) as feedstock[2] or as fuel. Consumers may not be in a position to indicate the proportion of each fuel used as feedstock or fuel due to technical or economic[3] (differential pricing) reasons. Lack of such disaggregated data may overestimate/underestimate fuel inputs into the manufacturing sector and distort comparisons over time and across countries.

Since the fertilizer industry is one of the few highly energy-intensive ones, its energy use creates another conceptual problem. As fertilizers are exclusively used as input into agriculture/plantations, it becomes necessary to include the energy use in fertilizers as an *indirect* energy input into the agricultural sector in addition to the *direct* consumption of fuels/electricity used for operating farm machinery. In many countries[4] where the consumption of fertilizers is growing very rapidly, inclusion of indirect energy input (used in the manufacture of fertilizers)[5] will present a realistic picture of the increasing energy intensity of agricultural production. Such information would be useful for policy makers for the purpose of formulating programmes of agricultural development.

Estimates for Traditional (Noncommercial) Sources

Although biomass and animate sources account for a large proportion of energy consumption in developing countries, the data base for these sources is rather weak. Estimates of consumption will have to be made through a set of well-planned surveys[6] in different regions and seasons to reflect these variations. Macro-level estimates[7] would have to be made giving mean values and standard deviations. Breakdown of ag-

2. In fertilizer and petrochemical industries these fuels can be used as fuel or as feedstock for producing intermediates such as ammonia/ethylene.

3. In the Indian fertilizer industry in the sixties and seventies, there was no excise duty on naphtha used in the fertilizer industry while furnace oil was heavily taxed. Under these circumstances, entrepreneurs used naphtha both as feed and fuel.

4. According to the FAO, indirect energy (mainly fertilizers) accounted for 80 percent of total commercial energy used in agriculture in developing countries. In India, consumption of electricity and petroleum products used in the manufacture of fertilizers increased by 257 percent over a decade (1971-1982). See Bhatia (1984).

5. However, this may create another problem of double counting of energy input into fertilizers while analyzing the changes in energy use in the manufacturing sector. Since data on value of output (or value-added) in manufacturing will include figures for fertilizers, a comparison of energy-output ratio would require that fuels/electricity used in fertilizers should be added in the consumption of manufacturing sector.

6. For a methodological discussion of energy surveys, particularly in rural areas, see Bhatia (1985), Desai (1985), Islam et al. (1984). Also see Hoffman (1984) for a discussion of data on consumption and determinants of noncommercial energy use. Howes (1985) presents a critical review of issues and methods in rural energy surveys in the Third World.

44

gregate figures into different sources such as firewood logs, twigs and branches, dead trees, crop residues of various kinds, bagasse, dungcakes, etc., would have to be made on the basis of field surveys. It would also be necessary to take into account the methods of obtaining these fuels, e.g., from forests, roadside trees, homesteads, local woodlots, community land, own farm, purchases from market, exchange, etc. Since traditional fuels are collected by different members of the family over different seasons, records of these fuels are not generally kept. Since data on these fuels are available in nonstandard units (such as backloads, headloads, buckets, baskets, cartloads), it becomes necessary to actually measure the quantities of these units before assigning weights. Gross energy contents of these fuels, variations in moisture content over seasons, and efficiency of use of these fuels are other types of data that need to be collected in the surveys.

These 'estimates' of consumption have to be cross-checked with the estimates of supplies of biomass resources and animal stock.[8] For example, estimates of consumption of crop residues for fuel have to be checked against the figures of output of crop residues (using straw-grain ratios) and their estimated consumption for fodder and other purposes. Production of animal residues can be calculated with the number of (stock) cattle and norms of dung production per day. Use of dung for manure purposes may be estimated from agricultural surveys, and the ratio of dung collected to total dung produced may be used to determine an 'order of magnitude' figure for dung used as fuel.[9] In the case of woodfuels, the figures on logs may be checked against estimated fellings (official as well as illegal) in the forests. It must be kept in mind that a number of assumptions would have to be made in estimating the supplies of biomass resources, collection of these resources, uses of these resources as fodder and fertilizers, and the estimated consumption for fuel purposes. All these assumptions should be checked against micro-level information from agricultural, industrial and livestock surveys. Despite all the difficulties, it is necessary to make these estimates so that there is some way of checking against figures obtained by multiplying per

7. For some of the estimates for India see Government of India (1974, 1979), NCAER (1979, 1982); for Bangladesh, see Islam et al. (1984), Tyers (1978), Manibog (1982); for Pakistan, Government of Pakistan (1978); for Sierra Leone, see Davidson (1985); for other countries, see Islam et al. (1984).

8. For details of the methodology of assessment of renewable sources, see Bhatia (1985a). For estimates of biomass resources and their utilization, see Desai (1980), Islam (1983), Islam et al. (1984), Reddy (1981), Kennes, Parikh and Stowlijk (1984), Nair and Krishnayya (1985), Parikh (1983).

9. For an excellent description of the methodology of assessment of biomass resources used as fodder and fuel in India, see Desai (1980, 1985). Also see Howes (1985), Bhatia (1985a).

capita consumption norms with population figures. The data so obtained should be displayed in terms of various end-uses, such as household cooking and space heating, industrial heat and steam-raising, brick-making, animal work in agricultural operations, etc.

Methods of Aggregation and Presentation

After obtaining estimates of consumption of various fuels and energy forms, it may be useful to aggregate these for the purpose of ascertaining their relative shares in the total. One has to be very careful in aggregating different forms of energy by using conversion factors, because it may distort the picture regarding relative contribution of alternative sources. Generally, various fuels are aggregated in terms of their gross energy content, and the totals are presented in terms of million tonnes of oil equivalent (MTOE) or million tonnes of coal equivalent (MTCE) or million (trillion) Btu's or trillion kilocalories/joules. Such aggregation does not take into account the differences in quality of fuels in the sense that oil is a more convenient fuel and has a higher efficiency (than, say, coal) in many end-uses.[10] Since there are wide differences in efficiencies[11] with which various fuels are being used, aggregation in terms of 'useful energy' would indicate a higher share of oil products in the total.[12] Hence, data in original units should be used wherever possible, and aggregation should be done only to indicate rough orders of magnitude.

Estimation and aggregation of animate energy presents another set of difficulties which has to be taken into account. For example, the contribution of animal power to the total energy use can be estimated in three ways: (i) number of work animals may be estimated and converted into horsepower using, say, a figure of 0.5 horsepower per animal;[13] (ii) one may estimate the fodder consumption of these work animals and use its energy equivalent as the energy input of work animals, and (iii)

10. The Indian Government (1974, 1979) uses the method of Million Tonnes of Coal Replacement (MTCR) where one tonne of oil products is assumed to be replaced by 6.5 tonnes of coal because of higher efficiency of using oil in rail transport, cooking, etc. In coal equivalent terms (or kilocalorie terms), a tonne of oil would be equal to only 2.0 tonnes coal.

11. Figures on efficiencies of using various fuels in selected end-uses are difficult to get. But it is known that coal used in steam traction or biomass resources burnt in traditional stoves have very low efficiencies of utilization.

12. Shares of electricity consumption in the aggregate are also affected significantly whether MTOE is used or MTCR is used.

13. In the estimates provided by studies sponsored by the Indian Council of Agricultural Research on energy requirements of intensive farm systems, it is assumed that a pair of bullocks develops 1.0 horsepower. If 35 million pairs of bullocks are working in Indian agriculture, this method will assume that 35 million horsepower is available in

rather than taking the 'engineering equivalent' of work animals, it may be useful to estimate their contribution in terms of diesel oil required if all the work animals were replaced by tractors.[14] Each one of these methods has problems of interpretation and availability of data. Hence, it has been suggested[15] that data on the use of human energy and animal power in agricultural operations or rural transport should be shown separately in terms of person-days worked or animal-hours spent on specific tasks. Thus, it is important to recognize the difficulties of aggregating different forms of energy rather than mechanically applying some conversion coefficients and aggregating contribution of work animals with oil products and electricity.[16] Another problem is that for many end uses specific energy forms cannot be substituted for each other, for example, tractors and oxen.

Presentation of Data

Apart from the use of tabular and graphic presentation, there are two methods of presenting available data so that the energy transactions can be seen at a glance. These are: the Energy Balance Table (EBT) and the Reference Energy System (RES).[17] The EBT provides a method of presenting data on production, imports/exports, transformation losses and final consumption in a consistent matrix format. RES emphasizes the estimates of efficiencies with which fuels are converted or transformed, transported and consumed using various end-use devices. It provides the entire information with the help of a flow diagram distinguishing various stages from production to consumption of fuels. The energy balance tables are now available for a number of developing countries for the period 1971-1982. A time series of a consistent set of estimates of sectoral energy consumption obtained from the EBTs can be used to study the determinants of energy demand. The RES can be incorporated into planning models using mathematical programming techniques.

Indian agriculture. By assuming the number of hours these animals work, their contribution is calculated in terms of HP-hrs and converted into KWh/Kcal.

14. For alternative estimates see Bhatia (1977, 1985).

15. See Desai (1985).

16. As pointed out by Desai (1985), there are conceptual problems of adding energy in the form of mechanical work with energy from fuels.

17. For a detailed discussion of energy data systems, and the construction and use of EBT and RES, see W.N.T. Roberts (1983). Energy Balance Tables are available for 1971 to 1982 for a number of developing countries, see OLADE (1981), OCED (1984). For a description of the RES, see Mubayi and Meier (1984). For an RES for India, see Pachauri and Srivastava (1984).

Assessment of Shifts from Noncommercial to Commercial Sources

The estimates of consumption obtained by methodologies discussed in the earlier sections may be used to assess to what extent commercial fuels are replacing traditional (noncommercial) energy forms in cooking, lighting, transport, agricultural operations (ploughing, irrigation and crop processing) and small industries. In addition to these estimates, it would be useful to conduct special surveys on fuel usage in villages/cities where similar surveys were done in the past to get an idea of changes that have occurred in the intermediate period. Such assessments will also benefit from data on trends in prices of traditional fuels (say, in urban areas) and commercial fuels and increasing commercialization of traditional fuels.[18] It will also be necessary to relate this shift to fuel price paid by users, the availability of traditional fuels (deforestation, changes in crop mix, changes in stock and use of animals) and investments in renewable energy sources (biogas, solar cookers, etc.). A good understanding of these trends will help the policy maker to make more accurate demand forecasts on which to base plans for the future supply of commercial fuels and/or of programs of renewable energy sources.

One particular aspect which can be of importance is the increasing mechanization of agricultural operations in many developing countries. This may take the form of substitution of traditional sources (animal power) by mechanized devices over time. In some countries (e.g., India), the stock of work animals may remain more or less constant over time or may decline, and increased energy needs may be met by commercial sources. Besides, additional agricultural output may require increased dosages of chemical fertilizers (along with high-yielding seed varieties) which, in turn, would require larger quantities of energy products to be used as feedstock and fuel. Since increasing agricultural output (for meeting growing needs of food and for meeting export demand) is going to be necessary for almost all developing countries, growing energy intensity of agricultural output has to be explicitly considered.[19] Some of these issues are discussed later.

18. For a discussion on these aspects, see A. van Buren (1983).

19. For example, in India, total (direct and indirect) energy used in operating farm machinery and for manufacture of chemical fertilizers has increased two and one-half times over the last decade. Unit consumption in TOE per million rupees of value of output in agriculture has increased from 17 to 49 during 1970-71 to 1981-82.

Assessment of Unfulfilled Demand

In many developing countries consumption of energy does not reflect demand for the following reasons: (i) there may be constraints on supplies due to insufficient domestic production and/or imports;[20] (ii) there may be constraints on distribution or transmission facilities which restrain the supply of surplus fuels/electricity available in some locations to consumers;[21] (iii) due to low purchasing power of a large segment of the population (particularly in rural areas), the actually existing energy demand for basic needs such as cooking, lighting and space heating is not expressed in the market place. In the majority of cases such people meet only a part of their requirements, especially when scarcity of traditional fuels is increasing over time; (iv) the government may be reducing available supplies in attempts to curtail demand through rationing, load-shedding, quota allocations, etc.

Invariably, such constraints on demand or shortages have serious adverse impacts on economic activity and the welfare. Apart from the inconvenience to consumers, power breakdowns may result in the shutdown of industrial activities, lack of irrigation water to crops, increase in the consumption of material inputs, etc.[22] Since shortages may result in undesirable consequences and their impact would differ from one sector to the other, it is necessary to assess where the shortages (or unfulfilled demand) are occurring.

Apart from the possible adverse impacts on economic activity, continued shortages could result in consumer responses which may not be desirable from the viewpoint of society. For example, shortage of power to the industrial sector may result in captive power generation by consuming industries. Such captive generation or the installation of backup systems would involve higher unit investment costs (due to scale economies) and consume more valuable fuels[23] (e.g., diesel oil in place of

20. The supply of energy would depend on capacity utilization in the short run as well as creation of capacity in the medium/long run. Imports may have to be curtailed or checked due to balance of payments difficulties, nonavailability of supplies, etc. Capacity utilization in the short run would depend on factors such as weather conditions (for hydro), availability of imports (for thermal power), labour conditions and management difficulties.

21. In India, during the early eighties, coal supplies from the eastern region could not be moved through railways to western areas. It is estimated that currently (1985) about 500 MW of power available in the Singrauli area cannot be transmitted to other regions where shortages persist.

22. A number of recent studies have discussed the impacts of energy shortages. For a detailed description and analysis, see Munasinghe and Schramm (1983), Munasinghe (1979, 1980), Sanghvi (1983).

23. Available data on this aspect may not be sufficient for proper analysis. For example, in India, a large number of industrial units have installed diesel generators, and

coal or fuel oil). In agriculture, shortages may result in loss of output if power for irrigation is not available at the right time. Since the private profitability of irrigation usually is high, the consumer response would be to invest in backup systems[24] (e.g., a diesel engine) resulting in over-capitalization. In some cases, farmers use other farm machinery, such as tractors, for irrigation purposes, which results in inefficient use of diesel fuel. To the extent that small industries, brick-kilns and households could use logs, soft coke, kerosene, etc., interchangeably, the nonavailability of commercial fuels would result in excess demand for woodfuels with consequent deforestation. Similarly, shortages of commercial fuels may result in higher consumption of traditional fuels with adverse health effects on eyes and lungs. For households which are unable to obtain adequate supplies of cooking fuels, the adjustment may be in terms of less food or less nutritive food.[25] In some cases, shortages of kerosene/electricity for lighting may lead to use of vegetable oils which could be better utilized as food.

Although the adverse impact of energy shortages may be recognized, it may be difficult to provide estimates of unfulfilled demand and the resulting consequences. Two approaches are possible. The first requires direct estimation of unfulfilled demand through special surveys. The second would seek indirect evidence such as increases in prices of fuels which are not controlled, presence of black markets, evidence of increased deforestation due to household demand for woodfuels, reduction in fodder availability to cattle and reduction in use of organic fertilizers (due to diversion towards fuel). Both approaches may be used since these complement each other.

Studies on unfulfilled demand or shortages are rare.[26] Even when available, it is difficult to judge whether consumers have not over-estimated demand in anticipation of reduced allocations. Industrial and

the consumption of diesel oil for power generation has increased. Small diesel generators can be seen in commercial establishments in cities and small towns.

24. For example, in India, a high proportion of medium- and large-scale farmers are known to keep a diesel engine in addition to an electric motor. This is particularly true in the northwest region where irrigation is highly profitable. This over-capitalization is also on account of the fact that farmers are not charged for electric connection and electricity to irrigators is sold at only 20 percent of its true economic cost.

25. In some cases it is reported that households without adequate fuel supplies reduced their food consumption or reduced the consumption of food items which required more fuel (e.g., beans, lentils). See Cecelski et al. (1979), Cecelski (1984). Also, see Smith and Ramakrishna (1985) for the health effects of biomass fuels.

26. For example, in India, the Central Electricity Authority (CEA) has made estimates of anticipated requirements of power, supply and shortages for each year during the period 1974-75 to 1983-84. Their estimates show that power shortages as a percent of anticipated requirements have been more than 10 percent during this period with the exception of 1976-77 and 1982-83.

agricultural output may change due to factors other than energy availability and prices. Changes in fuel prices may be a result of a number of factors, including administrative changes, and, as such, it may be difficult to separate demand factors as distinct from supply factors. Consumption of food, use of fodder for animals and use of dung as manure may depend on a number of complex factors which would make it difficult to isolate the impact of shortages of energy. This is an important issue which needs to be addressed by further research.

3. UNDERSTANDING CHANGES IN ENERGY CONSUMPTION/DEMAND

Once a consistent set of data is available on energy consumption/demand, it is necessary to understand why and how energy consumption varies over time and across sectors, regions, industrial units and households. Only in-depth analysis of factors affecting energy consumption can indicate what type of policy measures are needed to adjust energy demand in the context of macro objectives. Such analyses may first be done at the macro level, relating energy consumption with population, level of economic activity (gross output, GDP), structure of economic activity, patterns of industrialization, urbanization, energy prices, technological changes, etc. These macro analyses should also be supported at the sectoral level (say, for industrial, agricultural and transport sectors), thus bringing out the significance of various technological, economic and organizational factors relevant to each sector. These studies should be supplemented by studies on specific fuels such as oil, electricity, coal, etc., at the aggregate as well as sectoral level to isolate the effects of certain specific variables that determine their utilization. Micro-level studies on consumer behavior of households and potential substitutions between energy and nonenergy inputs, including conservation measured in industrial units, should also be conducted.

Energy Demand Analysis at the Macro Level

Energy demand analysis at the macro level should consist of the measures considered below: (i) comparing the growth rate of consumption of energy with the growth rate of gross domestic product (GDP) for a given period or subperiods and estimating the elasticity coefficient with respect to GDP; (ii) estimating GDP-elasticity coefficients, assuming a log-linear relationship, through regression analysis, using data for the entire period; (iii) computing energy consumption per unit of real GDP (at constant prices) over time and interpreting the results with respect to increases or decreases in energy-GDP ratios; and (iv) using computable

٢٨٤

general equilibrium models which take into account the interactions affecting aggregate growth, sectoral patterns of growth, the balance of payments situation and energy prices. Such analyses should be done separately for total commercial energy and total energy (commercial and noncommercial) as well as for different sources of commercial energy such as oil, coal, electricity. Since the data base for estimated consumption of noncommercial sources is rather weak and these sources account for a major share[27] in the total, it is important to ensure that these inaccuracies do not affect the picture regarding commercial sources of energy.

Comparing Growth Rates

Using the first method of comparing growth rates over different periods is convenient for seeing the relationship between energy consumption and economic activity. It is of particular interest to policy makers when they want to see if the energy-GDP elasticity has increased or decreased over specific periods (e.g., pre-oil price hike to post-oil price hike, etc.). Since this ratio is very easy to calculate, it is tempting to use this aggregate measure without analyzing the underlying reasons which might be responsible for changes in growth rates of energy consumption and GDP. For example, if the ratio of the growth rate of energy consumption to that of GDP in the post-1973 period has been lower than that in the pre-1973 period, one may want to conclude that energy is being used more efficiently in the post-1973 period. This, however, could be an erroneous conclusion because the growth of energy consumption might have been restricted due to supply shortages and the growth rate in GDP might have been affected by extraneous factors[28] (e.g., weather, export performance). Another problem is that this method is very sensitive to the selection of base year and terminal year for estimating growth rates. Although it provides a quick summary, this coefficient has very limited value for assessing future energy demand, the possible effects of conservation efforts, and other factors such as the effect of differences in relative growth rates of energy-intensive versus low-energy using activities, etc. To make projections of future energy

27. In a recent study by Bhatia (1985d), it has been estimated that in 1982-83 in India, total consumption of commercial energy was 83 MTOE (Million Tonnes of Oil Equivalent) compared with 157 MTOE of noncommercial energy. Thus, noncommercial sources still accounted for about two-thirds of the total.

28. In the Indian study of energy demand management, Bhatia (1985d) has shown that there is no stable relationship between commercial energy consumption or GDP. For example, in 1974-75 commercial energy consumption increased by 6.64 percent, corresponding to a real GDP growth of 0.66 percent. In 1977-78, when the real GDP increased by 8.8 percent, energy consumption increased by 4.16 percent.

demand, a much more disaggregated analysis by sectors and end-uses must be undertaken.

A further problem which affects the interpretation of the energy-GDP relationship is that energy consumption may be affected by factors other than GDP, and that changes in GDP may occur independently of changes in energy use. Supply factors depending on weather conditions (which affect output of hydro plants), capacity utilization in energy-producing industries and balance-of-payments situations may restrict the availability of energy even though demand may exist. On the other hand, changes in GDP may take place due to weather conditions (as they affect agricultural output), labor conditions, management factors and various extraneous factors such as world market demands for exports, etc. Under the circumstances, basing projected changes of energy consumption to projected changes in GDP may yield highly erroneous results.[29] However, exceptional events such as prolonged droughts/crop failures, etc., can be controlled by standard econometric refinements.

Energy-GDP Ratios

Another apparently easy method of macro-level energy demand analysis is to compute energy-GDP ratios over time, or across countries/regions for comparison. However, the conceptual issues in using the energy-GDP ratio discussed above, the problems of measuring of energy consumption as well as GDP and the problems with cross-country comparisons are well documented in the literature.[30] The widespread use of and potential substitution by noncommercial fuels, the nonreporting of output of small-scale sectors, the presence of unaccounted-for output in the organized sector, the effect of choice of the base year for constant GDP values, as well as conversion factors for different currencies, etc., invalidate the use of energy-GDP ratios for cross-country comparisons. However, for a given country, the ratio can be used as a measure of short-term changes in the energy-GDP relationship. It has been suggested that[31] "the energy-GDP measure is inadequate as an index of energy demand. The necessity of dissecting the energy use and output-generating activities to understand energy demand adequately is para-

29. An analysis of supply factors for energy consumption and GDP has been presented for each year after 1974-75 in the Indian case study by Bhatia (1985d). The analysis brings out the fact that growth rates in the consumption of commercial energy are essentially determined by supply factors. Similarly, changes in agricultural output and other factors were responsible for annual fluctuations in real GDP. Thus, to use the energy-GDP relationship as a demand relationship would not be correct.

30. For details see Siddayao (1982a, 1982b), Smolik (1978), Desai (1978, 1979), Choe (1980), World Bank (1981).

31. Siddayao (1982a).

mount to any forecasting activity. The energy-GDP ratio is merely a preliminary indicator of the demand for energy. It is a short-cut approach and serves as a preliminary approximation of how energy demand may be related to (a) economic activity, (b) standard of living, and (c) stock of energy-using equipment. It should not be given a role more important than this."

Even when it is used as a summary measure, it must be kept in mind that three very important factors are lumped together in the E/GDP ratio:[32] (i) changes in the mix of goods and services in the GDP; (ii) changes in the energy intensity with which the particular goods and services are produced; (iii) changes in the composition of energy, i.e., energy sources which provide higher useful heat may be added or replaced (coal replacing oil or fuelwood).[33] Since the composition of GDP (agriculture, industry, transport) as well as energy intensity of industrial mix (heavy, light) are of importance in any energy analysis. Therefore, the summary measure of E/GDP has rather limited value.[34]

Use of Regression Analysis

The advantage of regression analysis is that a number of variables other than GDP can be included and income and that price elasticity coefficients are directly obtained if a double-log formulation is used.[35] Econometric models of different levels of complexity can be formulated to estimate elasticity coefficients (short-term and long-term) with respect to GDP and energy prices.. These models can also incorporate the effects of changes in population, urbanization, agricultural growth, structure of GDP, pattern of industrial growth, relative prices of energy and nonenergy inputs, etc.

It would be outside the scope of this paper to review the available literature along with the results obtained. For this, reference should be made to the literature cited in this paper.[36] Some of the important methodological issues and results are discussed below.

32. See Schipper, Meyers and Sathaye (1984), Hoffman (1984).

33. In India, the E/GDP ratio has increased over the last decade. Rather than showing that energy efficiency has gone down, it may show that a higher percentage of coal is being used now than before.

34. See a discussion on Korea in Schipper, Meyers and Sathaye (1984).

35. See Siddayao (1982a, 1982b), Pindyck (1979), Hoffman (1981), Hall (1983), Kouris (1981), Berndt, Morrison and Watkins (1981). See the report by EMF4 Working Group in the *Energy Journal* (1981) and the World Bank (1981) review of the literature. Also, see J. Reilley (1982), Bohi (1981), Munasinghe and Schramm (1983), Dunkerley (1982), Dunkerley et al. (1981).

36. For empirical results see Siddayao (1982a, 1982b), Bhatia (1985), Fernandez (1980), Reilley (1982), Hoffman (1980), World Bank (1981).

One of the important limitations of these empirical studies in the context of developing countries is that consumption data do not reflect demand. For reasons discussed above, these reflect demand that is constrained by supply factors. One of the ways to tackle this problem is to include supply variables in the regression equation. Such variables could be weather conditions (affecting hydro output), an index of capacity utilization in energy-production industries, an index of labour relations, of transport capacity, etc. However, under these circumstances, it would not be correct to call it an energy demand function. It would, then, be a functional relationship explaining variations in energy consumption and would incorporate factors affecting demand as well as supplies.

Second, there is a trial and error approach needed in the selection of explanatory variables, the functional forms of the equations and the lag structure involved. There are also difficulties involved in econometric estimation where such problems as multicollinearity, serial correlation, etc., are encountered, thus affecting the statistical significance of parameters being estimated.

Third, there is the problem of specifying the causal relation between energy consumption and economic activity. In demand analyses, causation is postulated from economic activity to energy demand. However, in many developing countries (particularly in sectors such as mechanized agriculture, energy-intensive industries and transport) it may be the supply of energy which may be affecting economic activity.[37] A thorough analysis of various factors affecting energy consumption and economic growth in a simultaneous system would require use of macroeconometric[38] and computable general equilibrium models as discussed below.

Some empirical results of estimation efforts undertaken to understand the relationship between energy consumption and GDP have been reported in two studies.[39] The population variable has been either included explicitly or incorporated directly in the income and energy consumption variables by stating these in per capita terms. The price variable is usually taken to be the price of oil. In country studies this is

37. As discussed in Bhatia (1985d), there is reason to believe that availability of energy may affect a part of agricultural output (particularly in a bad agricultural year) and industrial output of continuous process industries and energy-intensive industries. Shortages of energy would adversely affect capacity to transport goods (by railways or roadways) which, in turn, would affect industrial/agricultural output.

38. One of the macro-economic models using Indian data is by Rahman (1982).

39. See Siddayao (1982a, 1982b). See also UN ESCAP (1981). These five countries are: India, Nepal, Sri Lanka, Fiji and Pakistan. The other countries included in the analysis are: Burma, Indonesia, Malaysia, Philippines, Singapore, Thailand, China, Korea, Cambodia, Laos, Bangladesh and Afghanistan.

taken to be the domestic price, but in international studies the price of the benchmark crude is most often used. The results were mixed, depending on the estimation form and period of estimation. In some cases quite unexpected and/or statistically insignificant results were obtained. Some of the highlights of the results are: (i) for the period 1965-78, income elasticity coefficients for all but five developing countries were greater than unity. Estimates for a shorter period, 1973-78, showed changes both upwards and downwards. In five out of twelve countries where these comparisons were possible, the income elasticity coefficients were higher than unity. Among these countries were India, Indonesia and China; (ii) the estimated price elasticity coefficients for several countries were statistically insignificant. There were also some unexpected results, namely positive values for price elasticity. The statistically-insignificant results during the period 1973-78 for most of the countries affected may reflect the pricing policies adopted by these countries which depressed energy prices below their true economic values. The generally low values of the coefficients which were statistically significant further reflect the inability of the countries to switch to other fuels during the short period, given the oil-based technology in place; (iii) these results show that this simple approach fails to take account of structural changes, such as changes in the industrial structure, or increased mechanization in agriculture, or the shift toward or away from the production of energy-intensive goods and long-run response to the changing price structure.

Some results available from a recent study[40] in India indicate the following: (i) for the period 1966-67 to 1982-83, using log-linear form, the estimated elasticity coefficient of commercial energy consumption with respect to per capita real GDP is 1.15, and the elasticity coefficient with respect to share of manufacturing is 0.706; (ii) however, when a real price variable (for energy prices) is considered along with real GDP, the statistical relationship is poor. The elasticity coefficient for per capita GDP is 0.62 while the elasticity coefficient for real price of energy is unexpected, i.e., is positive and significant. Thus, the use of regression analysis in single equation models of demand has serious limitations, and their use should be discouraged, particularly for demand projections. Similarly, one should be very cautious in using regressions of electricity or oil on GDP and projecting the demand for these forms of energy for investment planning or analysis of balance-of-payments. Such analyses

40. See Bhatia (1985d). The elasticity coefficient for total (commercial plus traditional) sources with respect to GDP is very low, i.e., 0.064, mainly due to the assumption of constant per capita consumption of noncommercial sources of energy.

have been used in some country level and international studies[41] but these have limited value as they do not consider possible changes in the energy-mix over time.

Use of General Equilibrium Models

Recently, a number of studies have been undertaken in which energy demand analysis is done with the help of computable general equilibrium (CGE) models.[42] The approach emphasizes the energy-economy interactions where two-way feedbacks between energy demand, aggregate growth, the sectoral patterns of growth, balance-of-payments situation and external shocks, etc., are explicitly considered. With such models, a number of different types of policy analyses can be carried out. These include studies of the effects of changes in the levels and composition of government expenditures/taxes, and changes in external conditions, such as the international price of petroleum, and foreign borrowing. Since these models require the use of large computers, their applications to developing countries have been confined to a few research groups in industrialized countries. Although these models are very powerful in linking different variables affecting energy-economy interactions, these should be used with caution because some of the assumptions made may not be valid and the quality of data/parameters used may not be good. Particular care is required with assumptions made with regard to the processes of adjustments since developing countries are characterized by administered prices,[43] rigidities of supply, imperfections in labour and capital markets and restrictions in imports. Similarly, parameter values for elasticities of substitution between energy and non-energy inputs, income and own-price elasticities of de-

41. For Thailand, Atilaksana (1983) has estimated the elasticity coefficient for per capita electricity consumption with per capita GDP as 2.98308. This has been used for demand projections. Wolf et al. (1981) also use income and price elasticities for projecting oil demand for developing countries.

42. The major studies are: Blitzer and Eckaus (1985) for a group of five East Asia countries; Blitzer and Eckaus (1983) for Mexico; Choucri and Lahiri (1984) for Egypt; Gupta (1983) for India and Dick et al. (1983) for four developing countries (Kenya, Ivory Coast, South Korea and Turkey). Subba Rao et al. (1981) have used an input-output model to estimate the marginal increase in energy requirement for a unit increase in particular types of final demand. For an example of the use of microcomputers for energy planning in Sri Lanka, see Munasinghe et al. (1985).

43. For instance, in a study on India by Gupta (1983), it was assumed that changes in import price of crude oil are passed on to the consumers in terms of higher/lower petroleum product prices. As discussed in Bhatia (1985d), such adjustments are rarely done since the prices are administered by the government and increases in product prices have been kept in check by reducing the share of excise duties and sales taxes in retail prices. Similarly, lower import prices are also not passed on to the consumer.

mand, cross-price elasticities of demand, etc., should be used with caution. Besides, the level of disaggregation of energy sectors (as well as non-energy sectors) is such that it is difficult to trace the impact of changing the prices of one or more petroleum products on interfuel and interfactor substitution.

Energy Demand Analysis at the Sectoral Level

In view of the problems associated with EDA at the macro level, it is necessary to quantify determinants of demand at the sector level, e.g., for agriculture, industry, transport, household, etc. The techniques applied to each sector can be adapted to the availability of data and the causative mechanisms involved; for instance, it is possible to use different models for agriculture, industry, transport and households. However, if independent sectoral models are developed, the estimates they provide are unlikely to be mutually consistent, and a further mechanism for assuring their mutual consistency has to be introduced.

Since energy balance tables provide a time series of consistent data on sectoral final consumption, they can be a suitable starting point for sectoral energy studies. The attempt should be made (i) to quantify the relationship between the level and pattern of energy consumption and the level and pattern of economic activity variables; prices of energy and other outputs; changes in demographic variables, tastes and technology, etc. Some of the issues these relationships raise, along with a few empirical studies, are discussed below.

Energy Use in the Agricultural Sector

Energy consumption in the agricultural sector should include both commercial and animate sources of energy. The estimated consumption should also include *direct* energy used in operating farm machinery as well as *indirect* energy (naphtha, natural gas, fuel oil or electricity) used in the manufacture of fertilizers, pesticides and farm machinery.[44] Since agriculture is both a consumer and producer of various forms of energy, it is necessary to assess, over time, the contribution of agricultural residues and animal residues towards meeting the basic energy needs for cooking, particularly of the rural and urban poor. Analyses of energy use in agriculture at the macro and micro levels in

44. It is considered important to include both the direct and indirect energy consumption in order to show the fast-increasing energy-intensity of the agricultural sector, at least in India.

developing countries have been recently reviewed.[45] According to the FAO (1972), the average use of commercial energy per hectare of cultivated land in developing countries in 1972 was only one-tenth of that in developed market economies. However, from 1972-80, the estimated consumption of commercial energy in agriculture in developing countries increased by 117 percent, the bulk of the increase being accounted for by growth in fertilizer consumption. The corresponding increase in crop output was only 28 percent, indicating that the intensity of commercial energy use with respect to crop output was increasing over time. This pattern is true for all the developing countries analyzed, although the elasticity coefficient is relatively lower for Latin America compared to that for Africa or the Far East. The increasing commercial energy intensity of agriculture is also evident from the data for India where total energy consumption over the period 1970-71 to 1981-82 has increased by 257 percent, while the real value of crop output has increased only by 23 percent.[46] Similar studies for other countries could be carried out to quantify the interrelations between commercial energy use and agricultural growth; effects of mechanization on the availability of food, fodder, fuel and fertilizers and the effects of changes in relative prices on interfuel substitution, level of mechanization, etc. Since agricultural systems vary from one region to the other, the relationship between crop yields and energy-input has to be analyzed under varying assumptions of input prices, tillage requirements, type of irrigation, crop systems, tenancy conditions, farm-size, etc. Although a number of micro-level studies are available, these tend to be more in the nature of engineering studies. There is a need for socio-economic studies on labor-use, efficiency of use of irrigation water and fertilizers as well as the impact of tillage on output. Such studies can indicate policy aspects of improving the efficiency of energy input (direct and indirect) in agriculture.

Energy Consumption in the Industrial Sector

Energy use in the industrial sector can be analyzed at the aggregate, sub-sector (group of industries), industry or enterprise level. At

45. See Bhatia (1984). A summary of main points is also available in Bhatia (1985c). Parikh and Kromer (1985) explore the interrelations between energy and agriculture in Bangladesh with the help of a general linear programming model.

46. Regression results using data for 1970-71 to 1982-83 for India, assuming log-linearity have been obtained for the following variables: (i) *direct* energy consumption, i.e., oil and electricity used in operating farm machines; (ii) direct energy consumption of the previous year; (iii) real value of GDP from agriculture; (iv) real price of energy. The results show that short-run electricity w.r.e. GDP is 0.232, that the real price variable is significant and that the short-run price elasticity is -0.74.

۲۹۲

the aggregate or industry-group levels, time-series data can be analyzed to quantify the changes in energy use due to different growth rates of branches and subsectors, and changes due to increases or decreases in the energy intensity of individual subsectors. The former is referred to as structural change while the latter is referred to as technical change.[47] The aggregate consumption of energy in the industrial sector[48] can also be analyzed using a regression equation where value of output (in constant prices), structure of industrial output (e.g., share of heavy industry) and real price of energy are the independent variables. Such a study would provide estimates of output and price elasticities.[49] Jankowsky (1980) has presented an analysis for four countries--Brazil, India, Korea and Kenya--where he found that industrial energy demand is a function of the absolute size of the industrial sector, the structure of output and the energy intensity of production. In Kenya, industrial demand grew faster than industrial output while in Brazil rapid industrial growth resulted in comparable increases of energy consumption, despite the growing role of heavy industries in the total. In India, the structure of fuel sources--namely the high dependence on coal--appeared more significant in explaining industrial energy demand than did the structure of its output mix.

Sometimes a particular form of energy (e.g., electricity or oil) is selected for specific analysis. The demand model could be specified by including the following independent variables: industrial output, electricity consumption lagged by one or more years (time period), lagged prices of electricity and its substitutes (petroleum and coal). The parameters of this model could be estimated using time-series data for a number of industries and responsiveness of electricity demand to these

47. A study on energy use in Mexican manufacturing by Sterner (1985) for the period 1970-81 shows that the increased energy intensity of manufacturing industry is found to depend entirely on the adoption of more energy-using technology, whereas the changes in output structure have had little or no effect.

48. In India, unit consumption of commercial energy has declined from 73.3 TOE in 1972-73 to 60.3 TOE in 1981-82, a reduction of 18 percent. This indicates that the average energy-intensity of industrial production is declining. Over a 10-year period, 1972-73 to 81-82, commercial energy consumption has increased by 43 percent while the real value of output in manufacturing has increased by 74 percent, implying an elasticity coefficient of 0.58.

49. In the Indian study, the regression results show that the short-run output elasticity is 0.75 while the energy price (real terms) elasticity is -0.436. The long-run output elasticity is estimated as 0.75/(1-0.15) or 0.88. The long-run price elasticity is -0.513. See Bhatia (1985d).

independent variables.[50] Similar studies for various oil products for Greece, Spain, Turkey, Brazil and Mexico are reported in Pindyck (1979).

In addition to these efforts at explaining aggregate and product-wise energy consumption in the industrial sector, a few studies have been undertaken in developing countries which estimate the elasticities of substitution between energy and nonenergy inputs as well as the elasticity of substitution among different fuels.[51] This approach has been widely used in developed countries. Since energy demand is a derived demand, the objective of these studies is to estimate elasticities of substitution between one energy type and another and between the energy aggregate and other factors of production. This is done by using a two-stage approach. In the first stage energy demand is a derived factor demand (assuming that factor inputs are chosen to minimize the total cost of production) based on a translog cost function.[52] Factor inputs include capital, labor, materials and energy. In the second stage expenditures on energy are broken down into expenditures on oil, natural gas, coal and electricity, now under the assumption that fuel inputs are chosen to minimize the cost of energy. The use of this two-stage approach requires certain additional assumptions about the underlying structure of production. In particular, it is assumed that the production function is homothetically separable[53] in the capital, labour and energy aggregates, i.e., (i) that expenditure shares for fuels are independent of the expenditure shares for capital and labour, and (ii) that expenditure shares for fuels are independent of total energy expenditures. Using this approach and the underlying assumptions, price elasticities of demand and elasticities of substitution have been estimated. For example, in the two industries selected for empirical verification,[54] (cotton textiles and

50. A study of demand for electrical energy in Korean industry has been reported by Uri (1983). The long run price responsiveness in all industries is estimated as -1.1, suggesting that a 1 percent change in the price of electrical energy is accompanied by a 1 percent change in consumption. The effect of output on industry electricity energy consumption is important with the effect being roughly the same across industries, with the exception of paper and printing and nonferrous metals.

51. The studies for developing countries are: For India: Williams and Laumas (1980), Apte (1983) and Chaube (1982); For Indonesia, Pitt (1985).

52. For methodological details and empirical estimates for developed countries, see Pindyck (1979). For a detailed review article, see World Bank (1981).

53. Weak separability of energy aggregate from capital, labour and material implies that the marginal rate of substitution (MRS) between any two energy inputs is independent of the quantities of capital, labour or material.

54. (See Dubey (1982)). These two industries broadly satisfied the assumptions implied in the use of a cost minimization approach. Due to the presence of a large number of firms and the lack of any regulation of output prices, the assumption of perfect competition in the (textile industry) output market was justified. The major input, raw

YQC

small-scale iron and steel) coal was found to be a substitute for both oil and electricity. These results could be used for adjusting energy prices towards efficiency prices. For Indonesia (Pitt (1985)) energy-use in various industries has been found to be responsive to changes in energy prices; 21 out of 27 own-price energy elasticities have absolute values above 0.50. Own price elasticities for labour are generally much smaller in magnitude than they are for energy. The results also show that a doubling of energy prices would induce a reduction in energy consumption in the range of 30 to 40 percent and would result in a rise in total cost of less than 2 percent. Although interesting and significant results can be obtained by using the translog cost functions, the applicability of this approach in the context of the special characteristics of energy and output markets needs to be very carefully assessed due to the following reasons: (i) data on energy demand refer to consumption rather than demand; (ii) consumption decisions may be more a function of availability than of prices (or relative prices); (iii) user choices may be heavily affected by auxiliary costs (e.g., appliances, equipment) associated with each fuel source, and by non-price considerations which may outweigh price differentials; (iv) the bulk of commercial energy may be used by the government and public enterprises whose management may not be responsive to changes in energy prices; (v) entrepreneurs may be in a position to pass on higher energy costs to consumers because of regulated price by government or lack of competition, with the result that they would have no incentives to adjust energy use in response to relative price changes.

Energy Consumption in the Transport Sector

Although a lot of literature exists on energy demand in the transport sectors of developed countries, empirical studies for developing countries are rare. The transportation sector energy demand in developed and developing countries may differ in a number of ways: by the share of different modes (passenger cars versus railways and bus transport), by different end-uses that a particular mode may serve (e.g., railways used more heavily for passenger transport relative to freight in developing countries) and energy-intensities of different modes (e.g., road transport may be more energy-intensive than railways). Besides, it may be important to consider other objectives than just energy-efficiency such as employment generation, use of domestic fuels instead of imported fuels, etc.

cotton was also easily available. In the steel industry, almost 80% of the output was in the public sector where inputs are allocated by the government and output prices have been regulated. However, for small firms using steel scrap as input, there are no controls on output and input prices.

Available energy studies[55] have analyzed: (i) aggregate demand for transportation services; (ii) energy demand by type of transport service, e.g., total passenger transport, by mode, intracity passenger transport, freight transport by rail or road; and (iii) demand for various transport fuels such as diesel oil, electricity, petrol (gasoline), etc. A review of some of these studies shows that the short-term price elasticity for petroleum products, particularly for gasoline in the transport sector, is relatively inelastic. The results also show that the demand for these products is income-elastic, particularly in the long-run. However, this may well be due to failure to strictly observe ceteris paribus conditions, e.g., cars replacing mopeds while incomes are rising.

Increased road freight transport relative to rail transport and the growth of international air freight has raised the average freight energy requirements per ton-kilometer. On average, rail transport has only one-third to one-quarter of the energy intensity of truck freight transport. However, even though the energy intensity is greater than rail, the convenience of truck transport may improve overall economic efficiency. Besides, there may be capital cost and employment factors which may determine the policy options regarding inter-modal choices in freight and passenger transport. Improvements in energy efficiency in the transport sector may have to be considered along with high capital costs of new equipment and the foreign exchange outflows on account of oil imports.[56]

In their econometric studies of the energy demand in transport in Mexico Berndt and Botero (1983) concluded: "In terms of affecting demand for energy quantitatively, output or income appear to be considerably more important than price, especially in the railroad and air transport sectors. For the gasoline and diesel fuels in the motor vehicle mode, however, price is of substantial importance."

Energy Use in the Household Sector

In some developing countries, the household sector may account for a more significant share of oil products consumption (kerosene for lighting and cooking) and electricity than is generally assumed. More-

55. See Noll (1982), Berndt and Botero (1983), Parikh (1980), Pindyck (1979), Dewan (1978), Bhatia (1985d).

56. In India, there has been a significant interfuel substitution by the railways such that the elasticity coefficient of commercial energy with respect to output in the transport sector is only 0.53. Due to intermodal shift (railways to road) and fuel substitution within railways, the energy efficiency of transport has improved. However, this is less due to conservation or more efficient use of energy but more due to shift from coal to diesel and electricity.

over, with program of rural electrification and problems of nonavailability of noncommercial fuels, consumption of commercial energy in the household sector is likely to increase over time. It is difficult to predict the extent and pattern of change. Hence, there is need for some analysis of past consumption data both for rural and urban households and of available options for the future.

Existing studies[57] use both macro-level data as well as data obtained from household expenditure surveys or special purpose energy surveys. Most of these use aggregate national and/or regional data on total energy consumption and/or specific sources to estimate elasticities with respect to income, price and demographic variables. The studies on Bangladesh, India, China, Pakistan and Indonesia use aggregate data to estimate fuel consumption for various end-uses (cooking, lighting, appliances, etc.) for different consumer categories such as urban/rural, small/medium/large farmers, high/low income levels. The level of detail as well as the sophistication of analysis varies. In some, only bivariate tables have been given while some others have attempted econometric analysis to estimate income and price elasticities. For example, for all commercial fuels, for Pakistan (Iqbal (1984)), the values of short-run and long-run income elasticities have been estimated as 1.23 and 2.27, respectively. The figure for long-run income elasticity seems to be higher than estimated in other studies for developing countries. While the long-run income elasticity for the consumption of gas plus electricity is 4.18, it is much lower (1.26) for coal plus kerosene. This supports the hypothesis that gas and electricity are superior to kerosene and coal. The own-price elasticities of gas plus electricity and coal and kerosene are negative and statistically significant at the 10 percent level.

Data on household energy consumption in eight non-OPEC developing countries have been analyzed to explore the level and composition of total fuel consumption by households at different income levels within each country (Fernandez 1980). Comparisons are made between urban and rural areas in the same country as well as among countries. The data show that noncommercial fuels are the dominant energy source in both urban and rural households at all but the highest income levels. The results on income elasticity of fuel consumption indicate that the elasticity of consumption of all fuels combined is in the range of 0.6 to 0.8. In urban areas, the elasticity of consumption of commercial fuels is

57. Some of these studies are: Cecelski et al. (1979), Dias-Bandarnaike and Munasinghe (1983), Amarullah (1984), Fernandez (1980), Cody (1980), Havrylshyn and Munasinghe (1980), Strout (1978), Wilcox (1980), Berndt and Samaniego (1983), Bhatia (1985d), Parikh (1980), Kennes et al. (1984), Reddy (1981), Nair and Krishnayya (1985), Pitt (1985a), Iqbal (1984), Sathaye and Meyers (1985), Glapke and Fuzzolore (1985), Zhu et al. (1983).

much higher than for noncommercial fuels while in rural areas the elasticities for the two fuel categories are more nearly equal. Although these findings should be regarded with caution because of technical difficulties in estimating the elasticities, it is clear that fuel consumption increases noticeably among higher income households within each country. The study also shows (based on the analysis of data from Korea and Pakistan) that although household size affects the estimates of income elasticity, no conclusions are possible regarding the magnitude of the influence of household size. It does appear, however, that estimating income elasticities without controlling for household size leads to overestimates of the true elasticities for Korea and Pakistan. Another conclusion is that electricity and gas consumption are dramatically higher among higher income groups. This large increase is the result of two reinforcing effects: (i) an increase in total fuel consumption among high income households, and (ii) their strong tendency to substitute electricity and gas for noncommercial fuels.

The interaction of fuel prices, capital costs of different energy conversion devices, and associated thermal efficiencies in the use of woodfuel, kerosene and liquid petroleum gas (LPG) for domestic cooking in Sri Lanka has also been analyzed (Havrylshyn and Munasinghe (1980)). Comparisons are made for cooking costs per family per annum when different fuels are valued at market prices as well as at social opportunity costs. Given the efficiency estimates for different stoves and calorific contents of different fuels, the economic analysis in a dynamic setting of changing fuel prices and fuelwood scarcities suggests that LPG may have an important role to play in easing the fuel problems in urban areas in Sri Lanka. Evaluated at social opportunity costs, the use of kerosene would be cheaper than firewood (using open fire). In view of rising oil prices and rapid increases in deforestation, the policy options should include a detailed analysis of kerosene, charcoal production and use of closed stoves. The paper also emphasizes the importance of other non-price variables such as reliability of supply, ease of handling and storage, convenience, and potential risks.

Assuming a log-linear relationship between per capita kerosene consumption and explanatory variables, regression analysis has been used for Indian data for the period 1965-66 to 1981-82 (Bhatia (1985d). The independent variables are: (i) per capita disposable income since kerosene being the item of household expenditure would be influenced more by changes in disposable income than by those in GDP; (ii) real price index of kerosene (1970-71 = 100); and (iii) kerosene consumption of previous year. The results show a statistically significant elasticity coefficient of 0.744 with respect to per capita disposable income. The elasticity coefficient for relative price of kerosene is negative (-0.46) and statistically significant.

In the Indonesia study (Pitt (1985a)), kerosene demand elasticities with respect to total household expenditure are twice as high (0.543) in rural Indonesia as in urban Indonesia (0.265), the average being 0.450. These elasticities are also inversely related to expenditure levels. Total expenditure elasticities of demand for firewood are negative, that is, firewood is an inferior good, in all geographic areas (except rural Java). The total expenditure elasticities for charcoal are 2.453 for rural areas, 0.181 for urban areas and 0.961 for Indonesia as a whole. The elasticity of kerosene consumption with respect to its own price is close to -1 for all the population groups. The elasticities of charcoal demand with respect to the price of woodfuels are somewhat smaller than those of firewood and the own-price elasticities of kerosene. The elasticity of demand for firewood with respect to the price of kerosene is not statistically different from zero. This suggests that there is no basis to the claim that the kerosene subsidy alleviates the deforestation problem in Java. The results also show that wealthier, urban households, for whom kerosene is the primary energy source, obtain a disproportionate share of total kerosene subsidy.

The papers on residential electricity demand in Mexico, Indonesia, Costa Rica and seven countries in West Africa have also estimated income and price elasticities using econometric models. The study on Mexico (Berndt and Samaniego (1983)) makes a distinction between access and consumption of electricity. This study considers that increases in income have an important "double-whammy" impact on electricity consumption, first in terms of increasing the number of households hooked up to electricity services, and second in terms of increasing the consumption of households already having access to electricity. While each of the components has estimated long-run income elasticity values of less than unity (in Mexico), their joint effect has been estimated to be considerably larger than one.

The study on Costa Rican electricity demand (Dias-Bandarnaike and Munasinghe (1983)) explicitly considers the quality of electricity supply by distinguishing between three levels of quality--medium, low and very low--on the basis of voltage-variation (in practice voltage-drop) level. The results obtained indicate that the quality of supply could have a significant impact on planned electricity consumption, and thus must be taken into account in demand forecasting. For example, *ceteris paribus*, residential consumers who faced low and very low qualities of electricity supply in Costa Rica were found to consume only 85 percent and 70 percent respectively of the electricity they would have consumed if the supply quality had been medium. The second important consideration, as brought out by the study, is that shifts in the planned demand curve due to changes in reliability imply that the benefits of electricity consumption must also change. For example, in Costa Rica, the total benefits of electricity consumption of households at the medium reliability level were 1.176 and 1.429 times the total benefits at low and very low qualities of supply, respectively. These results can be used in investment planning where the determination of optimal power system reliability levels requires trading off the increased costs of strengthening the system (to improve the quality of supply) against the increased benefits of consumption to users.

4. RESEARCH ISSUES IN ENERGY DEMAND ANALYSIS

Based on the above discussions, the following research issues have been identified:

(i) Although some preliminary studies have been done to identify the role of traditional sources, develop estimation techniques of the supply of biomass fuels, and survey methodologies for data on consumption, there is need for much more empirical work in this important area. A concerted effort is needed to identify the end-uses where these sources are of importance, review available data at macro, regional and village levels and identify important data gaps. However, it is surprising to note that in some countries (e.g., India), a large volume of data already exists which has not been properly analyzed. Efforts should be made to analyze these available data for major sectors such as household cooking, space heating and lighting (both in rural and urban areas), agricultural operations, small-scale industries, brickmaking, rural transport, etc. The studies should focus on issues such as changes in supply and demand for biomass fuels; determinants of the use of biomass sources as fuel, fodder and fertilizer; changes in efficiencies of conversion and utilization; processes of generation and use of traditional sources of energy in the contexts of socio-political characteristics of a village, etc.

(ii) Since meeting basic energy needs of populations at reasonable cost is an important policy objective, it is necessary to identify regions and households which are currently facing deficits in biomass fuels or are likely to face them. It is necessary to analyze how current needs are being met and what the welfare consequences of existing and increasing shortages are. The studies should be on a refined basis with the relevant region defined in terms of biomass fuel transport units. They should analyze available energy survey data (where available) combined with re-visits to the field. An important component of these studies would be an analysis of the processes of commercialization of biomass fuels, shifts to commercial fuels and welfare (disposition of income) implications of such shifts for both urban and rural poor.

(iii) There is a need to review the methodologies and collate the results of available studies on energy demand analysis and management. Such a review should collate and compare empirical information on the role of energy prices, possibilities of inter-fuel and inter-factor substitution, economics of conservation vis-a-vis production (or imports), etc.

(iv) In view of the importance of the impacts of energy shortages, there is need for a few country studies to examine this subject. Such research would involve (i) analysis of various methodologies of estimation of actual energy shortages or unfulfilled demand for different sectors and categories of consumers, (ii) some tentative estimates of the magnitude of these energy shortages and the processes through which they were managed (e.g., rationing, quotas, breakdowns, load-shedding, rostering of electricity loads, sudden imports, changes in consumer habits) and (iii) quantification of the impacts, if any, of these shortages on agricultural output, industrial output, employment, regional development, income distributions, consumer welfare, consumer health, etc.

(v) In the context of energy demand, there is need to evaluate the social profitability of the allocation of various scarce resources (capital, foreign exchange, management skills, technical manpower, organizational capabilities) to energy conservation vis-a-vis augmentation of supplies (through imports, higher production or better distribution).

REFERENCES

Amarullah, M. (1984). *Electricity Demand in Indonesia: An Econometric Analysis*, Pusat Penyelidikan Masalah Kelstanikan, Jakarta.

G. Anandalingam (1984). The Economics of Industrial Energy Conservation in Developing Countries, mimeo, Tata Energy Research Institute.

[4]

IAEE Convention Speech:

Energy, Exhaustion, Environmentalism, and Etatism*

*Richard L. Gordon***

Editor's Note: The author, Dr. Richard L. Gordon, won the IAEE's Outstanding Contributions Award for 1992. The following article is based on his acceptance speech given at the 16th international conference of the IAEE held in Bali, Indonesia, from July 27-29, 1993. The Association awards a prize annually for outstanding contributions to the profession of energy economics and to its literature.

INTRODUCTION

Two radically different views compete in energy economics. The first, stressed by academic economists, considers energy as a commodity subject to general economic laws. The tendency to technical progress is viewed as the most critical economic law involved. The second position—favored by those in the energy industries, the national and international agencies concerned with energy, and environmental groups—is that energy is unique in various (usually undefined) ways. I examine here first the question whether the consumption of energy and other so-called exhaustible resources is excessive. Modern economics

The Energy Journal, Vol. 15, No. 1.

* Appreciations are due the Association for honoring my efforts, my teachers and students for instructing me in sound economics and how it applies to energy, and The Pennsylvania State University for providing a fertile climate in which to work, the referee who commented on the substance of the paper, and M.A. Adelman, as he has done so often in my career, providing invaluable suggestions for revisions. This paper was the basis for my talk at the 1993 IAEE Convention in Bali. It offers reflections inspired by my work rather than fresh scholarship. In response to comments and reflection, this version tries to make clear how widely held are the views expressed.

** Professor of Mineral Economics and MICASU Faculty Endowed Fellow, College of Earth and Mineral Sciences, The Pennsylvania State University, 204 Eric A. Walker Building, University Park, PA 16802-5010.

in contrast to nineteenth century gloom, is the optimistic science stressing the ability of market economies to overcome adversities. In particular, markets clear more quickly and at lower prices than those fearful of market disturbances expect. Conversely, the dominant theme of academic writings is that governments have done more harm than good in energy.

This experience has broader implications that are best treated elsewhere. What is critical here is that the views expressed here are almost universally supported by academic energy economists, whatever their political outlook. The pro-market stance of academic energy economists including me arises from prolonged observation of energy developments. Assertions that ideology or abstraction is dominant are calumny.[1] For example, the skepticism certainly does not result, as its critics often charge, from the illusion that real economies resemble the idealized general equilibrium model of the text books. The need to incorporate the complications ignored by simple theory was part of energy economics long before consideration of such problems attained wider acceptance.[2]

The key dispute in energy is between those who believe that only human ingenuity limits economic development and those who feel that resource availability ultimately constrains material economic growth.[3] The evidence clearly supports the proposition that the human ingenuity has long prevailed over resource scarcity and suggests this situation will persist for at least the next half century. Another aspect of the debate whether energy is over- or underpriced because of market or government failure.

Three main subpoints are made here. The first is that the Hotelling (1931) model of exhaustible resource behavior was badly misused. The theory had great value for better organizing a sensible discussion of resource scarcity

1. I shared the belief that the vigorous defenses of market economies were unacceptable polemical tracts until, in the middle 1980s, I started reading the material. The excesses proved to worse than the reactions produced by favorably citing the material.

2. This acceptance is reflected by growing attention to transaction costs and to the role of contracting in industrial organization and microeconomics texts such as Kreps (1990). A less salutary development is a growing tendency to overstate the deficiencies of the general equilibrium model. Examination of the workings of a simplified economy is a useful pedagogical starting point.

3. This point was made many times. Basically, the optimism proceeds from Adam Smith's arguments about the long-run effects of the division of labor. Malthus and Ricardo took a step backwards from Smith by stressing the barriers to increasing the wealth of nations. One key modern pioneer in applying these ideas to natural resources was the geographer Erich Zimmermann. He asserted "*resources are not, they become* (emphasis in the original)." Unfortunately for the preservation of his influence, the concept was embedded in a massive survey of natural resource supply conditions when he was writing. Thus the bulk of the book is hopelessly outdated. To surmount this problem, in 1964 Henry L. Hunker extracted the key discussions and provided updatings that, of course, are by now also obsolete.

and initially was used in that fashion. Later writers lost touch with reality. This produced sterile, empirically questionable discussions.

My second point is that whatever the practical relevance of the Hotelling analysis, its policy implications are probably nil (or are that exhaustible resource consumption should be *increased*). The final concern is possible tendencies to overstate the case for restricting energy to reduce environmental side effects. In developing these arguments, I attempt to show that each has broad support as well as distinguished criticism.

Many resist this message. Two important types of opposition come from opposite directions. First is the energy establishment of landowners, producers, and the agencies of national governments and international organizations devoted to promoting producer interests. The others are the advocates of various forms of resource pessimism who fear that separately or synergistically resource exhaustion, pollution, and population growth are leading to ruin. Such concerns are predictable.

Regrettably, some economists with no such vested interests, including a few very distinguished ones, have become so enchanted with elegant theories assuming resource pessimism that they have come to believe they are valid. Robert M. Solow (1992) dismissed as "mindless wish fulfilment" the proposition that "ingenuity and enterprise can be counted on to save us from the consequences of consuming too much and preserving too little, as they have always done in the past."[4]

THE HOTELLING ANALYSIS AND RESOURCE PESSIMISM

In treating the two points about resource exhaustion, I refrain from taking the easy route of mocking amateur efforts. I examine the economic theories of resource constraints initiated by Harold Hotelling (1931).[5] Hotelling's article raised almost all the issues treated in the subsequent literature. The Hotelling model languished in obscurity from its publication to the proliferation of neoHotelling models in the 1970s and 1980s.

4. An interesting contrast was T.C. Koopmans (1974) who thoroughly examined the applied as well as the theoretic literature on exhaustion and accepted the view that exhaustion was not a practically relevant problem.

5. L. C. Gray's (1914) model dealing with dealt with optimization by an individual mine facing constant prices is the earliest known predecessor of Hotelling and is much less complete. The Hotelling analysis was one of several in which he used the calculus of variations, a technique for minimizing the value of integrals, to deal with important economic problems.

Previously resource economists correctly analyzed the issues without employing Hotelling.[6] The classic presentation by Barnett and Morse (1963) of the ingenuity-driven model of natural resources proceeds with only a cursory note (p. 46) that the Hotelling analysis exists. As a test of a scarcity problem, the authors use the basic Hotelling proposition that with impending exhaustion prices and costs of the products being depleted would rise.[7] Barnett and Morse found that the price and profitability record for exhaustible resources is the opposite from that implied if resources were exhaustible. Prices and profits fell.

Starting in the late 1960s, many economists used some form of the extended Hotelling analysis. These contributors include many with close experience with natural resource markets, and some of their work antedates the height of the vogue for Hotelling models. This work is worth reviewing because it illustrates that Hotelling's work does not clearly dictate a pessimist attitude about resource scarcity. The pioneers indeed used Hotelling to support optimism.

The first explicit statement of which I am aware that the proper interpretation of Hotelling model was that exhaustion was not a pressing problem was by Orris C. Herfindahl (1967). In 1967, I slightly strengthened Herfindahl's conclusions and tried to develop the analysis further.[8] Herfindahl suggested that exhaustion costs (or as Scott 1967 termed them, following Keynes, user costs) were too low to matter; I indicated that the constraint might be nonbinding and user costs could be nil.

I also tried to the limits of my mathematical skills to explain Hotelling's general model in which the effects of using up the lowest cost resources also are considered. I correctly deduced that delaying cost increases was an additional benefit from delaying production. The addition to the model of benefits from increased quality preservation necessarily implies that less incentive must be

6. The most important exception to assertions about the obscurity of the Hotelling analysis is a paper by Paul Samuelson (1957) showing the model provides optimality conditions for the storage of *renewable* resources between crops.

7. Another key aspect of the book is recognition that conservation and populism had a heavy element of protecting established interests from the challenge of newcomers.

8. Since he and Anthony Scott were published in a book (Gaffney 1967) of conference papers and my writings appeared in journal, I had priority of publication. However, my work (1966 and 1967) was inspired by reading their manuscripts.

provided by simply adding to the quantity preserved.[9] Many subsequent authors starting with Cummings (1969) developed more elegant representations. Their cost-increase-avoidance term consists of the cumulative discounted value of the cost increases in future periods imposed by increasing output at any time.[10]

Thus, a fundamental defect of much 1970s literature, ignoring the implications of depletion of low cost reserves, could have been avoided by more carefully reading Hotelling and prior commentators on his work. Hotelling devoted the first pages of the article to a simple case in which, throughout time, costs are constant and thus equal to average costs. This case and a variant in which costs are essentially zero are too often presented as *the* Hotelling analysis. No zero-cost case arises in Hotelling; it is in Irving Fisher (1930) where he deals with how shipwrecked sailors allocate a fixed stock of "hard-tack" —presumably some kind of biscuit that never deteriorates. Its price should rise at *r* percent. It could also be called the Saudi Case since it so closely approximates the unique situation of the Saudis.

In more general cases, the *r* percent rule reduces to the basic economic principle that nothing is an asset unless it earns the market rate of interest. This can occur in an infinite number of ways. They can include Hotelling-like price rises due to hoarding or the creeping up of costs as we resort to more expensive-to-exploit resources. They can also involve cost declines of both the Gray type of movements down the marginal cost curve by cutting output or by Barnett and Morse type technical progress. The theory then indicates how profits are made on mineral exploitation depends on the circumstances.

Another aspect of the literature that remains unsatisfactory is the treatment of uncertainty. The analyses tend to be overly complex, incorporating endogenous efforts to explore in response to concerns over exhaustion. Such models incorrectly characterize exploration. It proceeds regularly simply because companies recognize that good opportunities remain.

9. As I have confessed elsewhere (Gordon 1981a), my 1967 article contained two errors. I postulated that when demand increased at more than percent a year, exploitation would be delayed until demand growth declined below *r* percent. Gordon (1981a) also contains my demonstration of why this conclusion is wrong. I show why, in fact, if exploitation does not start before growth falls below *r* percent, it will not start at all. My second error was to argue that efficiency required production cost minimization in each time period. Goldsmith (1974) demonstrated that efficiency only required equating price to the sum of production and exhaustion costs. It is efficient to offset lower production costs with higher benefits to preserving resources. In addition, I correctly derived the optimality conditions for selecting the time to start and cease extraction and used them to confirm Herfindahl's assertion that the conclusion of conventional price theory that competitive firms never produce at outputs below that at which minimum average cost occurs also applies to firms producing exhaustible resources.

10. The most important extensions of Cummings work were those by Baumol and Oates (1988), and by Modiano and Shapiro (1980). Both papers showed that in a discrete time analysis, the solution could be derived using the constrained maximization concepts of differential calculus

Key flaws are the implicit assumptions that exploration immediately produces accurate estimates of the economically recoverable minerals contained and that decisions are based only on known discoveries. The first assumption is factually incorrect. For example, Trocki (1986 and 1990) surveyed the development of iron ore and copper supplies in the twentieth century and showed that depending upon the circumstances the development ranged from those occurring immediately to those that had to await the invention of technology to allow the use of the ore.

Finally, these analyses seem inadequately to treat expectations. Theory suggests that decisions are based on all available knowledge including recognition that more resources will be identified in already discovered deposits and more discoveries will be made. Anticipation then means that actual exploration has little effect on prices. In short, the standard economic theory that decisions are based on rational expectations as defined by Muth better handles uncertainty than do models of endogenous exploration.

The most vigorous exponent in the 1990s of the view that exhaustion effects are small is M. A. Adelman (1993). Campbell Watkins (1992) also made a valuable contribution in his IAEE presidential address. Both improve upon Barnett and Morse by using a measure better related to the Hotelling analysis, the value of mineral reserves.[11] These arguments use data analogous to that presented by Barnett and Morse. Adelman and Watkins demonstrate that mineral reserve values do not increase over time.

The standard criticism of resource optimism is that we cannot be sure that ingenuity will last. This ignores the long history and, more critically, the deluge of indicators that enough promising ideas are emerging to guarantee several decades of continued progress. The literature on energy technology provides ample evidence that further advances will occur.

Definitive proof (of anything) is never possible, but a badly neglected principle for economic policy is that commitments to the future should be limited. We are poor at guessing when, how, and to what extent major changes will occur. This suggests that we can and should delay response until the dangers of alleged problems are clear. Better systems of data collection and appraisal are needed, but experience also warns that policy makers do poorly in data analysis.[12]

The record since 1974 provides chastening examples of incorrect pessimistic forecasts and premature investment in response to such predictions. We see endless updating by resource pessimists of their perpetually incorrect

11. Barnett and Morse (p. 225-6) recognized the potential value of such estimates but considered measurement infeasible. Adelman and Watkins secured usable data sets.

12. This last sentence was inspired by one of M.A. Adelman's suggestions to me; the call for better data analysis is his; the fear that the effort will fail is mine.

forecasts of impending shrinkages of energy supplies, "shortages" of investment funds, and thus the inevitability of rising oil prices. With only a lowering of the height of the trend line, the same discredited price-rise scenarios are reiterated. If anything, the work became worse in the 1990s. These forecasters worry about the transitional problems of decommunization and ignore that a move to market economies might massively reduce energy consumption, raise supplies, and provide a major addition to oil and gas supplies for the rest of the world.

EXHAUSTION AS A POLICY PROBLEM

Let us now turn to the second point that energy-supply-availability difficulties neither affirm nor refute the validity of another familiar assertion that energy must be managed. If ingenuity rules, *dirigisme* is probably, but not definitely, unnecessary. However, even if supply limits dominate, the market solution still seems far superior to planning as a means of adaptation.

The literature on exhaustible resources clearly establishes that resource exhaustibility creates no special market failures (see footnote 9). Moreover, the implications for optimal exhaustion of plausible market failures are unclear at best. The theorists agree that monopoly generally will retard depletion (although if the demand curves have special properties, unlikely to prevail in practice, monopoly may not affect the rate of exhaustion).

In 1966, I proved that even if discount rates are inappropriately high, the effect on exhaustion is ambiguous (Gordon 1966). A high interest rate has both an impatience and a cost impact. The impatience factor causes a rush to reap profits. The cost factor causes the industry to be less profitable and discourages production. Which effect predominates differs with the circumstances. This conclusion was derived by pursuing hints in the work of Scott and Herfindahl including draft versions of their classic articles. They made me aware that in exhaustion, as in other areas, the implications of changing interest rates are indeterminate in general.

While many others have made the same discovery, the empirical relevance is inadequately recognized. In such cases as Saudi Arabia, in which costs are far below price and depletion effects are small, impatience would dominate decision making in a purely competitive world. The Saudis, in practice, underexploit because their decisions are more strongly influenced by an effort to monopolize. For the more typical mineral venture with a much lower profit margin and no monopoly power, the disincentive-to-development effect is more likely to predominate. Thus, an inefficiently high interest rate, the market failure that many analysts—not including me, believe is highly relevant to exhaustible resources may cause overly protracted or too rapid extraction.

Moreover, the simple rationale for assuming inappropriately high interest rates, the inadequacy of risk-hedging institutions, is invalid. The critical problem is that the supporting theory assumes hedging is costless. Once the transaction costs of hedging are recognized, it follows that not all conceivable futures markets inevitably would have benefits in excess of costs. Appraisal of the adequacy of futures market becomes an empirical question. Observing how financial markets operate in advanced economies suggests a high capability for supplying those services whose benefits match their costs. For example, well-functioning futures markets for oil have emerged.

The real problem is how the convoluted tax laws of the world distort decision making. The complexities are such that neither the direction of impact nor the appropriate cure is apparent. That antidote, in any case, should not concentrate on exhaustible resources, but deal with all investments. A natural caution arises here from the largely unhappy experience of the US in singling out minerals for tax provisions, the depletion allowances, to compensate for the effects of the basic tax law. Thus, even if exhaustion were a real problem, policy issues do not inevitably arise.

THE LESSONS OF ENERGY ETATISM

Experience reinforces this viewpoint. The history covering many countries over many years and all energy sources produces nothing but bad examples. Denunciations of *dirigisme* comprise the vast majority of writings on energy policy. Academic writers on energy economics contend that everything from the crown jewels of energy to the dregs has been mismanaged. This malpractice seems independent of time, space, or nominal ideology.

The principal policy errors were the forces that produced the cartelization efforts of the main producing countries of OPEC. M. A. Adelman has provided compelling documentation that arrogant, economically illiterate policy makers in industrialized countries, particularly the U.S., failed to understand why they should whenever possible foster competition in energy and avoid undermining it.[13] Part of the problem was aid to high-cost domestic industries. Another part was the politicians' tendency to personalize. They should have recalled Adam Smith's point "It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their

13. He has in progress a history of world oil prices; an anthology of his published papers on oil (Adelman 1993) should appear almost simultaneously with this article. The statement in the text applies equally well to the papers in the anthology and the thus-far-circulated working papers for the price history.

regard to their own interest (1776, I, ii, 2, p. 26-7)." Adelman errs, if at all, by possibly crediting these policy makers with knowing what they were doing.[14]

Domestic energy policies around the world involve further errors The United States—currently the world's largest economy, with extensive energy resources, a particularly strong economics profession, and a tradition of free inquiry—is the best studied case by far. The biggest and best demonstrated deficiencies are in oil and gas. Here a succession of universally undesirable interventions occurred. Both price floors and ceilings were serious failures.[15]

Ancillary deficiencies arose in coal, nuclear power, electricity, and alternative energy. U.S. energy research and development had only one great success, the light water reactor, and it benefited from having a strong national defense basis.[16] Energy is also one of many victims of another dubious policy—the decision of the U.S. government to retain ownership of about 30 percent of its land area.[17]

Europe, having mismanaged coal for many decades, turned to designing poor policies to govern the development of North Sea oil and gas resources. The Japanese have imitated the European in overprotecting domestic coal; the big difference is that because the industry was smaller, so was the impact of the blundering. Topping all this was that Communist inefficiency affected all sectors of the economy. Since the Communist countries were large energy producers, the general problem had large energy impacts.

14. In his comments to me, he indicates that OPEC might have overcome the resistance.

15. Each key policy is the subject of ample, invariably negative review. See Mancke (1974, 1976) for a good summary of the situation up to then and for a later discussion, see Robert Bradley (1989). Bradley has prepared a massive survey of oil and gas regulation starting with the first World War, which should appear in 1994.

16. My comment about light water reactors was made with full recognition of the numerous attacks that have been made on the technology; such attacks ignore both the successes of light water reactors and the difficulties of commercializing the supposedly preferable technologies. I have conducted sustained and extensive work on coal policy around the world. The nuclear literature is diffuse, but Zimmerman (1987) provides a good start. The electric power problem has produced many valuable books. Schmalensee (1979) is the single best analysis of the conceptual issues; the essay collection edited by Moorhouse (1986) is among the best critiques. Discussions of the main energy research and development failures are too scattered to review here.

17. Again, we have many interesting contributions. Clawson (1983) gives an excellent overview of public-land policy. While good-single volume surveys exists for other parts of land law, minerals have been too much for any one book. Mead (1985) has done excellent work on offshore oil and gas; Nelson (1983) did a valuable survey of coal policy within the Department of the Interior. I have published several papers (1981, 1984, 1985, 1987, 1988) on other aspects of coal leasing. Leshy (1987) has provided an interesting view of the law covering hard-rock minerals including uranium that suffers from inadequate appreciation of the need to stimulate efficient private ownership laws.

UNDERSTANDING ENVIRONMENTALISM

My third, most difficult, and most controversial objective is making sense of a discussion that is far more problematic than its proponents admit. We economists correctly define environmentalism as efficiently internalizing externalities. Policies have strayed from this goal.

My concern here is with ideologues. Most who call themselves environmentalists are simply well-intentioned respondents to warnings of environmental damages. They are your neighbors and mine exposed to the warnings about world ruin and contributing as they would to a United Fund.

I seek to deal with one larger-than-realized faction of advocates who believe that a general resource scarcity problem exists and use environmental scares to secure support for an antigrowth agenda.

The modern economic argument for intervention to control environmental externalities is based on a misreading of a remarkable 1960 article by Ronald Coase. In preCoase days, people talked about markets inherently failing to internalize the externalities. Since Coase, advocates of environmental policies talk of the necessity, for intervention because of a large class of externalities for which the transaction costs are so high that the market is less efficient than governments at internalizing.

What Coase really said is that in these cases, it is unclear what, if anything, should be done. In many cases, inclusion of transaction costs implies that the true costs exceed the benefits. Whatever the payoff, the government may be no better than the market at reaching a correct decision. This view of externalities, even if correctly interpreted, is too narrow because it ignores professional prophets of doom who magnify the dangers of every possible threat.

My view is based on a regular review of government reports and scholarly studies on environmental issues and observations of practice. My cases are major ones on which the primary literature was extensively surveyed.[18]

The prime example is the attack on nuclear power.[19] The critics disagreed whether the concern was routine radiation releases, reactor accident, or danger of nuclear terrorism. As each problem proved manageable, the case turned to complaints about the absence, for which the critics were responsible, of a definitive waste storage program. (I consider Three Mile Island evidence that in free societies safe reactors will be built and Chernobyl only as proof of the dangers of totalitarianism.)

18. The literature suggests that many of the secondary concerns such as with dangerous materials in buildings, waste piles, and food are even more dubious than the problems I treat. While critiques from clearly impartial sources exist, accessible summaries are more readily found in polemics such as those by Ray (1990, 1993).

19. This debate is so widely publicized that citations are simultaneously too vast to provide here and unnecessary. No dispute exists about what are the issues; the debate is over interpretation.

The sulfur dioxide threat also may have been exaggerated. Sulfur-dioxide pollution control is justified mainly to prevent health impacts. However, the available health impact estimates differ widely and all rest on a shaky statistical foundation (see Lave and Seskin (1977) for the pioneering estimates, Ramsay (1979) for later lower figures, and the update in Portney (1990)). Studies of North American acid rain problems clearly showed that special problems of damages to ecosystems had costs far less than cleanup costs. The Congressional Office of Technology Assessment (1984) tried to salvage the case by reusing the Lave and Seskin numbers to show health benefits justified action.

Questions also arise about concerns about global warming. Meteorologists seem to differ radically in their certainty about whether rising global temperatures are certain to occur.[20] Modeling greenhouse effects is at least as complex modeling the macroeconomy, and thus the models of greenhouse effects should be viewed as skeptically as macroeconomic models. Even if global warming occurs, its impacts are unclear.

Economic analyses of abatement are ambiguous about whether the benefits of preventing global warming exceed the costs of control. (This work has appeared in many places, most notably in a special issue of *The Energy Journal* (1991) and an anthology edited by Dornbusch and Poterba (1991). However, the participants often are the same.) Nordhaus has taken the lead in arguing that vigorous abatement would not be cost-effective. He has distinguished support (and criticism). Cline's (1992) attack on Nordhaus relies on changing enough assumptions to make more vigorous action seem efficient. Cline still ends up admitting that attention must be given to improving knowledge (1992, p. 368-78).

Even more problematic is the claims made by energy conservation enthusiasts (Grubb 1990 and Grubb et al 1991) who insist that global warming can be cured almost costlessly by adopting energy-saving measures claimed to be socially profitable. The argument for conservation in any case is independent of global warming concerns and implausible. The only market failure that distorts energy use is underpricing of electricity by regulators, and deregulation is preferable to the conservation programs into which utilities were forced. The cost estimates are inadequately supported with observations of experience.

20. The accessible literature expressing skepticism is limited and difficult for nonspecialists to appraise. Of the three skeptics cited here, only, Lindzen, is a senior faculty member at a major university. Some of colleagues in Penn State's Meteorology Department nonrandomly sampled believe strongly in the reality of the problems and others indicate good reasons for skepticism and a lack of a clear professional consensus.

With Gore (1992) and many others, the attack is upon far more than environmental externalities.[21] At a minimum, many environmentalists are advocates of some variant of a broad resource-limited vision of the world. (See Bailey, 1993, for examination of the influence of the proponents of general resource pessimism on the environmental movement.) Having rightfully failed to convince people of the threat of exhaustion, the resource pessimists may be seeking new rationales that appear more plausible.

Parts of the environmental movement simply hate free markets. Frankland and Schoonmaker (1992), in their study of the German Green party, recognize that the party turned to environmentalism as an anti-capitalist doctrine with more appeal than socialism. The book considers but rejects charges that the shift is merely tactical. Frankland and Schoonmaker believe that the Greens are seeking unsuccessfully a "third way".

Even this partial account seems far from a record of rational efforts efficiently to internalize environmental externalities. In politics and punditry, one apparently can cry wolf without losing credibility.

The response of economists to this experience is also problematic. The literature uniformly denounces the policies practiced and supports the contentions made above that many programs are unneeded. These critiques, however, rank among the more restrained economic appraisals of regulation. The comments on environmental policy lean more to the incremental reform approach also often applied to antitrust and public utility regulation. The critiques thus are far milder than those directed at energy price controls.

Naturally, I cannot state unequivocally that all the environmental concerns criticized here are invalid. However, the no-admission-of-doubt posture of environmentalism combined with the many false claims should inspire greatly increased skepticism. As economists, we should apply to environmentalism the same doubts we addressed to the self-serving claims of producers of energy or anything else supposedly deserving aid.

21. As Lott (1992) argues, Gore's book seems the quintessence of the tendency to exaggeration criticized here. Reexamining the book suggested that its most clearly vulnerable feature is portraying environmental problems as poorly publicized. The argument I find most offensive is Gore's declaration (189) on exclusion of environmental bads from gross domestic product, "Philosophically, it is similar in some ways to the moral blindness implicit in racism and anti-Semitism...." For those readers unfamiliar with how GDP was defined, it be recalled that the designers were well aware that many important activities, good and bad, were excluded. The benefits of unpaid labor in the household or in charitable activities as well as environmental damages were omitted. The omissions were made to produce numbers that were computable at reasonable cost and could be considered resistant to political manipulation. Gore (346-8) unconsciously illustrated how politicizing a "green" GDP would be by his complaint that the Bush Administration's Council of Economic Advisors had produced too low a value of the threat of global warming.

CONCLUSION

Human ingenuity has been remarkable at advancing the real standard of living and warding off the pressures of resource depletion. The menu of promising developments available to us suggests this process will prevail for many more decades. The immediate need then for avoiding depletion is nil and, even if it existed, etatism is not the answer. The argument that an exception must be made for correcting externalities has been so overused that its validity is imperiled.

REFERENCES

Adelman, M. A. (1993). *The Economics of Petroleum Supply: Papers by M.A. Adelman 1962-1993*. Cambridge, MA.: The MIT Press.

Anderson, Terry L., ed., (1993). *NAFTA and the Environment*. San Francisco: Pacific Institute for Policy Analysis (jointly sponsored with the Fraser Institute and the Political Economy Research Center).

Bailey, Ronald (1993). *Eco-Scam: The False Prophets of Ecological Apocalypse*. New York: St. Martins Press (A Cato Institute book).

Balling, Robert C. Jr. (1992). *The Heated Debate: Greenhouse Predictions Versus Climate Reality*. San Francisco: Pacific Research Institute for Public Policy.

Barnett, Harold J. and Chandler Morse (1963). *Scarcity and Growth: The Economics of Natural Resource Availability*. Baltimore: Johns Hopkins Press for Resources for the Future.

Baumol, William J. and Wallace E. Oates (1988). *The Theory of Environmental Policy*, 2d ed. Cambridge: Cambridge University Press.

Bradley, Robert L. Jr. (1989). *The Mirage of Oil Protection*. Lanham: University Press of America.

Cline, William R. (1992). *The Economics of Global Warming*. Washington: Institute for International Economics.

Clawson, Marion (1983). *The Federal Lands Revisited*. Baltimore: Resources for the Future.

Coase, Ronald H. (1937). "The Nature of the Firm." *Economica* NS 4 (4, November): 386-405. Reprinted in George J. Stigler and Kenneth E. Boulding, eds., (1952). *Readings in Price Theory*. 331-51. Homewood. Richard D. Irwin, Coase (1988), and Williamson and Winter 1991.

_________(1960). "The Problem of Social Costs." *Journal of Law and Economics* 3 (November): 1-44. Reprinted in Coase, 1988.

_________(1988). *The Firm, the Market and the Law*. Chicago: University of Chicago Press.

Cummings, Ronald G. (1969). "Some Extensions of the Economic Theory of Exhaustible Resources." *Western Economic Journal* 7 (3) (September): 201-10.

Dasgupta, P. S. and G. M. Heal (1979). *Economic Theory and Exhaustible Résources*. Cambridge: Cambridge University Press.

Dornbusch, Rudiger and James M. Poterba, eds., (1991). *Global Warming: Economic Policy Responses*. Cambridge, MA: The MIT Press.

The Energy Journal (1991). Special Issue on Global Warming. Edited by Leonard Waverman. 12 (1).

Fisher, Irving (1930). *The Theory of Interest*. New York: Macmillan.

Frankland, E. Gene and Donald Schoonmaker (1992). *Between Protest and Power: The Green Party in Germany*. Boulder: Westview Press.

Globerman, Steven, and Michael Walker, eds., (1993). *Assessing NAFTA: A Trinational Analysis*. Vancouver: The Fraser Institute.

Goldsmith, Oliver S. (1974). "Market Allocation of Exhaustive Resources." *Journal of Political Economy* 82 (5) (October November): 1035-40.

Gordon, Richard L. (1966). "Conservation and the Theory of Exhaustible Resources." *Canadian Journal of Economics and Political Science* 32 (3) (August): 19-26.

_________(1967). "A Reinterpretation of the Pure Theory of Exhaustion." *Journal of Political Economy* 75 (3) (June): 274-86.

_________(1978). "Price and Output Behavior of Exhaustible Resource Industries." *Materials and Society* 2 (3): 247-51.

_________(1981a). *An Economic Analysis of World Energy Problems*. Cambridge, MA: The MIT Press.

_________(1981b). *Federal Coal Leasing Policy, Competition in the Energy Industries*. Washington: American Enterprise Institute for Public Policy Research.

_________(1984). "Access to Federal Lands for Profit Making Purposes–An Economic Overview." *Materials and Society* 8 (4): 699-718.

_________(1985). "Levies on U.S. Coal Production." *The Energy Journal*. (Special Tax Issue) 6: 241-54.

_________(1987). "Coal in U. S. Land Policy." in John Byrne and Daniel Rich, eds., *Planning for Changing Energy Conditions*. New Brunswick: Transaction Books, Inc., 139-72.

_________(1988). "Federal Coal Leasing: An Analysis of the Economic Issues." Discussion Paper EM88-01, Washington, D. C.: Energy and Materials Division, Resources for the Future, July.

Gore, Senator Al (sic) (1992). *Earth in the Balance: Ecology and the Human Spirit*. Boston: Houghton Mifflin.

Gray, Lewis C. (1914). "Rent under the Assumption of Exhaustibility." *Quarterly Journal of Economics*. 28 (2, May): 466-489. Reprinted in Mason Gaffney, ed., (1967). *Extractive Resources and Taxation*. Madison: University of Wisconsin Press, 423-446.

Grubb, Michael (1990) *Energy Policies and the Greenhouse Effect: Volume One: Policy Appraisal*. Dartmouth: The Royal Institute of International Affairs.

Grubb, Michael et al. (1991). *Energy Policies and the Greenhouse Effect: Volume Two: Country Studies and Technical Options*. Dartmouth: The Royal Institute of International Affairs.

Heal, Geoffrey (1976). "The Relationship between Price and Extraction Cost for a Resource with a Backstop Technology." *Bell Journal of Economics* 7 (2) (Autumn): 371-378.

Herfindahl, Orris C. (1967). "Depletion and Economic Theory." in Mason Gaffney, ed., *Extractive Resources and Taxation*. Madison: University of Wisconsin Press, 63-90. Reprinted in Herfindahl (1974). 64-89.

_________(1974). *Resource Economics: Selected Works*. Edited by David B. Brooks. Baltimore: Johns Hopkins University Press, for Resources for the Future.

Hotelling, Harold (1931). "The Economics of Exhaustible Resources." *Journal of Political Economy* 39 (2) (April): 137-75.

Kneese, Allen V. and James L. Sweeney, eds., (1993). Handbook *of Natural Resource and Energy Economics*. Vol. 3. Amsterdam: Elsevier.

Koopmans, Tjalling C. (1974). "Ways of Looking at Future Economic Growth, Resource and Energy Use." in Michael S. Macrakis, ed., *Energy: Demand, Conservation, and Institutional Problems*. Cambridge, MA: The MIT Press, 3-15.

Kreps, David M. (1990). *A Course in Microeconomic Theory*. Princeton: Princeton University Press.

Lave, Lester B. and Eugene P. Seskin (1977). *Air Pollution and Human Health*. Baltimore: Johns Hopkins University Press, for Resources for the Future.

Leshy, John D. (1987). *The Mining Law: A Study in Perpetual Motion*. Washington: Resources for the Future.

Levhari, David and Nissan Liviatan (1977). "Notes on Hotelling's Economics of Exhaustible Resources." *Canadian Journal of Economics* 10 (2) (May): 177-192.

Lindzen, Richard S. (1992). "Global Warming: The Origin and Nature of the Alleged Scientific Consensus." *Regulation* 15 (2) (Spring): 87-98.

Lott, John R. Jr. (1992). "Goring the U.S. Economy." (Review of Senator Al Gore, *Earth in the Balance: Ecology and the Human Spirit.*). *Regulation* 15 (3) (Summer): 76-80.

Mancke, Richard B. (1974). *The Failure of U.S. Energy Policy*. New York, NY: Columbia University Press.

_________(1976). *Squeaking By: U.S. Energy Policy Since the Embargo*. New York, NY: Columbia University Press.

Mead, Walter J. et al. (1985). *Offshore Lands: Oil and Gas Leasing and Conservation on the Outer Continental Shelf*. San Francisco: Pacific Institute for Public Policy Research.

Michaels, Patrick J. (1992). *Sound and Fury: The Science and Politics of Global Warming*. Washington: The Cato Institute.

Modiano, Eduardo M. and Jeremy F. Shapiro (1980). "A Dynamic Optimization Model of Depletable Resources." *The Bell Journal of Economics* 11 (1) (Spring): 212-36.

Moorhouse, John C. ed., (1986). *Electric Power: Deregulation and the Public Interest*. San Francisco, Calif.: Pacific Research Institute for Public Policy, 447-75.

Muth, John F. (1961). "Rational Expectations and the Theory of Price Movements." *Econometrica*, 29:3 (July), p. 315-35. Reprinted in Robert E. Lucas, Jr., and Thomas J. Sargent, eds., 1981, *Rational Expectations and Econometric Practice*. Minneapolis, Minn.: University of Minnesota Press.

Nelson, Robert H. (1983). *The Making of Federal Coal Policy*. Durham: Duke University Press.

Portney, Paul R., ed., (1990). *Public Policies for Environmental Protection*. Washington: Resources for the Future.

Ramsay, William (1979). *Unpaid Costs of Electrical Energy: Health and Environmental Impacts from Coal and Nuclear Power*. Baltimore: Johns Hopkins University Press for Resources for the Future.

Ray, Dixy Lee with Lou Guzzo (1990). *Trashing the Planet: How Science Can Help Us Deal with Acid Rain, Depletion of the Ozone, and Nuclear Waste (Among Other Things)*. Washington: Regnery Gateway.

Ray, Dixy Lee with Lou Guzzo (1993). *Environmental Overkill: Whatever Happened to Common Sense?* Washington: Regnery Gateway.

Samuelson, Paul A. (1957). "Intertemporal Price Equilibrium: A Prologue to the Theory of Speculation." *Weltwirschaftliches Archiv* 79 (2) (December): 181-219. Reprinted in Samuelson (1966). Vol. 2, 946-84.

_________(1966). *The Collected Scientific Papers of Paul A. Samuelson.*. Vol 1-2. Edited by Joseph E. Stiglitz. Cambridge, MA.: The MIT Press.

_________(1972). *The Collected Scientific Papers of Paul A. Samuelson*. Vol 3. Edited by Robert C. Merton. Cambridge, MA.: The MIT Press.

_________(1977). *The Collected Scientific Papers of Paul A. Samuelson*. Vol. 4. Edited by Hiroaki Nagatani and Kate Crowley. Cambridge, MA.: The MIT Press.

_________(1986). *The Collected Scientific Papers of Paul A. Samuelson*. Vol. 5. Edited by Kate Crowley. Cambridge, MA.: The MIT Press.

Schmalensee, Richard (1979). *The Control of Natural Monopolies*. Lexington, MA.: Lexington Books, D. C. Heath.

Scott, Anthony (1955). *Natural Resources: The Economics of Conservation*. Toronto: University of Toronto Press.

_________(1967). "The Theory of the Mine under Conditions of Certainty." in Mason Gaffney, ed., *Extractive Resources and Taxation*. Madison: University of Wisconsin Press, 25-62.

Smith Adam, 1776 , *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: W. Strahan and T. Cadell. Reprints including the 1979 version in the Glasgow edition of the works and correspondence of Adam Smith. Oxford: Oxford University Press and Indianapolis: Liberty Classics.

Solow, Robert M. (1974). "The Economics of Resources or the Resources of Economics." *American Economic Review*, papers and proceedings, 66 (2) (May): 1-14.

_________(1992). *An Almost Practical Step Toward Sustainability*. Washington: Resources for the Future.

Solow, Robert M. and Frederic Y. Wan (1976). "Extraction Costs in the Theory of Exhaustible Resources." *Bell Journal of Economics* 7 (2) (Autumn): 359-370.

Sweeney, James L. (1993). "Economic Theory of Depletable Resources: An Introduction." in Allen V. Kneese and James L. Sweeney, eds., *Handbook of Natural Resource and Energy Economics*. Vol. 3. Amsterdam: Elsevier.

Trocki, Linda K. (1986). "An Analysis of the Role of Exploration in the Opening of New Iron and Copper Mines." Unpublished Ph.D. thesis in Mineral Economics, The Pennsylvania State University.

Trocki, Linda K. (1990). "The Role of Exploration in Iron and Copper Supply." *Resources and Energy* 12 (4) (December): 321-38.

U.S. Congress (1990). *Clean Air Act Amendments of 1990*. Washington: U.S. Government Printing Office.

U.S. Congress, Office of Technology Assessment (1984). *Acid Rain and Transported Air Pollutants: Implications for Public Policy*. Washington: U.S. Government Printing Office.

Watkins, G. C. (1992). "The Hotelling Principle: Autobahn or Cul de Sac?" *The Energy Journal* 3 (1): 1-24.

Williamson, Oliver E. and Sidney G. Winter, eds., (1991). *The Nature of the Firm: Origins, Evolution and Development*. New York, NY: Oxford University Press.

Zimmermann, Erich W. (1964). *Erich W. Zimmermann's Introduction to World Resources*. Edited by Henry L. Hunker. New York: Harper & Row. (Updated and revised by the editor from Zimmermann's part I of *World Resources and Industries*.)

Zimmerman, Martin B. (1987). "The Evolution of Civilian Nuclear Power." in Richard L. Gordon, Henry D. Jacoby, and Martin B. Zimmerman, eds., *Energy: Markets and Regulation*. Cambridge, MA.: The MIT Press, 83-106.

Part IV
Markets Versus Governments

[21]

Energy Policy and the Role of the State in the Market for Energy

DIETER HELM, JOHN KAY and DAVID THOMPSON*

I. INTRODUCTION

Energy policy is now in a state of transition. The policy consensus which emerged in post-war Britain — based on integrated public monopolies in gas, coal and electricity, and central planning by government — was firmly rejected by the incoming Conservative administration in 1979. In its place was to be established a new market philosophy for the energy sector. The role of the public sector was to be reduced, first by encouraging competition in gas and electricity through the Oil and Gas (Enterprise) Act in 1982 and the Energy Act in 1983, and second by the privatisation of the public sector utilities (starting with British Gas in 1986).

Yet it is clear that this market philosophy has so far been applied to the energy sector in a limited and partial way. Coal seems likely to remain in the public sector for the indefinite future. Gas, although privatised, is still an integrated monopoly, subject to extensive regulation. The attempt to introduce competition in the electricity and gas industries has so far failed almost completely. Furthermore, the government retains an extensive network of controls over North Sea oil development and production.

Energy policy thus stands at a crossroads. Its future path could see further development toward a competitive 'market for energy' which Nigel Lawson set out as the policy goal in 1982. Electricity privatisation requires a clear choice between the claimed advantages of centralised co-ordination and management and the less certain outcomes of the free play of competitive market forces. If the competitive solution fails, or is not pursued, then British

---

* Dieter Helm is Research Fellow at the Centre for Business Strategy (CBS) and a Research Associate at the Institute for Fiscal Studies. John Kay is Director of CBS and Professor of Industrial Policy at the London Business School. David Thompson was formerly a Programme Director at the Institute for Fiscal Studies.
This article forms the introduction to a forthcoming volume edited by the present authors, to be published by Oxford University Press — *The Market for Energy*. The book will contain many of the individual articles referenced here and grew out of the project on energy at the Institute for Fiscal Studies. Thanks are due to the Economic and Social Research Council for funding.

Gas may provide the model for privately owned integrated energy monopolies. A change in the political climate, or growing dissatisfaction with the results of liberalisation, might subsequently lead to a return to planning and State control or ownership. It is clearly an important time at which to take stock and to consider the various policy options and their comparative *merits* and *demerits*.

In this article, we provide an overview of the key issues. We will distinguish two contrasting policies for the energy sector. The first, which we will call 'the Post-War Consensus', was pursued by governments of both parties in the 1950s, 1960s and 1970s. This approach to policy places emphasis on centralised planning of the allocation of resources in the energy sector. The second approach is 'the Market Philosophy' outlined by Nigel Lawson (1982). This approach emphasises the role of market forces in the determination of prices and in investment and production decisions, in the energy sector.

In order to compare the merits of these very different policies for the energy sector, we will use two familiar economic concepts — market failure and regulatory failure. Market failure arises when unregulated private markets fail to meet consumers' requirements with maximum efficiency. Section II considers the incidence of market failure in the energy sector and its relevance for policy-making. Regulatory failure arises when intervention by government in the operation of a market, perhaps to rectify a perceived market failure, has the effect of reducing efficiency. It will be clear that advocacy of the policies characterised by the Post-War Consensus requires a belief that market failure is endemic in the energy sector and has serious consequences for efficiency, while the possibility of regulatory failure is viewed with equanimity. Conversely the Market Philosophy assumes that market failure is not serious in the energy sector, or at least that it is less serious than the regulatory failures which have accompanied the scale and type of intervention implied by the Post-War Consensus. In Section III we consider regulatory failure and its importance in the publicly controlled energy sector of the Post-War Consensus. We go on to consider its likely incidence in a liberalised, and privately owned, energy sector. In Section IV we consider the current policy issues in the light of our analysis of market and regulatory failure, and draw our own conclusions in Section V.

II. ENERGY MARKET FAILURES

Most commodities in Britain are provided by private firms in competitive, unregulated markets. This has rarely been true of energy. A similar observation could be made in most countries in the western world. Why is it that governments have tended to think that markets which could be relied on to produce most other commodities could not deal adequately with the supply of energy? What is special about energy? In this section, we review the arguments in common, although often implicit, use. Some concern particular characteristics of the *demand* for energy; the majority, however, are aspects of energy *supply*.

There are, we believe, three main questions which those who are sceptical about the possibility of a market for energy would pose. First, energy is a particularly important commodity. Without it, individuals suffer acute physical distress and may even die; almost all industrial processes grind to a halt. The production of energy affects all other sectors of the economy to a degree that is characteristic of few, if any, other commodities. Can market forces give this adequate recognition?

Second, the time-scales associated with decisions about energy are exceptionally long. This is partly a result of the non-renewable character of many energy resources. It also follows from the sheer size and scope of energy projects — oilfield development and power-station construction are among the largest single investments in modern economies. Can market forces cope with the very long-term planning which these decisions require?

Third, even if a competitive market in energy were desirable, is it feasible? Many areas of energy supply seem to be natural monopolies — production by more than one firm is technically impracticable or would lead to wasteful duplication on a large scale. Ever since the Rockefellers sought to monopolise the US oil industry, energy production has been concentrated in the hands of a few major firms which have sought to influence the markets they face as well as to respond to them. No elementary textbook points to the energy industry to illustrate the workings of perfect competition.

1. Can Market Forces Take Proper Account of the Importance of Energy to the Economy?

The analysis of competitive markets assumes that individual consumers are best placed to choose the goods and services they want. To do this, they must be well informed about the costs and characteristics of alternatives, capable of judging between them, and consistent in their behaviour. Each year, however, a number of elderly people die of hypothermia. It would be difficult for even the most fanatical admirer of the operation of market forces to argue that this outcome is the result of welfare-maximising behaviour. Moreover, the fact that the market fails in this disastrous way in a small number of cases must raise the possibility that it works less than perfectly in many more.

There are two distinct problems here. Consumers with resources adequate for their needs may use insufficient energy because they are poorly informed — about the price of energy, how to operate appliances, or their heating requirements. More generally, consumers may lack appropriate resources to achieve a minimum standard of living. Since energy comprises a substantial proportion of the household budgets of the poor, the pricing policy of energy utilities is likely to have a considerable impact on poverty. It has often been argued that the poor should be protected directly, through lower fixed or standing charges for connection, and lower unit prices especially at the peaks in winter. But this is an expensive and inefficient means of helping a small subset of poor consumers, and lack of income is better dealt with, as Dilnot and Helm (1987) suggest, through the social security system.

Security of energy supply is important to both domestic and industrial consumers. This requires that energy capacity should be available to levels in excess of normal requirements. But will profit-maximising suppliers of energy provide capacity which will rarely — and may never — be used? If they do, how will the costs be recovered? In principle, the costs of a security margin could be recouped by imposing extremely high charges when the spare capacity was brought into operation, and firms might be induced to provide such a margin by the prospect of the revenues they could gain in these circumstances. It is easy to see political and practical reasons why this is unlikely to happen, at least on the scale required.

There are two possible solutions to the security problem. One is that some regulation is implemented to require firms to provide appropriate levels of investment. The other is that the public sector itself provides the spare capacity. However, public provision of capacity to meet supply shortages affects the incentives offered to the private sector. Since the government will always ensure excess supply, prices — and hence private investment — will be depressed. This phenomenon is described in Helm and McGowan (1987).

The energy sector is a substantial proportion of gross domestic product (GDP), and oil in particular comprises a major part of the visible account of the balance of payments. Thus performance of the energy sector has a powerful effect on national economic performance. Energy policy in the North Sea oil and gas industries, the level of subsidy to British coal and the cost of building new power-stations have unavoidable macro-economic implications. The industrial relations problems of the coal industry have provoked repeated government interventions. The government has also been anxious to promote the development of nuclear power, partly in a (wholly unsuccessful) attempt to develop exportable advanced technology, partly with a view to weakening the bargaining power of the miners' union.

Energy may have effects on sectors of the economy other than through use. Externalities arise when the private costs of production and consumption are not equal to those of society, because costs or benefits spill over to those not directly involved. These social costs are typically considered to be large in the energy sector. They include the pollution effects of acid rain, the impact of nuclear risks on the general population, and the social consequences on miners and mining communities of declines in the coal industry.

There are two possible economic approaches to the existence of externalities, which we illustrate by reference to the acid rain example. The problem is that the Central Electricity Generating Board (CEGB) produces pollution for which it does not pay. Forestry and fishing in Sweden find that their costs of production are increased. Consequently their output falls. The first solution is the tax/subsidy method. In order to reduce the output of pollution, we tax the output of the power-station. To compensate the fisheries and forestry firms in Sweden, subsidy is paid out of these receipts. Hence power-stations reduce output and the level of activity in forestry and fishing rises. The alternative approach is to view pollution as evidence of a missing market. On this view, there is no market in pollution or in clean air, but such a market could be created. For example, if Sweden owned the right to clean air, it could force the CEGB to pay compensation for acid rain. Thus if the property rights in clean air are defined, the optimal level of pollution can be attained as a result of the trading that takes place when the polluter offers to pay compensation. Except in a few cases, the practical problems of implementing these solutions are obvious and it is not surprising that direct regulation of processes or output levels is generally preferred.

2. *Can the Market Cope with the Time-scales Implied in Energy Planning?*

Many people are concerned by the use of non-renewable sources of energy. While minerals are also non-renewable, they are not destroyed in use: oil and coal, once burnt, are never available again. It follows from this that decisions about energy use foreclose options otherwise available to future generations. Do markets take this into account?

Current decisions *do* reflect the interests of future generations, because an alternative to using resources now is to retain them in order to sell them, at a higher price, in the future. Conserving resources is an investment for the future and, as with any other investment, private firms will undertake it if it is profitable. The view that energy depletion policy necessarily requires intervention because the interests of future generations will otherwise be ignored is certainly mistaken, but it does not follow from that that they are considered to an appropriate extent.

In a competitive market, non-renewable resources will command a price above the cost of production or extraction, and that difference will increase at the real rate of interest, as shown by Newbery (1986) and Devereux (1988). If this did not hold, it would pay to deplete more now, rather than hold the resource back for future use. This rising price of natural resources generates a return on investment in conservation. It follows that if the market rate of return reflects the rate at which society would trade off present for future consumption — as it should and would in a competitive market — the competitive rate of resource depletion would also be the efficient rate.

Surprisingly, then, the fact of non-renewability does not, in itself, give rise to market failure. However, the assumptions of this model are sufficiently strained to preclude unqualified faith in the ability of markets to deal with resource depletion. The problems created for all public sector investment decisions by likely divergences between market interest rates and social time preference rates are well known. Current monopoly, or anticipated future monopoly, will distort depletion rates. Uncertainty about future ownership rights in the resource is a particular stimulus to rapid depletion.

Quite apart from the time-horizons implicit in depletion decisions, many investment projects in the energy sector have their effects over an extended period. The longer the investment time-horizon, the greater the degree of uncertainty over future returns. In itself, however, this is not an argument for intervention. Rather, the price of oil should reflect the additional risk. Indeed, future contracts should incorporate the risk premium in future prices of the output of the plan from the investment. In a perfect market, there would be a complete set of future contracts which perfectly incorporate these risk premiums. Failures in the investment decision arise either because future markets are incomplete or because the social discount rate deviates from that of the market. Surprisingly, given the degree of concern expressed about the uncertainty of future energy prices, there are few futures markets in the energy sector. The oil futures market is short-term and small in size. Long-term contracts are usually concerned with the mechanisms of supply rather than the reduction of price variables. In the electricity industry, Hammond, Helm and Thompson (1986) noted that the absence of future contracts beyond the annual setting of the private purchasing tariffs can adversely affect private sector investment. In addition, the discounting procedures adopted in CEGB planning do not correspond to those of the private sector (see Jones (1986) and Helm (1987b)).

One of the primary activities of the Department of Energy has been the monitoring and prediction of future levels of demand for energy (see Department of Energy (1983)). The justification for the DoE taking on a demand-modelling role is based on two premisses — that the government is instrumentally better informed than the market and that a single consistent set of forecasts dominates a more pluralistic approach. The private sector may possess inferior information upon which to base its investment decisions, when compared with that of government. This may be because there are economies of scale in empirical research (for example, in the use of forecasts from a large-scale demand-forecasting model), because the quality of research staff is better in government, or because the government possesses relevant 'inside' information about its own demand, other developments in the economy, or the plans of other firms.

The evidence for the above set of arguments is slim. The forecasting record of the British government has not been good or clearly superior to other attempts. An official view is likely to be destabilising if it is erroneous, since it will compound errors across the industries. Finally, if government failure tends to manifest itself in over-optimistic forecasting, the error is more likely to be an overestimate.

3. Can Energy Markets be Competitive?

The energy sector includes the largest monopolies in the UK economy — in coal, electricity and gas. The oil industry is dominated by a few very large firms. The existence and apparent inevitability of monopoly have proved a motive for intervention. Nationalisation was perceived as one solution to this problem, sufficient indeed to exempt these industries from the Monopolies and Mergers Commission up until the 1980 Competition Act.

A monopoly can be either *natural* or *artificial*. A natural monopoly arises where technical cost conditions are such that the industry can support only one firm in the industry. Artificial monopoly exists where, despite the technical possibility of entry, the single incumbent firm is protected from entry either by strategic barriers to entry or by statutory monopoly concessions.

Natural monopolies arise most frequently in networks — the electricity and gas grid, the North Sea oil pipes, the main telephone system, roads and railways. They can be either *local* (like Area Boards) or *national* (like the national gas and electricity grids and transmission system). When these exist, it is wasteful to duplicate provision and hence, for cost reasons, one producer is better than two or more. The problem that arises is that, in the absence of competitive pressure, the monopoly can exploit its dominant position by marking up prices. Furthermore there is little direct pressure to minimise costs.

Natural monopoly does not, however, necessarily coincide with the industry, and it tends to change over time. The national transmission system of the British Gas Corporation is a natural monopoly, but the retailing of domestic appliances is not (Hammond, Helm and Thompson (1985)). The electricity grid is a natural monopoly, but household wiring and sales of appliances are not. Thus the natural monopoly problem is typically confined to parts of the industry and not the whole. Consequently, the regulatory problem is not coextensive with the industry. (As we shall see below, this lack of coincidence gives rise to serious regulatory failure problems if a simple rule is applied to the industry as a whole.)

By contrast, *artificial monopoly* arises where dominant firms erect barriers to inhibit competition from rivals. Apart from the gas and electricity local and national networks, much of the energy industry has been characterised by this second type of monopoly, through a combination of statutory provisions and other entry barriers. There is little or no evidence of natural monopoly in coal and oil extraction and delivery, electricity generation and gas production. Auxiliary services, such as servicing, billing and appliance sales, are similarly potentially competitive.

These two different types of monopoly require different regulatory solutions. Natural monopoly is addressed through direct regulatory control of prices, output or rate of return, whilst artificial monopoly is remedied by competition policy, directed at reducing barriers to entry.

III. REGULATION FAILURE

In the previous section we identified the incidence and impact of market failure in the energy sector. The existence of market failure is, however, common to all markets. For intervention to be justified, these failures must not only be large, but must also be greater than those which result from government intervention. In this section we consider the problems associated with government intervention to rectify these market failures in the energy sector.

The policies which can be followed by a government which wishes to intervene where private markets fail can effectively be divided into two broad groups. In the first, private markets are replaced by public enterprises which are set the objective of directly following welfare-maximising policies. In the second, private firms are subject to regulatory constraints which aim to ensure that the pursuit of profit-maximising policies will yield efficient outcomes.

'Regulatory failure' arises where interventionist policies fail to remedy the market failures which they seek to correct or where intervention has unintended, adverse consequences for efficiency. In reality, the contrast between nationalisation and regulation can be drawn too strongly and the problems which arise are not very different in the two cases. The underlying causes of regulatory failure are related to objectives and to the availability of information (see Kay and Thompson (1987) for a more detailed discussion). Should public enterprises or regulatory bodies choose not to follow welfare-maximising objectives, then efficient outcomes will not be achieved. However, the successful specification of regulatory constraints to ensure the achievement of efficiency depends critically upon the information available to the regulated enterprise (whether public or private) and to the regulator. The policy problem can be characterised in a 'principal-agent' framework in which the principal (the regulatory authority or government department) relies on an agent (the utility) to achieve its objectives in circumstances where the objectives of principal and agent diverge and in which the two partners' access to information is asymmetrical (see, for example, Crew and Kleindorfer (1979) for an elaboration).

It can be seen that the problems of objectives and information interact. If the objectives of principal and agent coincide (that is, if the public enterprises and public regulatory bodies choose to follow welfare-maximising objectives) then the principal is likely to face good access to information but to have little need for it. If, alternatively, objectives diverge, then information is required to set regulatory rules but will not be (easily) available to the principal from the regulated enterprise. Where, as is the case for the UK public energy utilities, the enterprise is a monopoly, with a corresponding dominance in information and technical expertise, the information asymmetry can be acute. In what follows we will use this framework to assess the development of energy policy in the UK. We begin by looking at the performance of the nationalised energy utilities. Thereafter, we consider liberalisation and attempts to introduce competition with the State monopolies. Finally, we consider privatisation in the energy sector and the regulation of privatised energy utilities.

Policy and performance in the nationalised energy sector closely parallels those of the nationalised industries as a whole, not surprisingly given the importance of energy industries in the nationalised sector. In this section we trace the evolution of nationalised industry policy in outline only, before considering its application to the energy sectors. (For more general consideration of nationalised industry policy, and performance, see NEDO (1976), Pryke (1981) and Molyneux and Thompson (1987).)

Nationalisation by the Labour government between 1945 and 1951 was implemented on the basis of what is often called 'the Morrisonian model' (after Herbert Morrison). This established corporations which were publicly owned but which were separate from government and were intended to operate at arm's length from day-to-day political intervention. These corporations were typically given a national monopoly in the supply of goods and services.

The solution proposed to the perceived failure of private markets was to replace them with public corporations which, it was assumed, would seek to implement welfare-maximising policies. The crucial determinant of the success or failure of policy was thus whether, in the absence of any explicit constraints or incentives, the public corporations would choose to follow the efficiency rules which provided the rationale for their existence. The development of nationalised industry policy to the present day can be caricatured, not altogether unfairly, as a progressive recognition of the inherent improbability of this outcome.

This recognition is reflected in a sequence of White Papers which attempted to prescribe regulatory rules which would constrain the industries to act efficiently. The White Papers in 1961 and 1967 focused upon allocative efficiency. The level and structure of prices were to be related to marginal costs. The benefits of proposed investments, discounted by the opportunity cost of capital, were to be compared with the costs of the project. In contrast, the main focus of the 1978 White Paper was on productive efficiency — establishing the primacy of financial controls and introducing performance targets.

From the mid-1960s, then, the nationalised fuel industries were instructed on the steps to be followed to implement welfare-maximising policies. The procedure is admirably summarised in Posner (1973) and can be characterised as follows:

- prepare medium-term forecasts of the demand for energy and, within this, for component fuels;
- identify the investment paths required to meet this demand in each fuel industry;

- identify the efficient pricing policy for each fuel (using the marginal cost principles outlined in the White Papers) and check the consistency of the planned investment path;
- check the path of relative prices against the demand forecasts and iterate the process until the forecasts, prices and investment paths are consistent.

The function of prices, in this framework, is to provide a medium-term signal to consumers and a bench-mark against which to evaluate investment plans. Unanticipated short-term mismatches between demand and supply are remedied not through adjustment to prices but through the under-utilisation of capacity or through rationing (in practice usually the former for reasons discussed further below). Under this framework a medium-term view is taken by government of the likely future path of comparative advantage of the different fuels, and it is assumed that an orderly substitution can be implemented.

If this is implemented, then the consequences should be that consumers face prices which are stable in the short term and which give appropriate medium-term signals to inform consumer investment (in purchasing appropriate heating systems, cookers etc.). Energy supply is provided by the most efficient mix of fuels produced using the most efficient technology, and prices are set at efficient levels. This requires, however, that demand-forecasting is effective and the industries have incentives to implement the successive steps. In particular, they must produce the required level of output at efficient cost levels and set prices in relation to these costs.

Newbery (1986) assesses the success of energy planning by comparing forecasts prepared in 1973 for the year 1980 with the actual out-turn in that year. The differences between forecast and out-turn are striking and lead Newbery to conclude that policies built upon forecasts so prone to error are unlikely to be efficient. It is clear, however, that the technical standard of the Department of Energy's forecasts is high and that the forecasting models are able to track closely the relationship between the demand for energy and the underlying determinants of demand, in particular economic growth (Department of Energy (1983)). The forecasting failure identified by Newbery reflects both the sensitivity of energy demand to GDP growth (and the difficulty in accurately forecasting growth over the medium term) and a failure to achieve planned substitutions between energy sources (in particular, nuclear power for coal).

These problems are compounded by the probability that nationalised industry managers have concerns other than the maximisation of social welfare. There are many objectives which nationalised industry managers may choose to follow in preference to welfare maximisation, in particular output maximisation (Rees (1984)), expense preference (Williamson (1963)) and managerial slack.

Rees (1988) compares the performance of the electricity industry with that predicted by a model in which the firm maximises output subject to constraints, set by governments, concerning the minimum level of profits and the maximum level of capital expenditure, and where the allocation of labour resources constrains labour bargaining power. Key features of performance accord with the predictions of the theory: wages are above average (for relevant skill groups), reflecting the exercise of union bargaining power in a monopoly industry, and investment plans have been consistently, and substantially, over-optimistic.

More generally, a specification of the behavioural objectives of public enterprise managers in the framework of output maximisation or expense preference, subject to a government-imposed profit constraint, suggests that in periods when such constraints are weak, enterprises will fail to achieve both productive and allocative efficiency. In periods when the financial constraint is binding, however, this suggests that the achievement of technical efficiency will not be a serious problem and that the main failures will be allocative — both in relation to choice of production technique and in relation to pricing.

Pryke's (1981) assessment of the energy utilities' performance in the 1970s confirms that the weakening of financial constraints which resulted from the counter-inflation policy introduced in 1972 was followed by deteriorating performance in relation to productive efficiency. Analysis of developments since the 1978 White Paper, which elevated financial constraints to the centre for nationalised industry regulation, shows a sharp upturn in performance in relation to productive efficiency (see Molyneux and Thompson (1987)). In the case of gas, prices are set below efficient levels in a way which is consistent with output maximisation (see Hammond, Helm and Thompson (1985)). The over-ambitious investment plans of the electricity supply industry (see Rees (1988) and Jones (1986)) are also consistent with this goal. The output path in the coal industry, in which production has, until recently, continued at individual locations with supply costs well above current or prospective prices, also indicates significant allocative inefficiency (see Robinson (1988)). In these industries, however, there is little evidence of technical inefficiency (see Molyneux and Thompson (1987)). Most of these industries are, however, characterised by uniformities in the structure of prices (between different geographic regions or between time-periods) which fail to reflect variations in economic costs.

Thus regulation of public energy utilities in the UK has failed in a number of ways. The Morrisonian model — in which public enterprises are assumed to act as welfare maximisers — is obsolete. It is also clear that the system that replaced it — which essentially imposed rules or objectives but not constraints — also failed to ensure that productive or allocative efficiency was achieved. It does seem, however, that the redirection of nationalised industry policy in the 1978 White Paper — in which financial constraints and related efficiency targets were introduced — has had a beneficial impact, at least in relation to the achievement of productive efficiency.

The main efficiency failures in the energy sector are now allocative. The reasons for this relate both to the absence of product market competition and to specific regulatory failures. Particularly important are the inefficiencies in pricing and output paths in the coal industry which have been sustained from the earliest days of nationalisation because of the bargaining strength of the miners and because British Coal and the CEGB, its main customer, are both public sector monopolies.

Cross-subsidisation between energy industries (in particular from electricity to coal) has been accompanied by uneconomic cross-substitution within industries. Asymmetries in information between government regulators and monopoly industries have meant that requirements to relate the structure of prices to relative costs have been in some cases largely irrelevant. This summary thus suggests a fairly well-defined mix of successes and failures, and one which can be related directly back to the likely incidence of regulatory failure. This provides a frame of reference for analysing the recent reform of policy.

IV. THE NEW MARKET PHILOSOPHY

The *laissez-faire* economic philosophy of the incoming Conservative administration in 1979, and increasing dissatisfaction with the performance of public enterprises, led to the explicit attempt to opt for a market solution which abandoned the main features of the Post-War Consensus on energy policy. In 1982 Nigel Lawson, the then Secretary of State for Energy, set out the new objective — to create a market for energy.

1. What would a Market for Energy Look Like?

In a competitive energy market, *production* is carried out by many separately owned firms. There are no statutory restrictions on market entry other than general environmental planning requirements and those related to health and safety. A competitive market in the production of fuel is thereby established.

However, the infeasibility of competition in the *distribution* of some fuels (because natural monopoly exists, especially in electricity and gas) provides the opportunity for prices to be raised above efficient levels. Local distribution networks are separately owned, however, and a regulatory ceiling is fixed on the distribution charges for each area. This ceiling is set in a systematic relationship to existing charges in all other areas, thereby preventing significant exploitation of consumers, but also providing opportunities for distribution companies to increase profits by beating the average level of performance of all companies.

Distributional concerns relating to the energy sector are dealt with directly through the tax and benefit system. The development of natural resources is left to the market, but resource taxes are used to achieve the desired distribution of resource rents between producers, consumers and the government. External costs and benefits are dealt with through regulation (e.g. on emissions) or through specific taxes and subsidies.

Under this framework, prices perform a short-run allocative function. Consumers with a high preference for price stability can achieve this through purchasing of futures via long-term contracts. The path of output is determined by the investment decisions of the various market participants, which are based on their perceptions of the future path of costs and prices. The mix of fuels supplied is therefore essentially market-determined. Because these decisions are formulated in private markets, this policy framework is assumed to provide incentives which ensure that efficiency is achieved rather than simply planned for.

Evidently such a market approach is very different, in appearance and probably in consequence, from the structure which prevailed in 1980. It is also very different, in appearance and probably in consequence, from the structure which prevails in 1988.

The major legislative components of the new Lawson policy have been:

- the 1982 *Oil and Gas (Enterprise) Act* which ended British Gas's statutory monopoly in gas supply and distribution, forced the British Gas Corporation (BGC) to dispose of its oil interests (as Enterprise Oil) and provided for common-carrier provision and hence competitive entry;
- the 1983 *Energy Act* which extended the principles of the 1982 Act to the electricity supply industry. Most notably, it provided for the compulsory publication of the prices that Area Boards would pay for privately generated electricity (the private purchase tariffs) and tariffs for rent of the network;
- the 1986 *Gas Act* which privatised BGC, transferring it to the private sector as a single company subject to regulation by a newly created authority, OFGAS (the Office of Gas Supply).

We will consider how, and in what ways, these legislative initiatives have fallen short of the creation of a market for energy in two stages. First we consider the attempts to introduce competition into the production of energy, and then we consider the regulation of privatised energy utilities.

The central policy dilemma in creating a competitive market was and is the selection of an appropriate liberalising strategy. Much of the energy sector remains characterised by natural and artificial monopoly. Whilst the former type of monopoly requires careful regulation, the latter demands attention to the enhancement of competitive pressure: markets need to be liberalised and the terms of entry for rivals set to prevent artificial obstruction from barriers to entry.

Thus, in order to appreciate the impact of the 1982 Oil and Gas (Enterprise) Act and the 1983 Energy Act, we need to look at the underlying entry-preventing strategies.

Any liberalising strategy must be based on a prior view as to the degree of potential entry and the sorts of entry barriers which may inhibit it. As we noted above, the energy sector industries combine elements of both natural and artificial monopoly. Competition is possible only in the potentially competitive segments of the industry, and the natural monopoly elements should therefore be separated and subjected to regulation.

Non-natural monopoly elements are open to competitive entry. The dominant incumbent can, however, employ a number of strategies designed to create and sustain an artificial monopoly. In addition to the statutory barriers to entry, recent industrial theory has highlighted a number of strategic activities by which dominant firms may deter rivals. As applied to the electricity supply industry in the UK, Hammond, Helm and Thompson (1986) show how these strategies may deter entrants and hence may have undermined the liberalisation intentions of the 1983 Energy Act. The principal barriers they identified are excess capacity, entry costs, the bankruptcy constraint, and the strategic choice of objectives.

Where an industry is characterised by excess capacity, the potential entrant alters its expectation about the response rivals will give to entry. It will decide to enter only if the rivals' response is sufficiently muted to leave profitable opportunities for the entrant. If the incumbent has spare capacity, the cost to it of a loss of market share will be greater than if it were at full capacity. Thus entry is more likely to produce retaliation, and hence make entry less attractive.

Entry costs are of two forms — fixed costs and sunk costs. Fixed costs were dealt with under natural monopoly. Sunk costs are ones which are irrecoverable to the entrant, should it subsequently decide to leave the market. As these increase, the incentive to enter declines because the size of the initial investment at risk, and hence the costs of failure, rise. In the energy sector, the principal sunk costs relate to planning the entry decision, acquisition of energy skills, and the imperfection of the second-hand market for pipes, generating capacity, and so on. Much investment in the energy sector is specific.

The bankruptcy constraint is of considerable importance in the energy sector, because of the considerable presence of the public sector, and thus the fact that debt is underwritten. The potential entrant perceives that in a price war, the public sector incumbent is better able to withstand the short-term loss of profitability and hence is more likely to win. The existence of financial resources and a weak bankruptcy constraint is therefore a credible threat to the entrant that the incumbent would find it worthwhile to challenge the entrant, and hence reduces the entrant's expected return. This financial or bankruptcy barrier to entry is reinforced by another aspect of State ownership, namely the impact of managerial objectives on rivals' entry decisions. As we noted above, the separation of ownership and control encourages the presence of managerial rather than profit objectives. Following Rees (1984), we would expect an output-maximiser to lower prices and total profit to gain extra market share. This has a dual impact on potential entrants. The potential profit to the entrant is lowered because the nationalised firm lowers the general level of prices in the industry *and* is more likely to retaliate to loss of market share to the entrant, because market share more directly affects output than it does profits.

The 1983 Act did force the industry to publish the prices which would be paid to private sector producers (private purchase tariffs — PPTs) and the prices to be charged for the rental of the network (the common-carrier or network charge). The three problems with this liberalisation were the setting of the tariff levels by the dominant incumbent, the absence of longer-term contracts, and the failure to set up an independent body to review the administration of these tariffs. The fundamental impact of public ownership was not altered by the 1983 Act: control of the tariffs enabled the excess capacity to remain, and its costs to be passed on to consumers as higher prices without encouraging entry, because entrants received only 'avoidable' variable costs. New entry is rarely promoted by letting the dominant incumbent set its rivals' price.

The dominance that the incumbents exercised could be addressed through restructuring. Such an opportunity presented itself in the privatisation of BGC. The options for breaking up the industry were set out by Hammond, Helm and Thompson (1985), but due to political pressure, the desire to maximise sale returns, and the crucial role of management, the industry was not broken up. It remains to be seen how much restructuring takes place in the electricity supply industry.

Enhancing competition through liberalisation thus achieved very little in the UK electricity and gas industries, and government attention shifted towards improving the productive efficiency of the existing dominant firms. The government argued, as we noted above, that changing the ownership of nationalised industries would in itself improve efficiency, subject to an appropriate regulatory structure. So far, a detailed regime has been developed only for British Gas.

The regulation of private energy utilities in the United States, where private ownership of energy production is most common, has been subject to well-known difficulties. The most usual regulatory instrument — rate-of-return, or rate-base, regulation — involves the specification of price ceilings which are based on the enterprises' actual costs and which provide for a pre-specified 'fair' return on the enterprises' capital assets. This system provides no additional profit in improving efficiency and reducing costs. Conversely, any increases in costs can be passed directly on to consumers in higher prices. The efficiency incentives usually associated with private ownership are therefore almost completely eliminated. Furthermore, because the price ceiling is determined to provide a specified return on capital assets, there is an incentive for profit-maximising firms to adopt production techniques which are too capital-intensive (the Averch-Johnson (1962) effect). There may also be adverse incentives for the efficiency of pricing policy with incentives to under-price capital-intensive (for example peak) outputs (see Sherman and Visscher (1982)).

This regulatory failure is reflected in studies which have compared the costs and efficiency of regulated private utilities with those of similar publicly owned enterprises. Generally the findings of these studies show either no systematic differences in performance between the public and private sector or a differential which marginally favours the publicly owned enterprises (for reviews, see Millward (1982), Domberger and Piggot (1986) and Yarrow (1986)).

The underlying cause of this regulatory failure lies in the asymmetry in information between the regulatory authority and the regulated enterprise. If the regulatory authority knew what level of costs constituted efficient performance by the enterprise, then it would be able to structure the regulated price ceiling accordingly. In the absence of this information, prices are regulated in relation to achieved costs, with the adverse consequences for efficiency which have been noted. It can be seen that the nature of this regulatory failure — asymmetries of information in the achievement of efficiency between a monopoly enterprise and the relevant regulatory authority — is essentially the same as that discussed earlier in relation to UK nationalised industries.

The regulation of newly privatised enterprises in the UK has, however, been founded on an ostensibly different basis. The 'RPI-*X* formula', as it has become known, was recommended in the Littlechild (1983) report for the regulation of the privatised British Telecom. It has subsequently been adopted in the cases of the British Airports Authority and British Gas.

The details of the British Gas formula ('RPI-*X* + *Y*') are complex but the essential elements are straightforward (see Helm (1987a) and Price (1988)). The formula places a ceiling on the prices which BGC is permitted to charge its domestic customers. This ceiling allows BGC to pass on directly to customers any increase in the cost at which it purchases gas (*Y*) but requires that its non-gas costs (that is, labour costs, expenditure on materials, etc.) can only increase by the general rate of inflation (RPI) minus a factor *X*.

The basic principle underlying the 'RPI-*X*' system is that, with revenues constrained by the formula, the regulated enterprise's profitability will be determined by how effectively it controls its costs. There are thus rewards and penalties in relation to the achievement of efficiency which are largely absent in the case of rate-base regulation.

In theory, at least, the management of an enterprise which failed to achieve productive efficiency would find itself replaced (via take-over or at a shareholders' meeting through the actions of shareholders seeking to maximise their return).

The proper test of this regulatory system will lie ultimately in the performance (in terms of costs and efficiency) of the newly privatised enterprises. It will clearly be some years before any, even preliminary, verdict can be reached on this. Nevertheless, it is already clear that the practical implementation of RPI-*X* falls some way short of the idealised conditions necessary to generate the favourable incentives to efficiency discussed above.

These enterprises are not profit-maximising in any usual sense. The limits to the enforcement of profit maximisation are well known (see, for example, Helm (1988)), and in the case of the newly privatised utilities their size, market power, diffusion of shareholding and immunity from take-over make them very different from the textbook model of the profit-maximising firms.

Information necessary to assess efficiency appears, if anything, less readily available than before. If neither shareholders nor customers can monitor this, however, then the effectiveness of the RPI-*X* system will turn on whether or not the regulatory authority can determine what constitutes an efficient level of performance, and may design the regulatory price ceiling accordingly. In capital-intensive industries, however, if the price ceiling is set too 'tightly' then the consequence is likely to be underinvestment and, eventually, supply failure (see Yarrow (1988) for a discussion).

The regulator's task is thus not straightforward. Yet it is clear that unless the regulatory authority can form a judgement on the enterprise's efficiency which is largely independent of actual costs and performance, then the RPI-*X* system starts to become very similar, in practice, to rate-of-return regulation, with all its associated weaknesses. This will be particularly likely if there is concern to avoid supply failure in capital-intensive industries.

One method of generating such an independent judgement is to make comparisons of the performance of particular activities in different geographic areas. Yarrow (1988) outlines a system in which each local distribution network for electricity supply (the twelve Area Boards) is under separate ownership. The regulatory price ceiling applied in each area can be established by reference to the costs and performance of the other Area Boards. Similarly shareholders will be provided with comparative information on their company's performance. Effectively the system establishes 'yardstick' competition (Shleifer (1985)) in which, although each distribution network is a natural monopoly, the regulatory framework requires each company to match the efficiency of other distribution companies if it is to maintain normal levels of profitability.

This 'regionalised' solution was not, however, adopted in the case of British Gas. Nor does it appear that the regulatory framework has established the information base upon which such comparative analysis could be carried out. There must therefore be very limited grounds for optimism that the regulatory system established for the privatised BGC will be effective in improving the efficiency of its activities.

The new Market Philosophy has thus far failed in its ambitions. The energy sector has not seen significantly enhanced competitive pressure, and the newly privatised British Gas has survived with its integrated monopoly intact, and subject to less control than when in the public sector.

But the fact that the new policy has failed does not imply that it is infeasible. Many other factors have steered the privatisation programme away from the competitive objectives (see Kay and Thompson (1986)). In the next section we suggest a number of factors which would enhance the liberalisation strategy and form the components of a more competitive energy policy.

V THE WAY FORWARD

In this article, we have reviewed the institutional background and regulatory framework which have evolved in the post-war period, and we have examined the policy changes instituted since 1979. We have suggested that the rationale of energy policy lies in the identification of underlying market failures and the corresponding government failures that arise in corrective policies. The success of energy policy depends upon a proper analysis of such failures and a policy designed to intervene in the light of such failures.

In analysing the market for energy, we have focused on the two types of major failure — natural monopoly and artificial monopoly. The appropriate policy responses were then set out — regulation to prevent abuse of natural monopoly, and competition policy to ensure fair entry terms where competition is feasible. The creation of a market for energy depends upon the setting of these two policies.

Though it has been thought that nationalisation 'solves' natural monopoly by replacing profit maximisation by the pursuit of social welfare, it must now be relatively uncontroversial to claim that it in fact 'solves' very little in itself. The naïve view which dominated thinking in the early post-war period was clearly mistaken, and the painful attempts at control in the 1960s and 1970s reinforced this observation. Only effective regulation can mitigate the abuse of natural market dominance.

If ownership change via nationalisation failed to 'solve' the monopoly problem, the lesson has hardly been learnt for privatisation. Fundamentally, the privatisation debate in the UK has focused on the wrong question — the appropriate competitive and regulatory structure is a more important determinant of performance than ownership. The structure of regulation and competition is the central issue, not a secondary consequence of privatisation.

The important issue for monopoly is regulation. RPI-$X(+Y)$, we noted, suffers from a number of serious drawbacks. The emphasis on prices rather than costs is more apparent than real, and freedom to set individual prices to exploit monopoly power or support predatory intentions remains. Performance would be improved by placing targets on monopoly sectors as separate cost centres, rather than on the industry as a whole.

This location of targets gives rise to two further regulatory questions — how much of what sort of information ought to be produced, and what the optimal structure of the firm should be. Cost and profit centres should be located according to the underlying characteristics of the industry. Natural monopoly elements should be separated from potentially competitive ones, and the latter broken up into competing units. Hence individual coal-pits, individual power-stations and individual oilfields ought to be separated. For the gas and electricity industries, the national network/grid should form a separate company, and the Area Boards each additional companies.

Information for regulation produced competitively gives regulators appropriate 'yardsticks' by which to compare differential performance. Thus the electricity regulators could compare Area-Boards, pit costs could be contrasted, and gas showrooms ranked by performance. The regulator's need for information gives rise to the potential for his capture by the industry upon which he relies for his information. The wider the range of sources, the greater the independence and effectiveness of the regulator. If the regulator himself is independent of the industry, his position is enhanced. Allowing the Department of Energy to function as the sole monitor of performance is one of the central drawbacks of nationalisation.

Thus a market for energy requires identified inescapable monopoly elements to be set up, ideally as separate firms, but at least with separate accounts. It employs targeted price regulation, set and monitored by independent regulators.

We have argued that monopoly regulation needs to be supplemented by appropriate measures to stimulate competitive entry, for it is competition, not ownership, which is most likely to stimulate efficiency gains. Competition does not, however, arise spontaneously. Where the incumbent retains substantial market power, regulation for entry is required. The regulator needs to set and referee entry conditions to ensure the incumbent does not create and exploit strategic advantages over rivals. This latter point is especially important where the incumbent is large — e.g. British Coal, the CEGB and BGC.

Competition in the energy sector typically comes from three sources — within the industries, between the industries, and internationally. Current energy policy has failed to give adequate incentives and safeguards to entrants in each of these areas. Competition within the industry depends upon the actual and expected level, structure and revision of prices. These in turn depend on the institutional method by which they are set and revised. As with monopoly regulation, independent rather than incumbent methods are more conducive to competition.

Competition between industries is typically unregulated, and argued to be intense in certain markets. Indeed, it is this sort of competition, or its absence, which has been the gauge of monopoly and the need for regulation. Thus the domestic gas market is regulated, but not the industrial one. However, a precondition for effective competition is consistent regulation of each of the industries considered separately. If, for example, the coal industry charges above marginal costs to the electricity industry, and the electricity industry ups marginal cost pricing taking the distorted coal price as exogenous, and the gas industry charges below marginal costs, then it follows that electricity/gas competition in the industrial market is unlikely to allocate resources efficiently. The criterion for inter-industry competition is

٤٣٢

consistency between individual industry regulation. Given, therefore, the framework devised for British Gas, the options for electricity regulation are heavily constrained.

Competition may also be international. Increasingly, electricity and gas are becoming traded European commodities. Firms and utilities would, in a liberalised energy market, choose the cheapest source of supply. In electricity, in practice, this implies that France and Scotland could provide powerful checks on monopoly profit for an English CEGB. The key to competition here is access to a network of common-carrier provisions.

A competitive energy policy thus has three components — a structure which follows the natural characteristics of an industry and its market failures; a regulatory system to control natural monopoly elements and enhance entry conditions; and an information system to allow the monitoring of performance. The choice of industrial structure strongly influences the required degree of regulation for competition and quality of information. The less competitive the structure, the greater the need for compensating regulation for competition. To date, restructuring has been largely avoided, RPI-*X* suffers from a number of identifiable drawbacks, entry conditions have been inadequately addressed, and the dominant firms have effective informational monopoly. The government has a long way still to go to attain a competitive 'market for energy'.

REFERENCES

Averch, H. and Johnson, L. (1962), 'Behaviour of the firm under regulatory constraint', *American Economic Review*, vol. 52, pp.1052-1069.

Crew, M.A. and Kleindorfer, P.R. (1979), *Public Utility Economics*, London: Macmillan.

Department of Energy (1983), 'Energy projections methodology', DEN/S/7(NE), Sizewell B Inquiry.

Devereux, M.P. (1988), 'The British experience of taxing oil extraction', in D.R. Helm, J.A. Kay and D.J. Thompson (eds), *The Market for Energy*, Oxford University Press, forthcoming.

Dilnot, A.W. and Helm, D.R. (1987), 'Energy policy, merit goods and social security', *Fiscal Studies*, vol.8, no.3, pp.29-48.

Domberger, S. and Piggott, J. (1986), 'Privatisation policies and public enterprise: a survey', *Economic Record*, vol.62, pp.145-162.

Hammond, E.M., Helm, D.R. and Thompson, D.J. (1985), 'British Gas: options for privatisation', *Fiscal Studies*, vol.6, no.4, pp.1-20.

—, Helm, D.R. and Thompson, D.J. (1986), 'Competition in electricity supply: has the Energy Act failed?', *Fiscal Studies*, vol.7, no.1, pp.11-33.

Helm, D.R. (1987a), 'RPI-X', *Public Money*, June.

— (1987b), 'Nuclear power and the privatisation of electricity generation', *Fiscal Studies*, vol.8, no.4, pp.69-73.

— (1988), 'Mergers, take-overs, and the enforcement of profit maximization', in J.A. Fairburn and J.A. Kay (eds), *Mergers and Merger Policy*, Oxford University Press, forthcoming.

— and McGowan, F. (1987), 'Electricity supply in Europe: lessons for the UK', Institute for Fiscal Studies, Working Paper no.87/10.

105

Energy Policy and the Role of the State in the Market for Energy

HM Treasury (1961), *Financial and Economic Obligations of the Nationalised Industries*, Cmnd 1337, London: HMSO.

— (1967), *Nationalised Industries: A Review of Economic and Financial Objectives*, Cmnd 3437, London: HMSO.

— (1978), *The Nationalised Industries*, Cmnd 7131, London: HMSO.

Jones, I. (1986), 'The application of risk analysis to the appraisal of optional investment in the electricity supply industry', in D.R. Helm, J.A. Kay and D.J. Thompson (eds), *The Market for Energy*, Oxford University Press, forthcoming.

Kay, J.A. and Thomspon, D.J. (1986), 'Privatisation: a policy in search of a rationale', *Economic Journal*, March.

— and Thompson, D.J. (1987), 'Policy for industry', in R. Dornbusch and R. Layard (eds), *Performance of the British Economy*, Oxford: Clarendon Press.

Lawson, N. (1982), Speech on energy policy, Fourth Annual International Conference of the International Association of Energy Economists, 28 June, Churchill College, Cambridge.

Littlechild, S.C. (1983), *Regulation of British Telecommunications' Profitability*, London: Department of Industry.

Millward, R. (1982), 'The comparative performance of public and private enterprise', in Lord Roll (ed.) *The Mixed Economy*, London: Macmillan.

Molyneux, R. and Thompson, D.J. (1987), 'Nationalised industry performance: still third-rate?', *Fiscal Studies*, vol.8, no.1, pp.48-82.

NEDO (1976), *A Study of UK Nationalized Industries: Their Role in the Economy and Control in the Future*, London: NEDO.

Newbery, D. (1986), 'Energy policy issues after privatisation', Centre for Economic Policy Research, Discussion Paper no.109.

Posner, M. (1973), *Fuel Policy*, London: Macmillan.

Price, C. (1988), 'Gas privatization: effects on pricing policy', in D.R. Helm, J.A. Kay and D.J. Thompson (eds), *The Market for Energy*, Oxford University Press, forthcoming.

Pryke, R. (1981), *The Nationalised Industries: Policies and Performance since 1968*, Oxford: Martin Robertson.

Rees, R. (1984), 'A positive theory of public enterprise', in M. Marchand, D. Pestieau and H. Tulkens, *The Performance of Public Enterprises*, Amsterdam: North-Holland.

— (1988), 'Modelling public enterprise performance', in D.R. Helm, J.A. Kay and D.J. Thompson (eds), *The Market for Energy*, Oxford University Press, forthcoming.

Robinson, P.W. (1988), 'The economics of coal', in D.R. Helm, J.A. Kay and D.J. Thompson (eds), *The Market for Energy*, Oxford University Press, forthcoming.

Sherman, R. and Visscher, M. (1982), 'Rate of return regulation and two-part tariffs', *Quarterly Journal of Economics*, vol.97, pp.27-42.

Shleifer, A. (1985), 'A theory of yardstick competition', *Rand Journal of Economics*, vol.16, no.3, pp.319-327.

Williamson, O.E. (1963), 'Managerial discretion and business behaviour', in R. Marris and A. Wood, *The Corporate Economy*, London: Macmillan.

Yarrow, G.K. (1986), 'Privatization in theory and practice', *Economic Policy*, no.2, pp.323-364.

— (1988), 'Regulatory issues in the electricity supply industry', in D.R. Helm, J.A. Kay and D.J. Thompson (eds), *The Market for Energy*, Oxford University Press, forthcoming.

106

[22]

The Case for an Energy Policy

DESPITE PAST EXPERIENCE, most proponents of energy policy seem to think it so obviously desirable as not to require explicit defence or even definition. They take refuge in generalities such as the supposed 'strategic' characteristics of energy or the supposed responsibility of government for the security of energy supply or, more recently, the necessity for government to take into account the environmental effects associated with energy production, transportation and consumption.

Occasionally, the case appears in more respectable form, discussing a number of externalities which markets are supposed to fail to take into account. However, believers in energy policy mostly seem unaware of how real-world governments behave and they commit the sin of comparing the results of imperfect markets with the results they infer from the actions of perfect governments: the behaviour of a perfect government cannot, of course, be observed since there is no such thing, any more than there is a perfect market.

It is best to consider the case for an energy policy by examining first the general advantages of permitting markets to work and the disadvantages of government action. Then we can consider the extent to which markets might fail.

The Case for Markets

The case for leaving energy affairs mainly to markets has two strands. One is positive – that markets have such desirable characteristics that they have no close substitutes. The other is negative – the alternative of government action is so undesirable that it should be avoided wherever feasible.

To begin with the negative case, public choice theorists and Chicago-school empirical researchers have so clearly demonstrated the serious problems inherent in government action that the traditional *deus ex machina* case which used to be made by old-style welfare economists (that government could always be brought in to set to rights the mess markets were making) is surely untenable. That, however, does not prevent the old-fashioned case from being put forward by politicians, civil servants, Parliamentary committees, journalists and even some economists.

The Problems of Government Action

To summarise, the principal problems of government action are:[50]

- *First*, elections are an exchange of policies for votes but the market outcome is flawed because voters have little incentive to discover in detail what political parties will do when in office since the costs of finding such information are extremely high relative to the infinitesimal influence any individual has on the outcome of an election. It is, therefore, rational for voters to remain ignorant. The rationality of voter ignorance leads to sloganising, similar to that in advertising campaigns in oligopolistic (and particularly duopolistic) industries.

- *Second*, in representative political systems, governments, once in office, are the sole suppliers of policy: thus they have monopoly power and so the familiar problems of monopoly arise.

- *Third*, monopoly power in this case is particularly serious because governments can use that power to gain access to a very deep purse which consists of other people's money;

50 See, for example, Anthony Downs, *An Economic Theory of Democracy*, New York: Harper, 1957; Gordon Tullock, *The Vote Motive*, Hobart Paperback No.9, London: IEA, 1976; James M. Buchanan *et al*, *The Economics of Politics*, IEA Readings No.18, IEA, 1978; W C Mitchell, *Government As It Is*, Hobart Paper No.109, IEA, 1988; and Arthur Seldon, 'Politicians for or against the people', in Gerard Radnitzky and Hardy Bouillon, *Government: Servant or Master?*, Amsterdam-Atlanta Rodopi, 1993

… into financing their schemes.

- *Fourth*, governments are heavily influenced by producer pressure groups which appear able to deliver substantial numbers of votes; those groups respond rationally to politicised markets by lobbying – since any favours the government grants them will bring benefits to their members but will be paid for by others.

Government Failure

For such reasons, government failure is a serious problem. Indeed, it is so serious that, as Arthur Seldon has persuasively argued, there is something to be said for taking a risk on under-government.[51] The political market-place tends to lead to short-term decisions, taken in the hope of securing votes, on the basis of very poor quality information about the preferences of voters – because of the rationality of voter ignorance and because such information as does reach politicians is filtered by producer pressure groups which rush into the information vacuum (as we have seen they did in the British energy market in the post-war years). Moreover, because the system relies on majority voting it cannot provide for minorities in the ways that markets will do. The overall result is that, except by chance, the pursuit of political and bureaucratic self-interest will not be beneficial to the community (in contrast to the pursuit of self-interest in a competitive market).

The Reverse Invisible Hand

Of course, one does not have to assume that politicians and bureaucrats are malevolent to obtain this result, any more than one has to assume benevolence for the invisible hand to produce a socially desirable outcome from decisions which serve the self-interest of the decision-takers. Public choice

51. Seldon, *ibid.*, pp.3-21.

theorists would simply claim that people in the public sector are much the same as people in the private sector. But they have substantial monopoly power through which they can coerce and tax. They are not omniscient: they cannot define or discern the 'public interest' in any given case. Nor are they altruistic: they have their own interests to pursue. Consequently, as Milton Friedman has put it, there is a 'reverse invisible hand':

> 'People who intend to serve only the public interest are led by an invisible hand to serve private interests which was no part of their intention.'[52]

In this light, government is not the solution: it is the problem. It is quite extraordinary, given all the evidence about the way real-world governments behave, that so many people continue to recommend large-scale government action. Keynes would have seen the reason. It can be explained only by the continuing tendency of practical men to be the slaves of defunct economists.[53]

Markets and Discovery

The positive case for using markets rests primarily on the characteristics of markets as discovery processes. Whether or not markets are 'perfect' is irrelevant – there is nothing genuinely perfect or even particularly desirable about the perfect markets of economics textbooks. As Kirzner has said:

> 'What keeps the market process in motion is competition – *not* competition in the sense of "perfect competition", in which perfect knowledge is combined with very large numbers of buyers and sellers to generate a state of perennial equilibrium – but competition as the rivalrous activities of market participants trying

52 Milton Friedman, *Why Government is the Problem*, Stanford, CA: Hoover Institution, 1993, p 11.

53 'Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist.' (J.M.Keynes, *The General Theory of Employment, Interest and Money*, reprinted in *The Collected Writings of John Maynard Keynes*, Vol.VII, London: The Macmillan Press for the Royal Economic Society, 1973, p 383 (First Edn., Macmillan, 1936).)

to win profits by offering the market better opportunities than are currently available. The existence of rivalrous competition requires *not* large numbers of buyers and sellers but simply *freedom of entry*. The competitive market process occurs because equilibrium has not yet been attained.'[54]

To see the fundamental reasons why markets work beneficially, one has to take an 'Austrian' perspective like Kirzner's which, following Hayek, sees knowledge as essentially dispersed: by definition, it cannot be gathered together in the hands of a few clever people in Whitehall or elsewhere.

The Impossibility of Accurate Forecasting

An awkward problem which faces each person is that he or she must make decisions which, by definition, are about the future. Yet all experience teaches us that we cannot know the future. We are all ignorant: *we do not even know what we do not know*. This dilemma is, strictly speaking, insoluble: each person must make forecasts, explicit or implicit, in order to run his or her life, yet forecasting is impossible. A competitive market, however, helps to solve this awkward problem because it is a mechanism for producing information which otherwise would not be known. In Kirzner's words:

> 'The competitive market process is needed not only to mobilise existing knowledge, but also to generate awareness of opportunities whose very existence until now has been known to no-one at all.'[55]

Effects of the Market Process

As entrepreneurs seek new ways of doing things, as consumers seek new products, discoveries will be made which would not

---

54 'The Perils of Regulation: A Market Process Approach', in Israel M.Kirzner, *Discovery and the Capitalist Process*, Chicago: University of Chicago Press, 1985, p.130.

55 *Ibid.*, p.131.

otherwise have been made. The forecasts on which people base decisions are imperfect but competitive forecasting and decision-making, and the discovery process which comes into operation as decisions are made, generate new knowledge.

Moreover, the market will also co-ordinate actions which otherwise could not have been co-ordinated, it will stimulate efficiency to an extent otherwise unachievable, and it will allow a degree of freedom of choice otherwise unrealisable. It works essentially by transmitting information (principally via price signals) from consumers to producers and back again.

Consumers have the power of exit from suppliers which do not suit them and so enjoy greater security of supply and lower prices than when they are in the hands of monopolists (state or other). Minorities find that their wants are met in ways which a political market-place (which serves the majority) is incapable of reproducing. Producers find that efficiency standards are automatically set for them by the actions of competitors because they cannot afford to fall behind. Thus there is a constant stimulus to innovation and entrepreneurship which promotes economic progress.

A very significant corollary of the Austrian view is that, because markets are essentially discovery mechanisms, their results cannot be reproduced by regulation, government control or similar means. *The results of competitive markets can be achieved only by the process of discovery.*

More Rules for Regulators?

This apparently simple idea yields some very significant conclusions. For instance, take the present debate in Britain (which is very relevant to the energy industries) about whether there should be more rules for regulators and whether regulators should be brought together in groups for consistency in decision-making. From an Austrian perspective, laying down more rules for utility regulators or combining regulatory bodies into one are actions at best irrelevant and indeed more likely to make regulation worse than better. Regulators can never know what the outcome of a competitive market would have been, so providing them with tighter rules or co-ordinating the actions of one regulator with those of other regulators cannot help decision-making, though it may well make it more rigid

Minimising Regulation

Much more important is to minimise regulation, confining it to genuine 'natural monopolies' (which are very few and far between these days and anyway tend to be eroded by changes in technology, as in telecommunications). Where regulation at present exists but the regulatee is not naturally monopolistic (such as in electricity generation), temporary pro-competition regulation is the answer, leading to the withering away of the regulatory body over time as it is replaced by a competitive market.

Such arguments for markets are a long way from the old-fashioned textbook case that there is a state which can be described as a perfectly competitive market the results of which one should aim to simulate because it produces 'socially desirable' results. The competitive market is a process, not a state. In this imperfect world, we are far better off relying principally on decentralised market discovery processes than on centralised government procedures. If we want the freedom which comes from allowing a wide range of preferences to be met and we want efficiency in providing goods and services, markets are far superior to the political system. Government is the prime (though not necessarily the sole) instrument for providing law and order and national defence, establishing and maintaining property rights, promoting and sustaining competition and – quite important in the energy field where one of our oldest-established industries has for years been in decline – for tempering the effects of decline and providing a safety net for the disadvantaged. But most governments stray far outside such bounds for the reasons already given and the reverse invisible hand acts to make matters worse, not better. Many of the services provided by the state are not so much demanded as supplied.

Apparent Problems of Using Markets

However, many people clearly distrust market outcomes. Economists who see markets in the static terms in which many textbooks still portray them show this distrust; moreover, economists have some self-interest in playing down the advantages of markets because government action is job-creating as far as they are concerned. Non-economists, not surprisingly, find the rather subtle case for using markets difficult to grasp. Governments, of whatever party, have an interest in promoting the advantages of intervention and most of the Press swallow this case unthinkingly.

Most of 'the second-hand dealers in ideas'[56] have accepted that central planning has failed but they have not drawn the logical conclusion – that, for essentially the same reasons that central planning failed, most government intervention will also fail. The demonstrations many years ago by von Mises and Hayek that the informational and calculational requirements of centralised planning and forecasting are such that it is very unlikely ever to be socially beneficial (though it may, of course, provide benefits to those who do the planning) apply, *mutatis mutandis*, to lesser measures of government intervention.[57]

Markets may appear impersonal, lacking an obvious guiding hand and suffering from numerous 'imperfections'. There is a strong temptation to try to improve on them – as indeed one should if the object is to stimulate competition, for example by easing entry to markets. But many people are also tempted to think that if only *they* could make the plans they could improve substantially on the market outcome.

Energy – 'Too Important to Be Left to the Market'?

That is particularly so in energy which, it is frequently remarked, is 'too important to be left to the market'. There

---

56 F.A.Hayek, 'The Intellectuals and Socialism', *University of Chicago Law Review*, Vol.16, No.3, Spring 1949.

57 See, for example, Kirzner, *op. cit.*, especially p 136

are large numbers of would-be energy planners who would like to over-ride preferences as they appear in the market-place, imposing on others their own preferences – whether for more 'energy efficiency', for particular aspects of a cleaner environment or for protection of particular fuel industries. There is a strong authoritarian streak, in addition to self-interest, in most of the pleas for 'long-term co-ordinated' energy policies or for setting particular shares for given fuels.

What are the failures with which, it is alleged, energy markets cannot cope?[58]

Security of Supply

It is argued that government has a rôle in promoting security of supply because it has some public good aspects (not all the benefits can be appropriated by the provider). Many British government actions in the energy market have been justified on security grounds although, more often than not, security is just an excuse for actions taken for quite different reasons. In any case, the effects have been perverse. British Coal was protected for many years on 'security' grounds, *yet the main result was to promote insecurity by making British consumers of coal and electricity dependent on a single supplier of coal.*

The theoretical argument is in any case unsound. Both consumers and producers have interests in secure supplies and will naturally diversify to provide them. The main reason the market under-provides is that it is rigged by political action: suppliers know that if they anticipate and provide for emergencies they will not be allowed to appropriate the benefits (for instance, by increasing prices). Therefore they have no incentive to make provision. There may be a case for some government action to provide against emergencies (say, some excess stocking) but it tends to undermine commercial incentives for secure supplies.

---

58 For more discussion, see Eileen Marshall and Colin Robinson, *The Economics of Energy Self-Sufficiency*, London: Heinemann Educational Books, 1984, and Colin Robinson, 'Depletion Control in Theory and Practice', *Zeitschrift für Energie Wirtschaft*, 1/86, January 1986.

Balance of Payments

Balance-of-payments reasons have been invoked by British governments for protecting indigenous energy industries (mainly coal but also nuclear) against market forces. The argument is complete nonsense. It cannot improve the balance of payments, or any other indicator of economic performance to buy high-cost home-produced goods and services when cheaper imports are available.

Protection Against Rising Prices of Imported Fuels

Another common argument is that we cannot afford to become dependent on imported fuels, as a competitive market might dictate, because in the long run the prices of those fuels will rise. Since history shows that the long-run tendency of real fuel costs and prices is to decline,[59] the empirical evidence for this case is weak. Nor is the implicit case – that some clever people can foresee trends in fuel prices to which markets are blind – at all convincing. One might have thought that those who are risking their money would be better predictors than would-be planners.

There is a possible insurance-premium-type argument for support for indigenous fuels on the grounds that imported fuel prices *might* go very high in the future. But calculating the correct premium – so as to incorporate not only the probability of those higher prices but also 'society's' attitude towards risk – would be a daunting task.

Safeguarding Future Generations

Competitive markets are said to neglect the interests of future generations. For example, producers of a resource might discount the future at 'too high' a rate, leaving too little of that

---

59 A classic study by Barnett and Morse (H.J. Barnett and C. Morse, *Scarcity and Growth. The Economics of Natural Resource Scarcity*, Baltimore, MD Johns Hopkins Press, 1963) showed that the real costs of extracting most resources in the USA, including fuels, had fallen substantially in the previous one hundred years. Subsequent studies confirm these results, despite the temporary increase in costs and prices in the 1970s. The main reason for declining costs appears to be improvements in technology.

resource in the ground for their successors. Resource markets are imperfect and will never achieve 'optimum' rates of depletion (in the Pareto sense) except by chance.

But, in all such matters, one must ask: What is the alternative? Would it have been better if, in the second half of the 18th century some resource conservationists – ahead of their time – had persuaded the British government to keep coal resources in the ground? We would have had plenty of coal now but we would have missed the discovery process which resulted from the exploitation of Britain's coal resources. Natural resources, when used, are not lost but transformed into technological advances, other new knowledge and increased use of capital.

Moreover, when considering market 'short-termism', one must remember that most participants in markets have a much longer perspective than most politicians. Compared to the alternative of political action, it is simply not true that markets take a short-term view. Clever people who believe they can form a much better long-term view than can market participants, even to the extent of perceiving how to promote the interests of future generations, should be (but evidently are not) immensely rich. If markets really are so short-sighted, anyone able to see beyond the end of his or her nose could make a killing by using that long perspective to invest.

Environmental Externalities

It is claimed that markets fail to take full account of some environmental effects because they are externalities which fall on individuals and organisations other than those taking the offending action, and that they are uncompensated or under-compensated. Although the case is for a government environmental policy rather than a government energy policy, it is necessarily linked to the energy industries because of their significant environmental impact.

The case may appear overwhelming to those who think in market-failure terms. It is less so to those who recognise the extent of government failure and observe the problems to which government action in the energy field has led. A particular difficulty is that, if there is to be a successful environmental policy run by government, that government must submerge all its short-term political interests and concentrate on achieving benefits which will accrue in the long term when the originating administration will be long gone and a different party may well be in power. It seems unlikely political behaviour. Unfortunately, it seems more probable that governments will use 'green' concerns as excuses for actions they wanted to take anyway which have little to do with genuine concern about the environment. That seems to have been the case with the fuel tax increases in Mr Lamont's last Budget in March 1993.

In general, it seems preferable to use market forces to the fullest extent possible (though opinions will differ on what that extent is) rather than to rely on political action. The politicisation of environmental matters will bring a short-term perspective to environmental affairs and deliver excessive regulation by people who can impose costs confident that they will always fall on others. For that reason a greater effort is required to find market solutions to environmental problems, whether 'global' or local. As with security of supply issues, the reason why markets do not work well in environmental matters is often because they are not allowed to do so – in particular, because property rights are not always clearly enough defined so that owners can defend their property against damage by polluters as against others.[60] The main problems reside in so-called 'global' effects where scientific knowledge is very poor and which are not so obviously amenable to property rights solutions.

6. Summary and Conclusions

BRITISH ENERGY 'POLICY' has a poor record extending back over many years. Until very recently, there was an increasingly protectionist trend because policy was dominated by producer pressure groups which took advantage of the desire of

60 Terry L. Anderson and Donald R. Leal, *Free Market Environmentalism*, San Francisco: Pacific Research Institute, 1991

politicians to capture votes. Privatisation has to some extent reduced government activity in energy markets. Nevertheless, the privatisation schemes themselves were heavily influenced by the wishes of the producer groups concerned which had interests in common with governments anxious to raise substantial revenues and to widen share ownership. The consequence was limited liberalisation of product markets. Liberalisation was left mainly to industry regulators and the competition authorities, delaying the benefits of competition and, given the power of the incumbents, leading to awkward regulatory problems.

This *Occasional Paper* has emphasised the advantages of competitive markets as well as the disadvantages of government action. In the light of these advantages and disadvantages, a policy specifically for energy seems not only unnecessary but undesirable. Whatever excuses are given, in practice 'policy' has generally arisen from pursuit of short-term political interests, supported by the producer pressure groups which thrive and lobby government whenever a market is seen to be politicised. In such markets short-termism and excessive lobbying are inherent. Adjustment processes are hampered and participants in the market are induced to invest resources in influencing government rather than in trying to improve the efficiency of their activities.

In energy markets, as elsewhere, there is no sense in recommending policies, based on supposed 'imperfections' and 'failures' in markets, which would be unlikely to work even if we were governed by far-sighted saints and angels. Given the world in which we live, if the aims are increasing efficiency, the prevalence of long-term views and representation of the interests of consumers, the best (though 'imperfect') answer is to rely primarily on energy markets. Attempts by politicians and civil servants, no matter how well-meaning they may be, to provide policies for energy are much more likely to be a hindrance than a help.