

Cambridge Working Papers in Economics CWPE 0469



UNIVERSITY OF
CAMBRIDGE
Department of
Applied Economics

Electricity Liberalisation in Britain: the quest for a satisfactory wholesale market design

David Newbery



The
Cambridge-MIT
Institute

*Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research*

CMI Working Paper 64

Cambridge Working Papers in Economics CWPE



UNIVERSITY OF
CAMBRIDGE
Department of
Applied Economics

**not to be quoted without
permission**



The
Cambridge-MIT
Institute

*Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research*

CMI Working Paper Series

Electricity liberalisation in Britain: the quest for a satisfactory wholesale market design

David Newbery*

Britain was the exemplar of electricity market reform, demonstrating the importance of ownership unbundling and workable competition in generation and supply. Privatisation created de facto duopolies that supported increasing price-cost margins and induced excessive (English) entry. Concentration was ended by trading horizontal for vertical integration in subsequent mergers. Competition arrived just as the Pool was replaced by New Electricity Trading Arrangements (NETA) intended to address its claimed shortcomings. NETA cost over £700 million, and had ambiguous market impacts. Prices fell dramatically as a result of (pre-NETA) competition, generating companies withdrew plant, causing fears about security of supply and a subsequent widening of price-cost margins.

JEL: L94, D43

Key Words: electricity, liberalisation, market design, market power,

INTRODUCTION

The standard model of the electricity supply industry (ESI) in almost every country before liberalisation was an effectively vertically integrated franchise monopoly under either public ownership or cost-of-service regulation. Investment in generation and transmission were (in theory) chosen to deliver the least-cost expansion plan (subject to government energy policy on fuel mix and plant choice), financed by low-cost borrowing underwritten by the franchise revenue base. Britain was no exception, with the entire ESI under state ownership since nationalisation in 1947. The Central Electricity Generation Board (CEGB) owned all generation and transmission in the whole of England and Wales, selling bulk power to twelve Area Boards, responsible for distribution and supply (retailing). In Scotland, the North of Scotland Hydro-Electric Board (NSHEB) and the South of Scotland Electricity Board (SSEB) each held regional franchises that included generation, transmission, distribution and supply. The Government set the annual External Financial Limit restricting (publicly provided) borrowing, which in some years could be negative, implying a net dividend payment to the Treasury. The tariff structure was moderately sophisticated, with a two-part zonal Bulk Supply Tariff charging for capacity (of both generation and transmission), and variable costs (energy and regionally differentiated losses). Area Boards offered a variety of tariffs, with various forms of peak-hour capacity charges. While the pricing may have been sophisticated, investment planning, and particularly investment delivery, was poor, slow and costly, and there were few incentives to deliver cost efficiency.

Liberalising and restructuring the ESI was intended to replace this command and control structure with its regulated charges by a decentralised market-driven system that would nevertheless deliver secure, reliable electricity efficiently and at competitive prices. At the time

* Professor, Department of Applied Economics, University of Cambridge, Sidgwick Avenue, Cambridge, England CB3 9DE (E-mail: dmgn@econ.cam.ac.uk). I am indebted to Karsten Neuhoff and Alex Henney for helpful comments.

the Government decided to restructure and privatise the ESI, there were few models available. The United States had evolved the contractual form of investor-owned franchise monopolies under state cost-of-service regulation that was criticised for poor incentives, stranded investments, and in some states, high prices. Chile, on the other side of the planet, had been reforming, restructuring and gradually privatising its ESI since 1978, while Norway already operated a spot market for club members for wholesale energy, but remained publicly owned. With no obvious model to follow, considerable political pressure to deliver a competitive outcome (in contrast to the earlier privatisations of telecoms, gas and water), and a tight timetable, the challenge was to design a set of markets and institutions to deliver these objectives. Of comparable importance, the design had to allow a smooth and predictable transition to a market-based system not just for electricity, but for the nationalised coal industry, three-quarters of whose (largely uneconomic) output was sold to the ESI, which in turn depended on coal for three-quarters of its output. The generation and distribution companies were to be sold to the general public and therefore needed predictable revenues on which they could be valued.

2 RESTRUCTURING AND PRIVATISATION

Two different solutions were adopted for Britain (and yet another for Northern Ireland). The *Electricity Act 1989* created the post of the Director General of Electricity Supply, the DGES, to regulate the natural monopoly wires businesses of the National Grid Company (NGC) and the Regional Electricity Companies, and to set price caps, which would be reset at periodic reviews every 4-5 years. He had a duty to ensure that reasonable demands for electricity were met, that licence holders were able to finance their activities, to promote competition in generation and supply, to protect customer interests, and to promote efficiency. The Office of Electricity Regulation, Offer, was set up by the Government as an independent body under the *Electricity Act*, headed by the DGES.

In England and Wales, the *Electricity Act* divided the CEGB, with its 74 power stations and the national grid, into four companies. Sixty per cent of conventional generating capacity (40 power stations with 30 GW capacity) were placed in National Power, and the remainder (23 stations of 20 GW) were placed in PowerGen. The original plan was to place the 12 nuclear stations with 8 GW in National Power, which had been given the bulk of the fossil generation in the hope that its resulting size would be financially viable. At a late stage the financial advisors made it clear the nuclear stations were not saleable at a reasonable price. They were transferred to Nuclear Electric and kept in public ownership until 1996. The high-tension grid, together with 2 GW of pumped-storage generation,¹ were transferred to the National Grid Company (NGC). These four companies were vested (i.e. created) as public limited companies (plcs) on March 31st 1990, at the same time as the twelve distribution companies, now known as the Regional Electricity Companies (RECs). NGC was transferred to the joint ownership of the RECs, and the RECs were sold to the public in December 1990. The pumped-storage generation of NGC

¹ Turbines pump water up to a hill-top reservoir during off-peak periods, allowing generation in peak periods or to provide rapid response to meet short-falls in generation.

was separated and sold to Mission Energy at the end of 1995, and the RECs sold their shares in NGC in a flotation on the Stock Market, also at the end of 1995. Sixty per cent of National Power and PowerGen were subsequently sold to the public in March 1991, with the balance sold in March 1995. Competition in generation was introduced by requiring all generators (public and private) to sell their electricity in a wholesale market, the Electricity Pool.

The *Electricity Act* also set out a time-table for introducing competition into supply. At privatisation, the 5,000 consumers with more than 1 MW demand were free to contract with any supplier (who could buy directly from the Electricity Pool), but all other consumers had to buy from their local REC, which had a franchise monopoly. In 1994 the franchise limit was lowered to 100 kW, and another 45,000 customers were free to choose their supplier. Starting in late 1998, the remaining 22 million customers had that right, and by mid-1999 the REC franchises finally ended.

The Scottish system, with about 10 GW capacity, was also restructured on March 31, 1990, when the North of Scotland Hydro-Electric Board became Scottish Hydro-Electric, and the non-nuclear assets of the South of Scotland Electricity Board were transferred to Scottish Power. Both were privatised as vertically integrated regulated utilities in June 1991, free to sell into the English market, using the English Pool price as the reference price for Scottish trading and operating under the same system of regulation.

The publicly owned nuclear stations were restructured again when the 5 newer Advanced Gas-cooled Reactors (AGRs) with about 5 GW, together with the new Pressurised Water Reactor at Sizewell, were transferred from Nuclear Electric together with the 2 AGRs from Scottish Electric to British Energy. British Energy was then privatised in 1996. Nuclear Electric's 7 remaining old Magnox reactors with about 3 GW (which had negative net value) were transferred to the publicly owned British Nuclear Fuels Ltd, the fuel (re)processing company.

2.1 Market and institutional design

The most interesting institutional change in restructuring the British ESI was the creation of the Electricity Pool - a compulsory bulk electricity spot market that determined the merit order and wholesale price of electricity in Britain. This operated as a compulsory day-ahead last price auction with non-firm bidding, capacity payments for plant declared available (determined as an exponential function of the reserve margin), and firm access rights to transmission (with generators compensated if transmission constraints prevented their bids being accepted). Each day generators bid their plant into the pool before 10 a.m. and received their dispatch orders and a set of half-hourly prices by 5 p.m. for the following day. Bids had to be valid for the 48 half-hourly periods, although generators could specify various technical parameters (minimum load, ramp rates, etc) in some detail to force a particular pattern of use over the day, and also influence whether the plant would set the price.

The half-hourly System Marginal Price (SMP) was the cost of generation from the most expensive generation set accepted (including start-up costs where appropriate), based on a forecast of demand and ignoring transmission constraints. Generators declared available received capacity payments and, if dispatched, the SMP, which together made up the Pool Purchase Price, PPP. All companies buying electricity from the pool paid the Pool Selling Price, PSP, whose difference from the PPP was the uplift, which covered a variety of other payments

made to generators. The System Operator (National Grid) used the same (rather ancient) software GOAL to dispatch plant as the former CEGB. As the successor companies had copies of GOAL, they could shape the rather complex individual plant bids (start-up, no-load, and three incremental prices plus various technical parameters) to optimise their revenue, rather than bidding the true parameters.

The various institutions required to manage the decentralised system were codified in the Pooling and Settlement Agreement, a multilateral contractual arrangement signed by generators and suppliers which provided the wholesale market mechanism for trading electricity. It defined the rules, and required almost all parties wishing to trade electricity in England and Wales to do so using the Pool's mechanisms. It provided the supporting financial settlement processes to compute bills and ensure payment, but did not act as a market maker.

National Grid Company (NGC) owns and controls high voltage transmission, and as the Transmission System Operator, was responsible for scheduling and despatch. Elsewhere the Systems Operator is often legally required to be independent of generation and transmission. NGC also acted as the Ancillary Services Provider, the Settlement System Administrator and the Pool Funds Administrator, though again the provision of these services can be and often are separated from the provision of transmission services.

In addition to the Pool, which acted both as a commodity spot market producing the reference price and a balancing market, most generators and suppliers signed bilateral financial contracts for varying periods to hedge the risk of pool price volatility. The standard contract was a Contract for Differences (CfD) which specified a strike price (£/MWh) and volume (MWh), and was settled with reference to the pool price, so that generators were not required to produce electricity in order to meet their contractual obligations. These CfDs could be one or two-sided, offering different hedging possibilities. Partly because the market structure was so concentrated, and partly because of the pass-through nature of the franchise contracts, other markets were slow to develop and remained very illiquid. The Electricity Forward Agreements market emerged as a screen-traded over-the-counter market that allowed contracts to be traded anonymously and portfolio positions balanced. It failed to evolve into a futures market, partly because of the illiquidity caused by the large number of products (four-hourly periods for working and non-working days, for SMP, PPP and uplift), but mainly because the underlying market was so uncompetitive.

Contracts are not only important for risk-sharing but were also critical in managing the transition from a vertically integrated company able to pass all its costs through to its captive customers to a market-based industry in which customers were free to buy from the cheapest supplier. The two major transitional problems facing the designers were that British deep-mined coal was considerably more expensive than imported coal (and was soon to be revealed uncompetitive against gas), and that nuclear generation had failed to set aside definable funds for decommissioning. The surplus available to build up a decommissioning fund after paying for operating and fuel cycle costs were likely to be far too low given the likely equilibrium Pool price. The second problem was dealt with by imposing a Non-fossil Fuel Obligation (NFFO) on the RECs (to buy electricity generated from non-fossil fuels, overwhelmingly nuclear power), and imposing a Fossil Fuel Levy (FFL) on all fossil generation (initially at the rate of 10.8% of the final sales price). This levy was paid to Nuclear Electric to build up a fund to meet its

liabilities (of about £9.1 billion, which can be compared with the privatisation proceeds from selling off the CEGB of just under £10 billion).

The first problem of transition was handled by a series of take-or-pay contracts between the generators and the still state-owned British Coal for the first three years at above world market prices. The generators in turn held contracts to supply the RECs for almost all their output, for up to three years, that allowed the costs of the coal contracts to be recovered from these contract sales.² There was the additional and very important benefit that the profit and loss accounts of the generators and RECs could be confidently projected for the first three years, and these provided the necessary financial assurance for the privatisation to proceed.

There are two routes to effective competition in generation. The first and more satisfactory route is to ensure that capacity is divided between sufficiently many competing generators that no one generator has much influence over the price. This option was ruled out by the tight Parliamentary timetable which gave too little time to reconsider plans and to divide the generation companies further once it became clear that nuclear power was unsaleable. At privatisation, the two fossil generators set the pool price over 90 per cent of the time (the balance being set by Pumped Storage, which arbitrated a limited amount of electricity from the off-peak to the peak hours). Nuclear Electric, Scotland and France supplied base-load power that hardly ever set the pool price. Green and Newbery (1992) calculated that a duopoly unconstrained by entry would have significant market power and would be able to raise pool prices to very high levels (shown in figure 4 below).

The second and indirect route to competitive pricing is to induce generators to sell a sufficiently large fraction of their output under contract, and expose them to a credible threat of entry if the contract price (and average pool price) rises above the competitive level. A generator that has sold power on contract only receives the pool price for the uncontracted balance. If this is a small fraction of the total (and it is usually about 10-20 per cent), then there is little to gain from bidding high in the pool. High bids run the risk that the plant is not scheduled, leading to the loss of the difference between the SMP and the avoidable cost, and the trade-off between lost profit on uncontracted marginal plant and higher inframarginal profits is increasingly unattractive as contract cover increases. Contracts and entry threats are complimentary - entry threats encourage generators to sign contracts, and contracts facilitate entry.

The advantage of the creating sufficiently many companies for competition is that it does not need to rely on the continued contestability of entry, and it works well even when the competitive price is well below the entry price, in periods of excess capacity. As this route was not chosen, contracts and entry threats were all that remained, at least if price regulation was to be avoided. On vesting, the three generating companies were provided with CfDs for virtually their entire forecast output, for periods of between one and three years. This both managed the transition to a free market and initially reduced their incentive to exercise spot market power to negligible levels, though not their ability to take advantage of transmission constraints and to game capacity availability.

² The details of the various contracts required are set out in more detail in Henney (1994, pp120-4)

2.2 Regulation of domestic suppliers, entry and the “dash for gas”

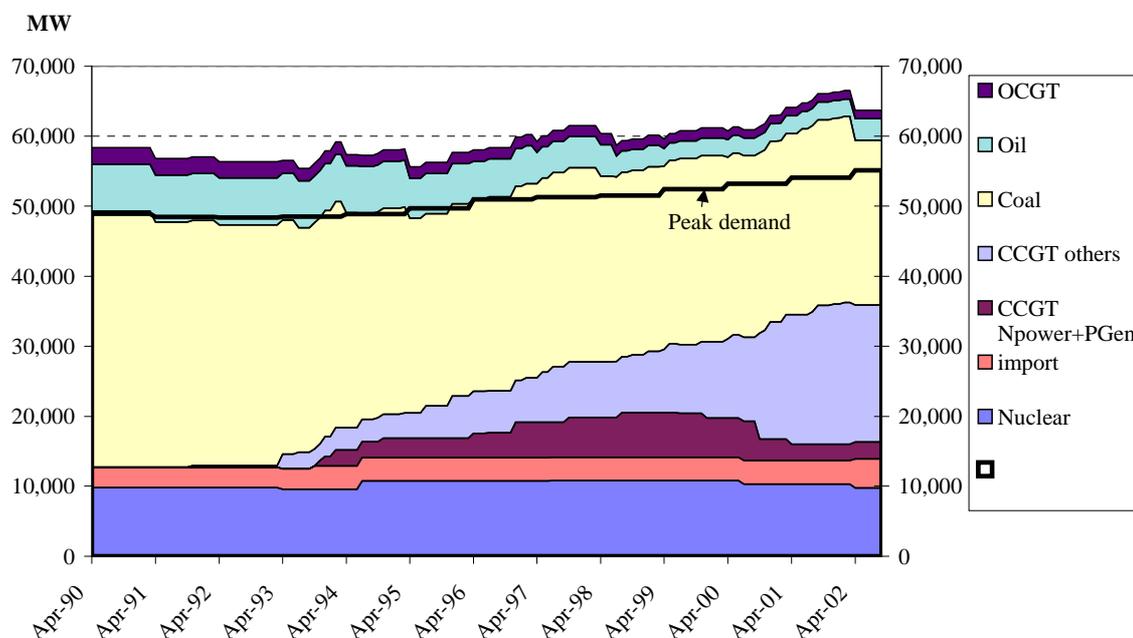
Initially a minority of the market was free to buy power in the Pool or by contract, and the captive customers required regulatory assurance that their prices would be reasonable. This was assured by allowing the RECs to pass through the regulated charges for transmission and distribution, and requiring them to demonstrate that they had purchased any power on behalf of their captive customers “economically”. In advising on market design and regulation, Professor Stephen Littlechild (who was subsequently appointed the first DGES) recognised that the main problem facing the ideal of a competitive ESI was the overwhelming market power of the incumbents. If restructuring could not be relied upon to deliver lower concentration, then entry was the only route to eventual competition. Entry of new merchant Independent Power Producers (IPPs) would be helped by the existence of long-term contracts for gas and electricity. The solution was to allow the RECs to offer long-term Power Purchase Agreements (PPAs) to IPPs, and to hold equity as an incentive to sign these contracts. The PPAs allowed the IPPs to sign long-term contracts for gas (usually take-or-pay) and to issue comparable duration bonds. The economic purchasing requirement was designed to reduce the risk of “sweet-heart deals”, and, critically, the franchise would end in 1998, limiting the possible damage to captive customers.

The solution was accepted, and substantial entry occurred. Within a few months contracts (generally of 15 years duration) had been signed for some 5 GW of gas-fired CCGT plant, which, in addition to the incumbents' planned 5 GW of similar plant, would displace about 25 million tonnes of coal, or nearly half the 1992 generation coal burn of 60 million tonnes. The new CCGT capacity amounted to about one-sixth of existing capacity, which was in any case more than adequate to meet peak demand. The “dash for gas” and the switch from coal more than halved the size of the remaining deep coal mining industry. The coal labour force had fallen from nearly 200,000 at the time of the 1984-5 coal miners' strike to about 70,000 by 1990, but pit closures reduced numbers to 20,000 by 1993 and less than 10,000 by 1998. Figure 1 shows the rapid entry of gas-fired generation, and the resulting evolution of capacity connected to the National Grid. The decline in CCGT owned by PowerGen and National Power reflects industrial restructuring discussed below.

Capacity payments were made to each generating set declared available for despatch, and were equal to the Loss of Load Probability (LOLP) multiplied by the excess of the Value of Lost Load (VOLL, initially set at £2,500/MWh and indexed to the retail price index) over the station's bid price (if not despatched) or the SMP (if despatched). This was set the day ahead and proved manipulable by declaring plant unavailable, and then re-declaring available on the day to collect the now raised payment. This practice was investigated by the regulator and new audit procedures were agreed to reduce the incentives for mis-reporting unavailability (Offer, 1992), together with new Pool rules for computing LOLP. This was now determined by the highest declared or re-declared capacity in the current and seven previous days, so that there was an eight-day lag between declaring a plant unavailable and its impact on LOLP. A somewhat perverse implication was that the actual LOLP could be unity (certain power cuts) while the value used to reward capacity could be almost zero. Newbery (1998c) argued that the computation of LOLP seemed excessive, given the high level of reliability over the first decade, and its overestimate may have contributed in part to the high capacity payments. On the other hand, the VOLL seemed rather low, as Patrick and Wolak (1997) found that large

consumers in one area were charged £7,153/MWh in each of the three peak (or “triad”) half-hours in 1994/5 for grid connection charges.³ As it was the product of VOLL and LOLP that determines capacity payments, these two possible errors may have been offsetting.

Figure 1 Plant Capacity connected in England and Wales, 1990-2003



Source: NGC *Seven Year Statements*, various years, and data from J Bowers

In the 1994-95 financial year, the generators earned £1,421 million from capacity payments, or £24.5/kW/year compared to £5,821 million from selling at the SMP. Capacity payments were thus 20% of total payments for generation (excluding other ancillary services supplied by generators to the pool), and far higher than in earlier years when plant was more fully contracted. These capacity payments would have been sufficient to build 3GW of new plant, or nearly 6% of total capacity. In the period 1995-97, the annual average capacity payment was over £30/kW/yr. During this period, the annual grid connection charge varied from 8/kW to -£10/kW. The cost of keeping a new open cycle gas turbine to provide reserve power might be £20/kW in interest and depreciation, and perhaps £6/kW for O&M (MMC, 1996), so capacity payments should have been more than enough for security of supply.

High capacity payments could also provide incentives for large generators to withhold plant, as Newbery (1995) demonstrated. Depending on the contract cover and plant margin, generators with a market share of about 30% might have an incentive to withdraw plant, exactly the opposite incentive to that intended. Green (2004) examined the evidence and found that this strategy did not appear to have been significant. Later, dissatisfaction with

³ Admittedly, these charges are not known accurately until after the peak, but large customers subscribe to moderately accurate forecasting services that can predict when prices are likely to be very high. The very low observed price response suggests that consumers value not adjusting load in response to high prices, and by implication attach an even higher value to not losing load.

capacity payments would be one of the factors causing the DGEN to review the workings of the Pool and recommend the changes that resulted in the New Electricity Trading Arrangements (NETA) of 2001, after which capacity payments were abolished.

In addition to dispatching stations, NGC as SO also had to resolve transmission constraints by paying out-of-merit generators to run if required (“constrained on”) or not to run (“constrained off”) in an export-constrained zone even if in the unconstrained dispatch. Under the vertically integrated CEGB the dispatch schedule automatically determined the security-constrained efficient dispatch, and the grid appeared adequately sized for such that organisational form, but in the market-driven unbundled industry, the costs of resolving constraints rapidly increased (to £255 million in 1993-4 or £4.3/kW/yr). NGC offered an incentive deal to the RECs to share the benefits of reducing these and other costs, an idea that taken up by Offer. The resulting price control for NGC contained incentives to reduce constraint (and other ancillary service) costs, essentially by sharing the costs with a cap and collar. NGC proved adept at contracting for some plant behind constraints, making minor reinforcements to the grid, and scheduling maintenance to minimise these costs, reducing these constraint costs to less than 10% of their peak value.

Another criticism of the Pool was that it was only half a market, lacking any demand side bidding. That was not quite correct, as NGC operated an annual tender auction for the provision of standing reserve to assist in its system management function. Standing reserve was provided by open-cycle gas turbine and pumped storage plant, but also by demand reductions and non-centrally despatched small generators, though all had to offer amounts in excess of 3MW. Large consumers could therefore specify their availability and willingness to reduce demand in various seasons and at various times of day, and NGC then accepted bids for which the total cost of providing load reductions were less than VOLL. In 1997/98 1,809 MW of centrally despatched generation and 458 MW of demand modification and small-scale generation were contracted (NGC, 1997). The offer curve of such bids suggests that while there was some moderately cheap demand side flexibility, beyond a quite modest level consumers needed a higher value than VOLL to be willing to curtail load, again suggesting that VOLL may have been under-estimated, and that short-run demand elasticities for electricity were very low (with current control and metering devices).⁴ In addition, the Pool developed a less successful form of demand-side bidding directly into the Pool, and again, its failure was an additional source of pressure to reform the Pool.

The Pool Purchasing Price determined the price of raw (unconstrained) energy and capacity, but generators and consumers are interested in the price at their location. It was appreciated that the theoretical solution to efficient spatial pricing is locational marginal pricing (LMP) developed by Bohn, Caramanis & Schweppe (1984). As NGC developed a more satisfactory solution to the rather hastily designed system put in place at privatisation, it was recognised that LMP faced a number of potentially serious drawbacks, not least of which was that its performance in the presence of considerable market power was untested (and largely

⁴ Enthusiasts continue to believe that low cost ICT will enable even domestic consumers to time-shift loads such as freezers, hot water and storage heaters, and air conditioners where their thermal inertia allows electricity to be stored for modest periods in the form of heat (or cold). Evidence that this is cheaper than carrying generation reserves remains sparse.

unknown). The additional basis risk of trading at a large number of grid points whose price could diverge considerably from the Pool price would require a large number of potentially illiquid contracts to cover risk. The possible gain in allocating the costs of transmission constraints and losses more precisely was not thought worth the loss in transparency and market liquidity. NGC therefore retained zonal access charges based on the incremental costs of reinforcing the grid to meet demands and supplies in that zone. NGC also publishes annual *Seven Year Statements* (looking ahead seven years) which update predictions of demand and supply by zone and indicate where new generation might best locate. The transmission charges are paid by consumers based on demand at the three half-hours of system maximum demand separated by 10 days (the “triad”), and by generators based on declared net capacity (or output in the triad if facing a negative grid charge).

The more serious weakness in locational pricing was that, in contrast to the CEGB period, transmission losses were not borne by generators, distorting the merit order, while firm access rights rewarded, rather than penalising, generators in export constrained zones. Scotland was the obvious example, and two successive attempts by Offer to introduce transmission losses were successfully appealed to the courts.

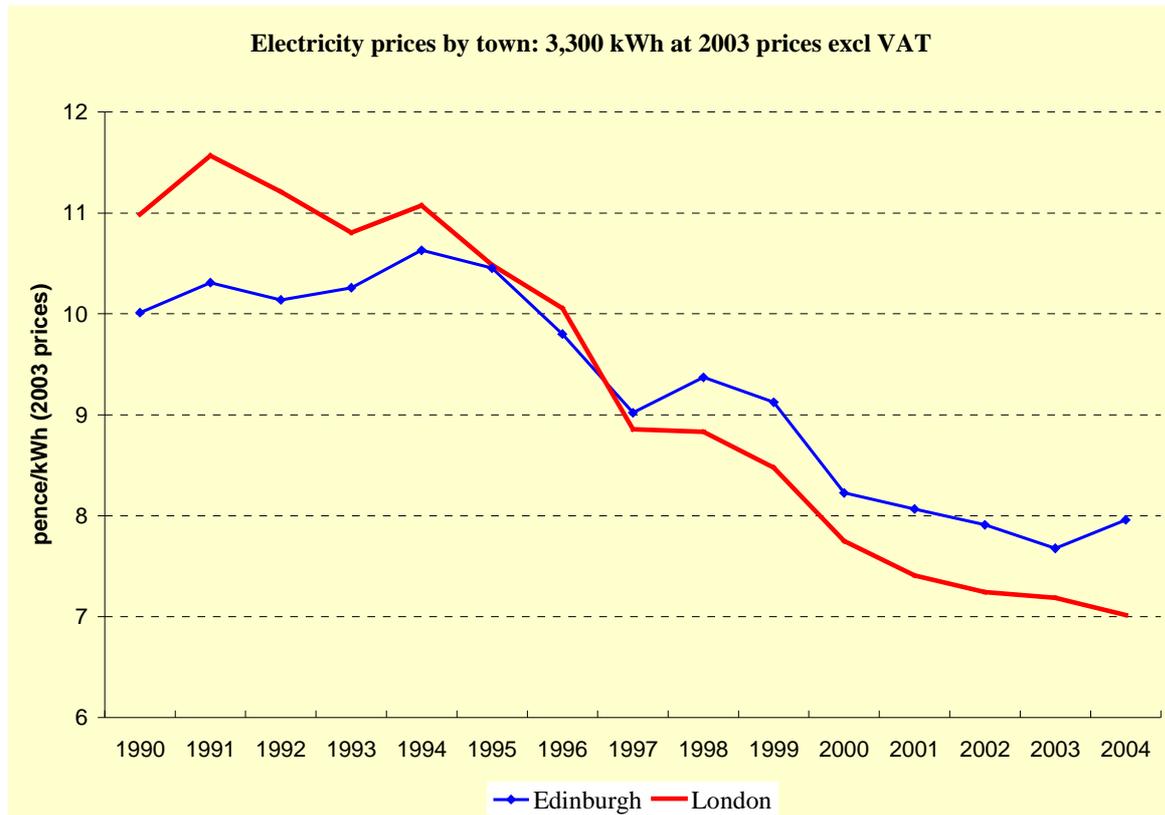
2.3 Performance after privatisation

Privatisation and restructuring the CEGB delivered substantial improvements in efficiency, as Newbery and Pollitt (1997) document. They estimated that after the first five years, costs were permanently 6% lower than under the counterfactual continued public ownership, with a present discounted value at the public sector discount rate of 6% equal to a 100% return on the sales value of £10 billion. Labour productivity doubled, real fuel costs per unit generated fell dramatically (even in the publicly owned nuclear company), and substantial new investment occurred at considerably lower unit cost than before privatisation. The contrast with Scotland was striking, where a similar social cost-benefit study by Pollitt (1999) found negligible efficiency improvements. One reason was undoubtedly that the two Scottish companies were not restructured, and remained vertically integrated, making it more difficult for competitors to gain access to their home market, even though nominally Scotland was able to trade in the English electricity Pool. Scotland was an exporter through a severely constrained interconnector that was not efficiently priced, and had only two local generators, reducing the prospects of competition. Figure 2 shows the average price of domestic electricity in Edinburgh, Scotland, and London, England. Initially, London was 10% more expensive than Edinburgh, but by 2001 Edinburgh was almost 10% more expensive than London.

This raises the question, currently exercising the regulator (now named Ofgem, the Office of Gas and Electricity Markets) and the proposed GB System Operator, NGC, on how access to these scarce interconnectors should be determined and priced (NGC, 2004). If full nodal pricing is thought problematic, then “market splitting”, in which the SO determines when constraints isolate markets, and then sets market clearing prices in each zone, as in Norway, would seem attractive. In particular, it would allow English generators to contract with Scottish consumers, and this counterflow would release more export capacity from Scotland, as the constraint only applies to net electricity flows. English generators would effectively be paid to export to Scotland an amount equal to the excess of the English

marginal price over the Scottish marginal price (which should include transmission losses), and would thus be able to compete effectively in that market. Under the existing system Scottish generators could price locally up to the English Pool price, as the shadow price of the export constraint was not made explicit and they did not pay for the quite substantial transmission losses.⁵

Figure 2 Domestic electricity prices at 2003 prices excluding VAT



Source: DTI *Energy Prices*, various issues. Figures are averages for credit customers taking 3,300 kWh/yr

The lesson that vertical unbundling (at least legal, and preferably ownership) is essential for effective competition has been accepted in the new Electricity Directive, and in the consultation for the British Electricity Trading and Transmission Arrangements (BETTA) that started in 2003.⁶

Privatisation, combined with unbundling and a transparent wholesale market, provided incentives for considerable efficiency improvements, but the concentrated market

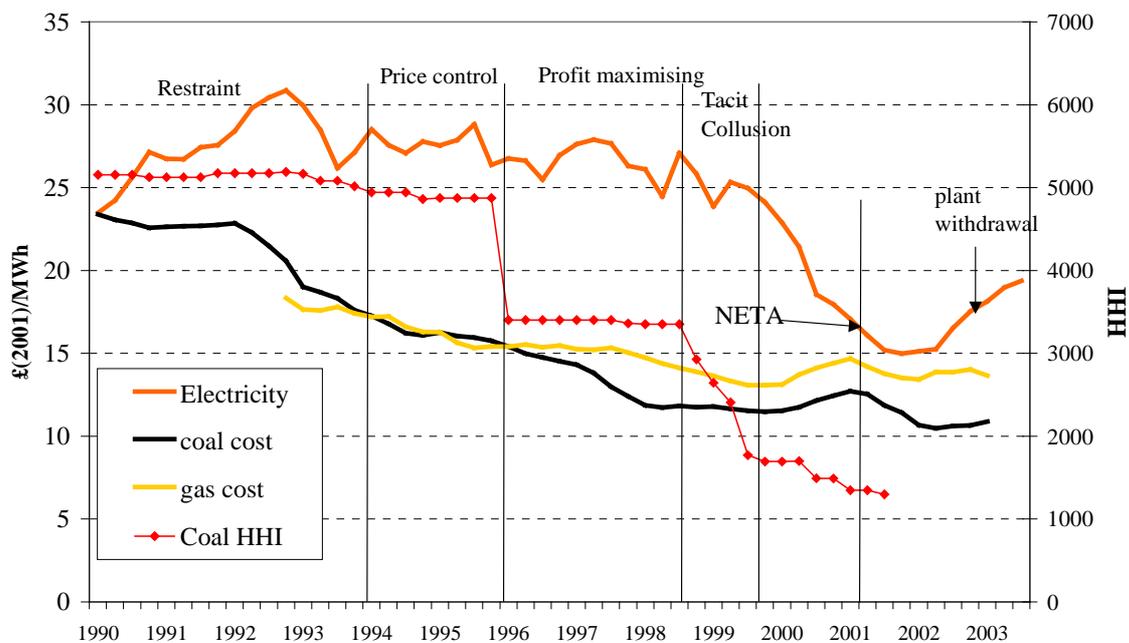
⁵ Marginal transmission losses from Northern generators to the load centres were often greater than 10%. Despite various attempts and judicial review Ofgem “is of the opinion that it is not legally possible for it to approve this Modification Proposal” (to introduce cost-reflective charging for transmission losses). Ofgem *Information Note* of 30 January 2004.

⁶ The Energy Bill introduced in November 2003 aims to create a single GB-wide set of arrangements for trading energy and to access to and use of a single GB transmission system, but is subject to Royal Assent in July 2004.

structure initially allowed the incumbent generators to retain these cost reductions as enhanced profits. The social cost-benefit analysis of Newbery and Pollitt (1997) found that while the overall simple sum of net benefits of privatising the CEGB was nearly £10 billion, consumers lost relative to the counterfactual in which fuel prices fell and the CEGB had set prices as in the past, while the owners of the generation companies gained very substantially.

Figure 3 summarises a long and turbulent period of pricing in the England and Wales wholesale market, during the entire life of the Pool until 2001, and under the New Electricity Trading Arrangements (NETA) thereafter. Hourly and daily price volatility was very considerably higher than the smoothed figures shown. Averaged over the year 1997/8, for example, the average spread between the highest and lowest half-hourly PPP prices on a day is 180% of the average price on that day, and the standard deviation of half-hourly prices over the year is 78% of the average PPP.⁷

Figure 3 Real wholesale electricity and fuel prices 1990-2003



Source: Pool data and RPDX data, DTI *Energy Prices* for fuels, HHI from J Bower

Figure 3 shows the fuel cost of generating electricity from coal at 36% thermal efficiency and from gas at 50% gross efficiency (55% net efficiency), and hence the margin between the yearly moving average wholesale price and avoidable cost. The line with diamond markers gives on the right hand scale the Herfindahl Hirschman Index (HHI) of market concentration of coal-fired plant (for most of this period the price-setting plant). This

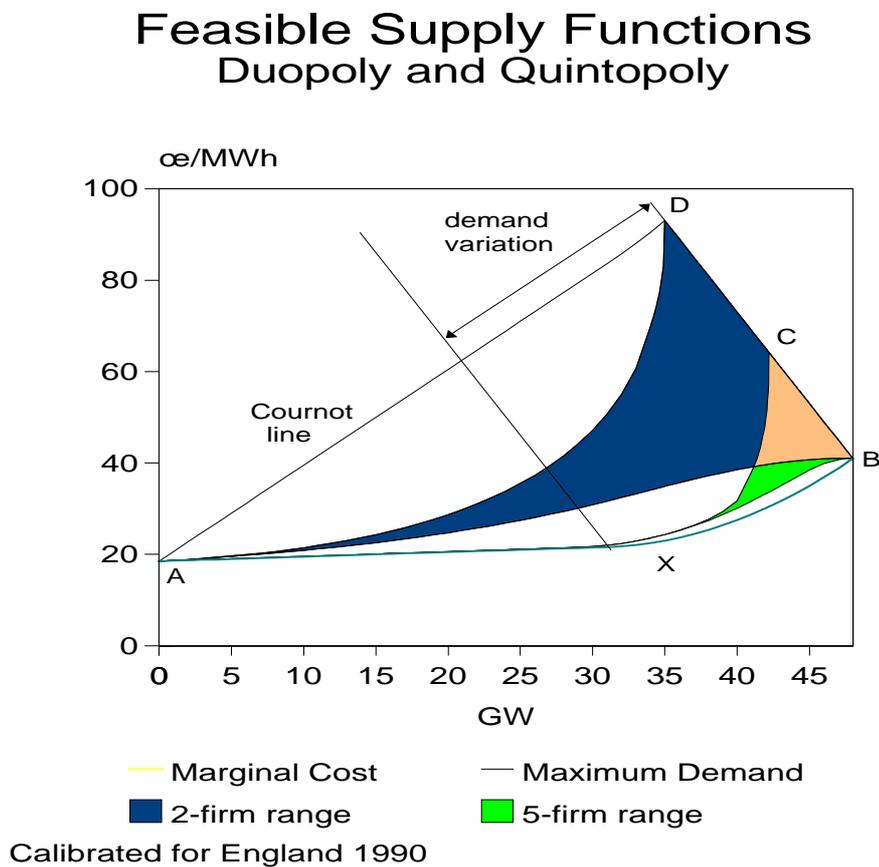
⁷ Some of this variability is predictable and can therefore be hedged. Characterising the unpredictability of prices, which is a measure of risk, requires correcting for predictable time variations over the day, week and year. The standard deviation of the difference between the actual half-hourly price and the moving average for that hour was £14/MWh or 55% of the average price.

is the sum of the squared percentage shares of available capacity, so that the initial value of just over 5000 represents the equivalent of a duopoly.⁸ The evolution of prices is divided into periods identified by Sweeting (2001). To understand them it is first necessary to discuss the determinants of imperfectly competitive equilibrium prices in an electricity pool.

3 CHARACTERISING MARKET EQUILIBRIUM IN A POOL

Modelling price formation to understand market power and market efficiency is a challenging problem that is not yet fully solved. Green and Newbery (1992) modelled the English Electricity Pool by adapting Klemperer and Meyer's (1989) supply function equilibrium (SFE) model. The model is difficult to solve and typically gives a continuum of equilibrium prices. Figure 4 reproduces their calibrated model for England and Wales, ignoring contracts and entry threats.

Figure 4 Equilibrium price range ignoring entry threats and contracts



This approach is attractive and appears to be supported by companies' claims that they bid supply schedules. It assumes a single-price gross pool with bids that hold for a reasonable period of time over which demand varies - as with daily bidding in the English Pool. In its simplest form it assumes that the supply functions bid are continuous and differentiable, and that demand is linear with constant slope but varies over the 48 half-hours. Each generator chooses a

⁸ The number of equivalent firms is $10,000/HHI$.

supply function that maximises his profits given the residual demand he faces, made up of the variable total demand *less* the total supplies bid by other generators. As his bid has to be valid over the whole daily range of residual demands, instead of choosing a single quantity to submit to the market that would determine a single price (as under the Cournot assumption), he has to choose a continuous function relating the quantity that he is willing to offer at each price realisation. The set of feasible solutions will be Nash equilibria in supply functions.

There are, however, difficulties with this approach. Most pools (and the English Pool in particular) restrict bids to a single price for each quantity offered, producing a step function or ladder rather than a continuously differentiable function. The Amsterdam Power Exchange is a good example, where their web site provides the bid and offer ladders for each hour, their intersection providing the market clearing price for that hour. Fabra et. al (2004) argue that this radically alters the nature of the equilibrium, and requires modelling the market as a last-price auction, following on the earlier paper of von der Fehr and Harbord (1993). They solve this if there is a single period and a known inelastic demand (up to a binding price cap), but cannot characterise the solution for bids that must hold for many periods (48 in Britain) with uncertain or varying demand. Hortacsu and Puller (2004) use data that is in step function form, which they then smooth to determine the marginal revenue of the residual demand facing each generator, and demonstrate that at least for the larger companies their bids appear to be profit maximising against this smoothed schedule. Newbery (1992) suggested that if generators randomised over the positions of the steps in a step function, they could replicate a differentiable supply function, but it remains an open question whether this would be an optimal response to such behaviour on the part of other generators.

Standard Cournot oligopoly models are simpler, can be defended in tight market conditions, but suggest a more deterministic outcome than supply function models with their range of indeterminacy. Increasingly, consulting companies are developing price-formation models, the best of which capture the strategic aspects of supply function models with more careful modelling of the non-convexities of start-up costs which can dramatically influence the cost of providing additional power for short periods.

Despite this apparent diversity of approach and the rather unsatisfactory theoretical foundations of bidding models, the evidence from various markets is consistent with the SFE story. Competition is more intense (closer to Bertrand) and prices closer to avoidable costs with spare available capacity, but as the margin of available capacity decreases, competition becomes less intense and outcomes closer to Cournot (as in the SFE). However, there remain two additional considerations before we can understand the English price evolution shown in figure 3. First, while the possession of market power is legal, abusing it is not, and dominant generators need to be aware of the threat of competition references. That is the simplest explanation of incumbent bidding behaviour from 1990-94, where an acceptable level of prices at which to aim was arguably the entry price. The second important feature of the Pool is that it is a repeated auction, repeated every day and with the evidence of bids and outcomes available with a relatively short lag to the participants.

The European Commissions provides a characterisation of collective dominance as a situation in which the market characteristics are conducive to tacit co-ordination and such co-ordination is sustainable, that is it is profitable and deviations can be deterred. The market characteristics that are conducive to tacit co-ordination include concentration, transparency,

maturity, with a homogenous product produced by companies with similar costs and market shares, facing an inelastic demand, and with barriers to entry. Evidence supporting such a finding would include excess price-cost margins, profits and an insensitivity of prices to cost falls. With the important exception of barriers to entry, the English Pool appeared to have all these defining characteristics and behaviour. Tacit co-ordination was therefore to be expected, and market surveillance should clearly take account of this possibility.

3.1 Tacit co-ordination in the Electricity Pool

Andrew Sweeting (2001) tested for tacit co-ordination by looking at individual company bids, subtracting all other bids from total demand to determine residual demand, and asking whether the bids were profit maximising given the residual demand (but ignoring contract positions). He finds that in the first period up until 1994 both incumbents bid less aggressively than would be (short-run) profit maximising, and that the prices were on average around the level at which entry was just profitable. If we were to conjecture what strategies collectively dominant incumbents might co-ordinate on, given close regulatory scrutiny, then keeping the price at the entry level while dividing the market in proportion to some objective criterion (such as plant capacity) would be plausible. It would also explain why both companies were keen to build new CCGTs even when their economics were marginal,⁹ for this would allow them to justify an increased market share (or in practice, in this prisoners' dilemma, maintain market share in response to investment by the other company).

As figure 3 shows, fuel costs continued to fall away from electricity prices, to the point that the regulator claimed that they represented excess profits. He imposed price caps (on both the annual average demand- and time-weighted Pool price) until the generators divested enough plant to improve competition. The companies sold 6,000 MW to Eastern (later TXU) with an earn-out clause of £6/MWh,¹⁰ ostensibly to compensate for the sulphur permits transferred with the plant and to reduce the buyer's risk, but with the additional consequence of raising their rival's marginal cost when bidding into the Pool.

Sweeting found that during the period 1996 (after divestiture when the price cap ended) to 1998, bids seemed to be best responses and thus each firm was non-collusively maximising profits. The price-cost margin increased as the regulatory threat of market abuse was replaced by (rather relaxed) competitive pressure, and the incumbents were probably quite happy to have sold plant at prices reflecting market power, in a market that was continuing to experience rapid entry. The ability to sustain a high price-cost margin depends on the volume of excess capacity, which was threatening to increase rapidly unless more coal plant were withdrawn or scrapped. At the same time Offer and Parliament (through the select committee that investigated the energy industries) were becoming increasingly convinced that the Pool was not working well, and that the detailed rules of the dispatch algorithm were being manipulated to increase profits

⁹ While it is true that IPPs also entered, they did so on rather favourable long-term contracts not available to the incumbent generators. Certainly given the early gas prices and CCGT efficiencies, and compared to the opportunity cost of coal, the economics of investment were very marginal, as the House of Commons (1993) argued.

¹⁰ That is, Eastern paid £6/MWh to the selling company for all electricity generated, increasing the effective marginal cost by that amount.

(Offer, 1998a-c). Henney (2001) notes other sources of discontent, notably the Labour Party's belief that the Pool used "an operating and pricing system that was not competitive and was weighted against coal" (Robinson, 2001).

4 THE NEW ELECTRICITY TRADING ARRANGEMENTS (NETA)

In October 1997, the Minister for Science, Energy and Technology asked the DGES to review the electricity trading arrangements might and to report results by July 1998. Offer's objectives, approved by the Government, were to consider whether, and if so what, changes in the electricity arrangements would best meet the needs of customers with respect to price, choice, quality and security of supply; enable demand to be met efficiently and economically; enable costs and risks to be reduced and shared efficiently, provide transparency; respond flexibly to changing circumstances; promote competition in electricity markets, facilitating entry and exit from such markets; avoid discrimination against particular energy sources; and be compatible with Government policies (Offer, 1998d, pp 83-4).

The process that led to the eventual ending of the Pool and its replacement by NETA have been extensively described and criticised elsewhere (e.g. Newbery, 1998b,c). Shuttleworth (1999), writing after the publication of Offer's *Interim Conclusion* (Offer, 1998d), noted that "it is difficult to find any rigorous analysis to underpin the reform proposals", while Newbery (1998c) concluded that "(T)he present review appears to have relied mainly upon unsubstantiated claims, inappropriate analogies, unquantified criticisms, and a remarkably uncritical assessment by the participants of the debate, without commissioning the kind of detailed analysis one might have expected from a regulatory agency claiming industry expertise."

The *Pool Review* (Offer, 1998e) argued that the complexities of price formation in the Pool allowed generators to exercise more market power than would have been possible had the market been structured more like a classic commodity market. It criticised the opaque method of determining price based on a scheduling algorithm(GOAL) devised for the vertically integrated (between transmission and generation) CEGB, as well as the capacity payments, and the concept of a single-price auction. It also criticised the Pooling and Settlement Agreement (PSA) for blocking desirable changes, because as a contract between parties it could only be changed with their agreement, and, given the voting arrangements, it was rare for any change to make all parties better off. The recommendations *Pool Review* were accepted and NETA went live on 27 March 2001. NETA replaced the PSA by a Balancing and Settlement Code with a well-defined method of making modifications, giving Ofgem more influence in the process. The Pool ceased to exist. Electricity was now to be traded in four voluntary, overlapping and interdependent markets operating over different time scales. Bilateral contract markets cover the medium and long run, while forward markets offer standard contracts (base-load, peak hours) for periods up to several years ahead. A short-term "prompt" bilateral market (OTC and exchange), operating from at least 24 hours to Gate Closure (3¹/₂ hours before a trading period, subsequently reduced to one hour in July 2002), allowed parties to adjust their portfolio of contracts to match their predicted physical positions. This short-term market would yield information to construct a spot price for each half-hour (e.g. the UKPX Reference Price Data).

At Gate Closure, the official end of the bilateral markets, all parties had to announce their Final Physical Notifications (FPN) to the System Operator (SO). The SO would then accept bids and offers for balancing the system. These bids and offers would be fed into the Balancing Mechanism to produce cash-out prices for clearing imbalances between traders' FPNs and their actual (metered) positions.

The most obvious difference between NETA and the Pool is that under the Pool all generation was centrally dispatched while under NETA plant is self-dispatched. The obligation to balance output with demand is now placed on each generator, with the SO's task confined to ensuring system stability. The Pool, that acted as both a wholesale market for all electricity and allowed NGC as SO to balance the system, was replaced by a Balancing Mechanism (also operated by NGC as SO) for the residual imbalances of parties that fail to self-balance. Whereas the Pool operated as a uniform single-price auction for buying and selling all power (including that needed for system balance), the Balancing Mechanism is run as a discriminatory (pay-as-bid) auction. NGC charges for balancing through the Balancing Services Use of System charge.¹¹

Elxon determines two cash-out prices: the weighted average of accepted offers determines the System Buy Price (SBP) and that of bids the System Sell Price (SSP). Any party found to be out-of-balance when metered amounts are compared with FPNs is charged either the SBP (if they are short, that is the FPN is more than the metered output (for a generator) or less than metered consumption (for a consumer), or they receive the SSP if they are long (and have to spill power). The critical feature of the original design of the Balancing Mechanism is that these prices are normally different ($SBP \geq SSP$),¹² and penalise each party's imbalances, whether or not they amplify or reduce the system imbalance as a whole. Figure 6 below gives an indication of this volatility.

As a result of the initially extreme volatility of the balancing prices a considerable number of modifications were made. One of the more important (P78, shown on figure 6) made the reverse balancing price (i.e. the price facing parties who were in the opposite position to the overall market, e.g. long when the market was short, and hence aiding balance) would revert to the spot price, and hence not penalise those helping balance the system relative to their selling in the spot market. The idea of moving to a single marginal balancing price has been mooted but so far rejected by Ofgem.

Note that there are two distinguishing characteristics of the Balancing Mechanism, either of which could be changed independently. The first is that there are (normally) two different prices for being short or long. The second characteristic is that these prices are determined from a discriminatory auction in which bids and offers pay or are paid as bid, and the average cost of securing the services is then charged out.¹³ One consequence of this

¹¹ Ofgem is wedded to the fiction that it is possible to distinguish between the cost of trades that are required to balance the *system* and the penal charges levied through the cash-out prices for *individual* imbalance, and has elaborate rules for drawing this distinction.

¹² The prices were equal by about 25% of the time, and SSP exceeded SBP very occasionally (0.1% of the time) in the first 18 months.

¹³ The Dutch balancing market is at the other extreme. It operates a uniform price auction to determine a single price for those 15-minute periods in which the system is either long or short for the whole

combination is that it is more risky for a generator to offer balancing services. If a generator has an accepted offer to increase output, and then suffers a loss of output, he is likely to have to pay more than he is paid. He may therefore prefer to retain the spinning reserve for his own insurance. A single final balancing price would make such an offer never any worse than self-insuring and normally better, and would thus promote a more liquid balancing market.

4.1 The evolution of market structure

During the period in which NETA was under discussion, National Power and PowerGen had to decide on their future strategy in the face of considerable uncertainty about market developments and impending excess capacity. Until 1995, the Regional Electricity Companies (RECs) had been protected against take-over by Golden Shares, but these lapsed and in the following few months eight of the 12 RECs were targeted. Six were successfully acquired, two by other UK regulated utilities, one by the vertically integrated Scottish electric utility, Scottish Power, and two by US utilities. The two RECs targeted by National Power and PowerGen were referred to the Monopolies and Mergers Commission and then blocked by the DTI. One of these was subsequently bought by another US utility group.

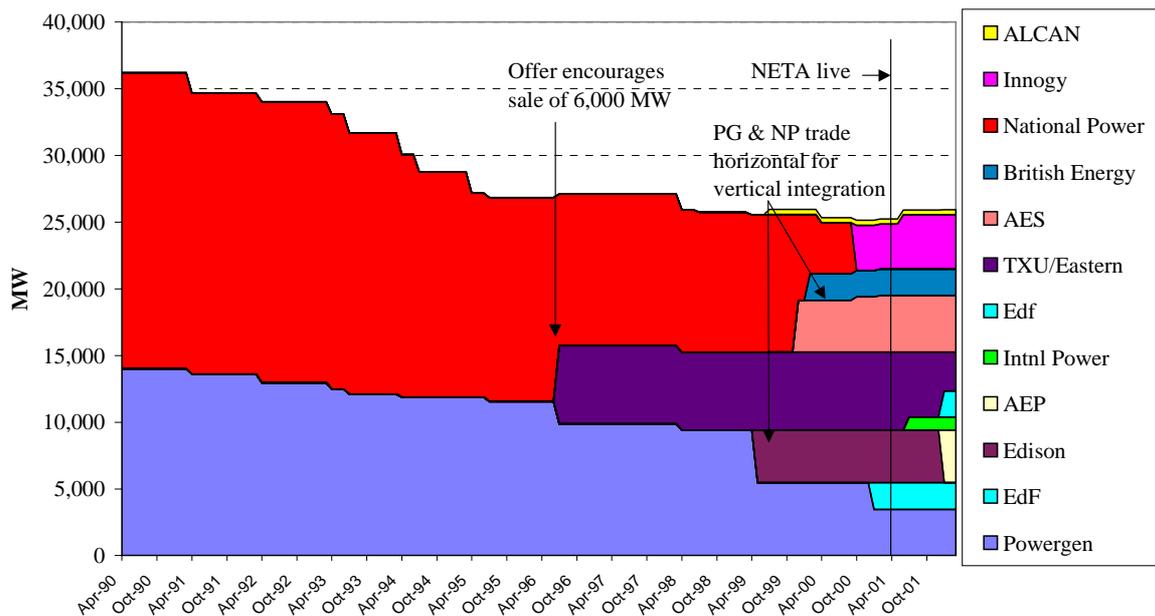
The logic of combining risky generation with the offsetting risks of downstream customers was amplified by the very favourable low debt position of these regulated utilities, and made them irresistible to the duopoly generators, but their market power made vertical reintegration unlikely to pass the competition authorities.¹⁴ The obvious solution was for the companies to divest generation so that they could pass scrutiny when they bid for supply companies, that in preparation for full retail liberalisation were being unbundled from the REC distribution businesses. The urgency of achieving this objective was increased by the uncertainty over the new trading arrangements. Thus on 25 November 1998 PowerGen entered undertakings with the Secretary of State to sell 4,000 MW of plant and to end the earn-out clause on its 1996 power station sales in return for clearance to acquire East Midlands Electricity's distribution and supply business. Similarly National Power agreed to sell the 4,000 MW Drax station in order to buy the supply business of Midlands Electric. The delicate task facing National Power and PowerGen was to sell the plant for attractive prices into a market that was in danger of being oversupplied with increasing gas generation. Here the new Labour Government helped by imposing a moratorium on building new gas-fired plant in 1997 to assist the coal mining industry during the period of sorting out the Pool (and also imposed a so-called Climate Change Levy that was actually a tax on energy rather than carbon, again protecting coal). The DTI estimated that this delayed the building of 5,800 MW of gas-fired capacity.

period, and charges those who are short while rewarding those long. There is the potential (not yet used) to add a penalty of 1Euro/MWh to both imbalances. If the system is both short and long within the 15 minute period it determines two prices, effectively one for each sub-period in which the imbalance is in one direction.

¹⁴ Risks could have been hedged by long-term contracts between generation and supply companies, but the transaction costs of writing long-term contracts to cover all contingencies (such as the ending of the Pool, the Emissions Trading System, Climate Change Levy, Renewables Obligation Certificates) might make vertical integration more attractive. Supply companies also suffer from credit risk as they are typically under-capitalised unless combined with generation or distribution.

The solution was to ensure that the price-cost margin remained high while plant was offered for sale, and Sweeting (2001) identifies the period from 1998-early 2000 as one in which National Power and PowerGen could have increased their individual profits if they had bid lower prices. That is consistent with co-ordinating on a higher-price equilibrium than short-run myopic profit maximisation would deliver. During this period plant was profitably sold, indicated by the falling HHI in figure 3, and the changing shares in figure 5.

Figure 5 Capacity ownership of coal generation, 1990-2002



Source: Data supplied by J Bower

The companies buying the plant were warned by Ofgem that there was no guarantee that prices would remain high, particularly given the impending arrival of NETA, which Ofgem was claiming would itself lead to prices at least 10% lower than otherwise. Nevertheless, Edison Mission paid £1.3 billion for the 2000 MW stations at Fiddler’s Ferry and Ferrybridge in July 1999, or £314/kW, and increased the plant output by more than 30%. With the new buyers keen to improve the returns on their purchases by increasing plant output, figure 3 shows that the earlier co-ordinated duopoly equilibrium was no longer sustainable and the price-cost margin collapsed *before* NETA went live, but after the fall in concentration (HHI). Edison Mission subsequently sold its two stations in October 2001 for less than half the purchase price (incurring a balance sheet impairment of \$1.15 billion on the \$2 billion purchase cost).

4.2 The impact of the new trading arrangements on market performance

The intellectual case for replacing the Pool through which all energy was traded by a voluntary Balancing Mechanism covering rather less than 2% of energy was that this would force both buyers and sellers to haggle over the price of electricity without the clear and

transparent signals delivered by the Pool. If parties were forced by a risky, opaque and potentially penal imbalance market to contract ahead, consumers would have time to shop around for better deals, making the market more competitive. This argument ignores two important facts. The first is that about 90% of electricity traded before NETA was under contract, and the annual contract round was a period of intense haggling. NETA did not change that.

The second is that the relative bargaining strength of generators and consumers depends more on the ability of generators to hold the market to ransom than the fine details of the spot or balancing market. This strength is best measured by the extent to which consumers can meet their total demand from all other generators, and the number of generators that are pivotal (i.e. essential) for meeting demand. The loss to a generator of not selling is the difference between the price and variable cost (shown in figure 3), whereas that to a consumer of not being able to buy power is potentially the difference between the value of lost load and the price, which may be hundreds of times as large. Pivotal generators therefore have very considerable bargaining power in any market design.

Three factors influence this bargaining power – the number of competing generators, the reserve margin, and (in the longer run) the ease of entry. Entry was extremely easy in the Pool, but, after vertical integration and with the removal of a guaranteed market of final resort, considerably riskier under NETA. The Pool reserve margin was normally quite adequate as a result of earlier entry, while competition had just become intense as the Pool ended, and was already demonstrating its impact on spot and contract prices. The claim that NETA therefore was necessary (and sufficient) to mitigate generator market power is unsubstantiated.

There was a rather more confused claim that replacing a single price auction (like the Pool) by a pay-as-bid or discriminatory auction would obviously lower the average price, ignoring the auction literature on revenue equivalence. A more sophisticated claim was advanced by Currie (2000), who argued that repeated single-price auctions encouraged collusion more than discriminatory auctions. Newbery and McDaniel (2003) argued that the theoretical, empirical and experimental evidence on auction design applied to the electricity market was ambiguous. Fabra et al (2003) developed simple models comparing the two auction designs and were able to demonstrate that with predictable and unchanging demand, a discriminatory auction would yield lower (short-run) prices than a single price auction, but would typically lead to a less efficient use of plant. They were not able to produce results for multi-period and repeated auctions.

Offer (1999) estimated the costs of switching to NETA at about £700 million (spread over a five year period) followed by additional annual costs of £30 million.¹⁵ Offer justified this cost by claiming that prices would fall 10% as a direct result (although this would be a transfer from generators to consumers, not a net social benefit). The Government helped by removing the gas moratorium, but in the event prices fell so far that CCGT entry was put on

¹⁵ “The costs of implementing and operating the new trading arrangements are estimated to be between about £136m to £146m per annum, for a five year period. Thereafter the operating costs are expected to be of the order of £30m per annum.” (Offer, 1999, p14). These continuing costs almost certainly understate the extra costs of maintaining 24-7 trading floors for balancing.

hold indefinitely. Ofgem was able to claim that the “Evidence of the first year of NETA shows wholesale prices around 40% below those under the former Electricity Pool.” (Ofgem, 2002). While this may exaggerate the fall, figure 3 shows that compared to 1999, prices at the time of NETA were indeed substantially lower. Newbery and McDaniel (2003) argued that the price fall was due to competition, not NETA, as prices fell before NETA, and as electricity cannot be stored, future prospects of changed trading should have had no impact on pre-NETA prices. John Bower (2002) (who supplied the plant data used in Figure 4) demonstrated this more rigorously using formal econometric tests. Evans and Green (2003) confirmed this finding for Bower’s specification, but raised the question whether the announced end of the Pool would unravel the collusive equilibrium by backward induction for the point at which the Pool (and transparent pricing) would end. They found support for a variable “para-NETA” that takes the value 0 until October 2000, and 1 thereafter. An alternative and simpler explanation is that indeed collusion ended, but because of the actual fall in concentration, an event set in motion, not by the anticipated NETA, but by the desire of the incumbents to integrate forward into supply, expressed as early as 1995.

In the medium run the average price in any electricity market will be determined by the conditions and costs of entry and exit (and possibly on the threat of price caps or other regulatory interventions that might reduce expected profits). The first publication of the Joint Energy Security of Supply (JESS) Working Group in June 2002 stated that “Capacity margins are healthy and are expected to remain so.” Their next report in February 2003 indicated no change. Shortly thereafter, generating companies started to experience financial distress and some went into administration. British Energy, the privatised nuclear company, had to be bailed out by the Government, and surviving companies started to scrap or mothball plant. The plant reserve margin fell to below the NGC’s target margin of 20%, and in the summer of 2003, NGC’s forecasts suggested a rapidly deteriorating situation. In response, Winter 2003/4 forward peak prices rose from £25/MWh to over £35/MWh and some mothballed plant was returned to the system.

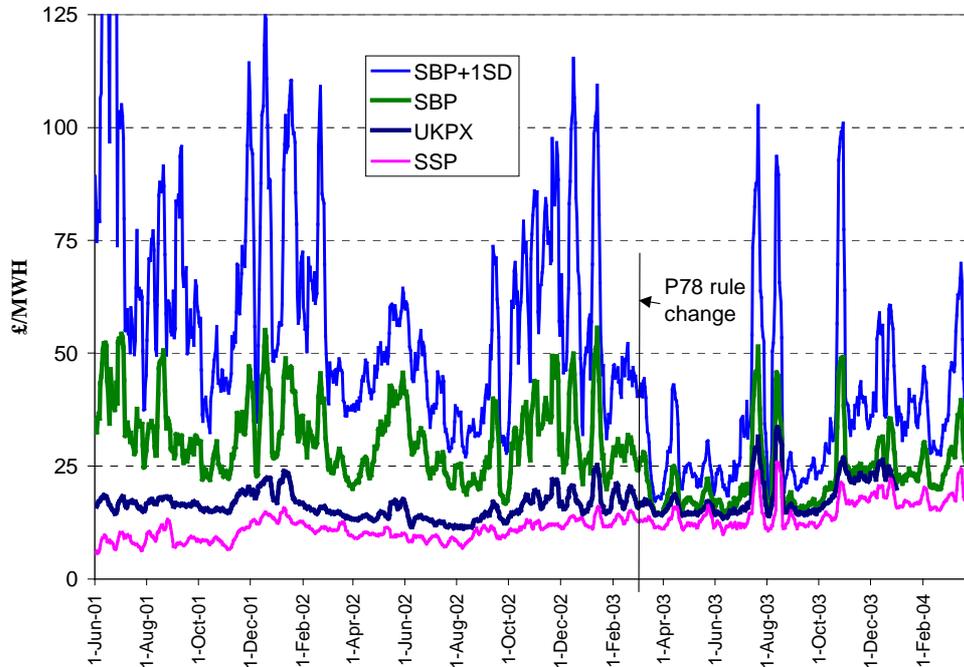
In the event the winter was mild (only 15% of winters in the past 75 years were as mild), demand was lower than the previous year, and there were no capacity shortages. The Third JESS Report in November 2003 claimed that forward markets were delivering the appropriate signals and participants were responding as they should, although the scare revealed a worrying lack of information about the status and likely time needed to return mothballed plant. The impact of plant removal and narrowing reserve margins can be seen in figure 3, where the price-cost margin has returned a considerable way towards a more sustainable equilibrium (although one that in 2004 was still too low to justify new build).

What can we conclude from this expensive change in market arrangements? First, the costs of balancing have been increased, to the detriment of non-portfolio generators (i.e. new entrants and British Energy) and intermittent suppliers like wind.¹⁶ The balancing prices are considerably more volatile and unpredictable than the Pool prices that served as a more liquid balancing market. Figure 6 shows 7-day moving averages of the buy and sell prices, and, to

¹⁶ Although the net surplus of the Balancing Mechanism is recycled, there are transfers between different types of participants, while there are extra real costs in requiring all participants to replicate the SO balancing function, especially in maintaining spinning reserve.

give a sense of the risk in the SBP, gives one standard deviation of the 7-day half-hourly buy prices, as well as the underlying spot price.

Figure 6 Spot and cash-out weekly moving average prices June 2001-April 2004

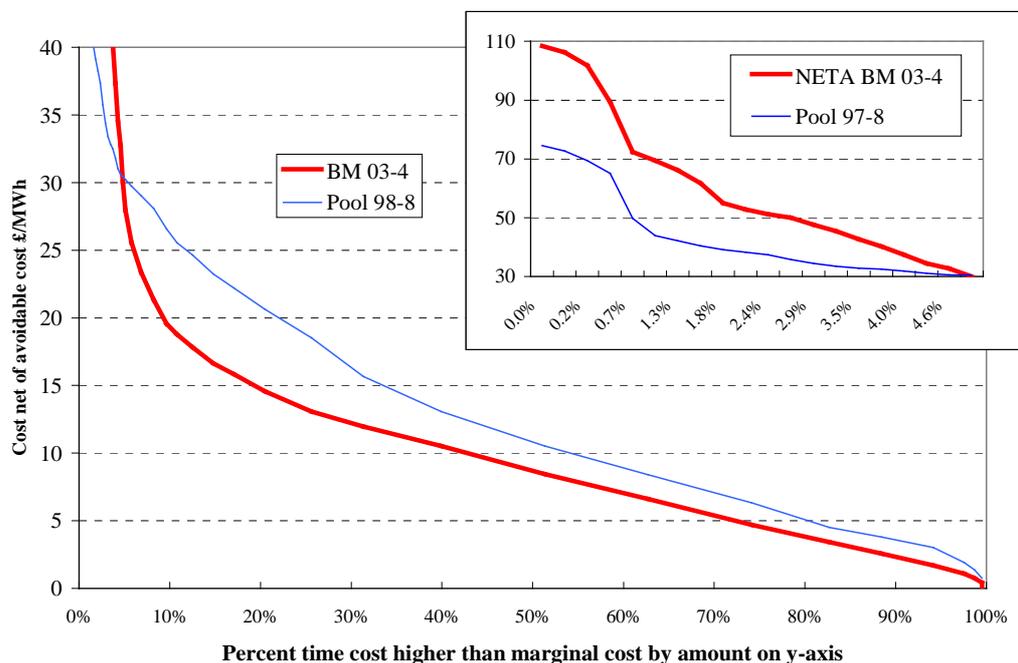


Source: Elexon price data: UKPX is the Reference Price Data for the day-ahead spot market

The costs of balancing will depend on whether the participant is a generator or supplier. As contestability is a key issue, the relevant question is whether the Balancing Mechanism (BM) unreasonably raises the risk to a small entrant. This can be estimated as the risk of having to pay the buy price (SBP) after a generator suffers a forced outage, in order to meet an assumed contract position. If a generator fails at a random moment and stays off-line for 24 hours, the cost will be the 24-hour average of the SBP from that moment. In the year before the P78 rule change indicated above, the expected cost of such an outage (relative to an assumed variable cost of £12/MWh) was £17/MWh or £0.4/kW/event compared to £13/MWh or £0.32/kW/event under the Pool for 1997-8. The variance was, however, twice as high as under the Pool. In the year following P78, the average cost had fallen to £11/MWh or £0.3/kW/event and the variance had also fallen to 150% that of the Pool.

Figure 7 illustrates the cost duration curve for balancing under NETA from April 1 2003-31 Mar 2004 compared to the Pool in 1997-98. Thus 5% of the time the cost would be £30/MWh for the following 24 hours in both the Pool and the recent BM and 1% of the time it would be £70/MWh in the BM compared with £44/MWh under the Pool. The risks in the early days of NETA were very much higher and led to claims that plant was inefficiently part-loaded to avoid penal imbalance costs, at considerably higher cost.

Figure 7 Cost of 24-hour failure under the Pool and NETA



One should interpret this finding with some care, as the Pool required bids to remain valid for 24 hours while bids and offers to the BM can be changed on a short time scale and in response to a perceived tightening of the market when a large unit goes off-line, making it more risky for generators to handle outages. Even if we ignore such responses, if a large plant were to go down, the demand in the BM might be such as to considerably increase the short-run cost, but without knowing the shape of the bids and offers it is hard to estimate by how much.

The effect on suppliers is that they would over-contract on average as the SBP is more penal than spilling the surplus at the SSP, and this would (slightly) raise the cost and risk of selling. This may have the desirable effect of encouraging contracting, which tends to mitigate generator market power in the spot market, although at the expense of increasing demand and hence market power in the contract market. The low liquidity of most electricity markets (and certainly the British markets) makes the cost of rebalancing contract positions high and again acts as an entry barrier.

Second, the BM mutes scarcity signals by paying generators their bid price and not the marginal price (in order to mitigate market power and possibly reduce volatility). This, together with the lack of integration with spot, forward and contract markets, and the lack of any capacity payment, makes the entry decision more uncertain and risky, and may lead to lower reserve margins. If so, then the lower reserve margins and the extra entry costs will allow a higher average wholesale price, refuting Offer's claim that NETA alone (i.e. regardless of market structure) would reduce wholesale prices by 10%. Again, this claim needs to be examined carefully, for example, for a peaking generator that can offer balancing services of an hour duration at very short notice. If this generator burns distillate at a cost of

£50/MWh, then if it were to bid the SBP it would earn £21/kW/yr under NETA (net of fuel costs) running the 409 hours the SBP exceeded £50/MWh, and £12.6/kW/yr under the Pool running the corresponding 662 hours. One obvious problem with this calculation is that while it is riskless to offer capacity to the Pool at the avoidable cost of £50/MWh, it requires skilled bidding to achieve the SBP in a pay-as-bid market. Nevertheless, while a peaking generator might have some difficulty paying its grid charges (which vary across the country) from +£9/kW/yr to -£7/kW/yr) and other fixed costs (perhaps £6/kW/yr) under the Pool, NETA appears to provide reasonable incentives apart from the problem of bidding. This could be circumvented by offering such services to NGC on contract to bid into the BM.

The early complaints of wind generators and CHP that they were discriminated against have not been adequately tested against more recent market conditions – certainly CHP output dropped dramatically but that was arguably because of the adverse spark-spread (electricity *less* gas cost). Wind power now typically sells on contract to supply companies who can better manage the imbalance risk within their entire portfolio, and is in any case massively rewarded by Renewable Obligation Certificates that increase the electricity price from around £25/MWh to £65/MWh.

5 CONCLUSIONS

The British experiments have demonstrated a number of important lessons for electricity market liberalisation. First, ownership unbundling of transmission from generation helps support a competitive wholesale market, which in turn puts pressure on companies to reduce costs. Scotland failed this test and failed to improve its performance. Second, efficient pricing of scarce interconnector capacity and charging correctly for losses might have allowed the Scottish market to import English competition, but was blocked by the courts. Third, while competition drives down costs, concentrated markets can sustain inefficiently high price-cost margins. Pivotal generators retain market power that is best addressed by reducing concentration, although entry that increases the reserve margin also helps. Tacit co-ordination is likely given electricity market characteristics, and is best addressed by encouraging contracts and reducing concentration. Fourth, investment in generation can be facilitated by a transparent Pool, domestic franchises, and wholesale market power. Britain replaced nearly one-third of its (already adequate) capacity by such means. Entry (including returning mothballed plant to service) is responsive to price signals (the forward spark-spread).

Fifth, unbundling and liberalisation increases risk for generators and encourages them to seek vertical integration with suppliers. This offers the opportunity for the regulator and competition authorities to trade horizontal for vertical integration and to reduce concentration, at the cost of increased entry barriers. A better alternative is to start from a more fragmented structure. That would allow one to consider legal restraints on such vertical integration to encourage more contracting and market liquidity, but we lack evidence on the costs and advantages of such enforced competition.

Sixth, the British contractual approach to liberalisation that requires licences to be held by both potentially competitive and natural monopoly segments has worked better than many of the Continental alternatives (and arguably the US's onerous duty on regulators to deliver "just and reasonable" prices). Licences require the holders to provide the regulator

with the information needed for adequate market monitoring, and allow market abuses to be swiftly and cheaply addressed.

Finally, such apparently basic issues as the desirability or not of capacity payments (or obligations) and the design of the wholesale and balancing markets remain unresolved. The ideal of a Pool with adequate competition, capacity payments, and a better governance structure for rule changes was never tried, and might have worked as well or better than NETA, with its emphasis on bilateral contracting and opaque balancing costs. On balance, NETA replaced the Pool's flawed governance structure by one more susceptible to incremental improvement (though at the cost of greater regulatory uncertainty), failed to increase either the liquidity of markets or the participation of the true demand side, increased trading costs, replaced capacity payments by a pay-as-bid balancing mechanism, and cost over £700 million.

Once it settled down and the obvious changes were made, NETA probably delivers similar outcomes as the Pool from existing generation. Entry is now more difficult than before, but that is not solely due to NETA. Vertical integration has reduced the demand for suppliers to contract, the end of the domestic franchise has removed the logical counter-party to contracts with new independent generators, but the removal of the Pool as a market of last resort almost certainly raises entry costs. Just at the time that FERC has embraced the concept of a Pool (with locational marginal pricing) as the benchmark for the Standard Market Design, Britain has abandoned a model whose main failing was its poor market structure and governance.

REFERENCES

- Bohn, R.E., Caramanis, M.C. & Schweppe, F.C., 1984, 'Optimal pricing in electrical networks over space and time', *Rand Journal of Economics*, Vol. 15, No. 3, pp. 360-376.
- Bower, J. (2002) 'Why Did Electricity Prices Fall in England and Wales?: Market Mechanism of Market Structure?', Oxford Institute for Energy Studies Working Paper, EL O2.
- Currie, D. (2000) "The New Electricity Trading Arrangements in England & Wales: A Review", The Beesley Lectures on Regulation, The Institute of Economic Affairs and the London Business School, 10 October 2000.
- Evans, J. and Green, R. (2003) 'Why did British electricity prices fall after 1998?' CMI WP26 (available at <http://www.econ.cam.ac.uk/electricity/publications/wp/index.htm>)
- Hortacsu, A. and S.L. Puller (2004), 'Testing strategic models of firm behavior in restructured electricity markets: a case study of ERCOT', CSEM WP 125 (available at <http://www.ucei.berkeley.edu/pubs-csemwp.html>)
- von der Fehr, N-H. M. and D. Harbord, (1993), 'Spot market competition in the UK electricity industry', *Economic Journal*, 103, 531-46.
- Fabra, N., N-H von der Fehr and D. Harbord, (2004), 'Designing electricity auctions', CSEM WP 122 (available at <http://www.ucei.berkeley.edu/pubs-csemwp.html>)
- Green, R.J., and D.M. Newbery, (1992), 'Competition in the British Electricity Spot Market', *Journal of Political Economy*, 100(5), October, 929-53.
- Green, R.J. (2004) 'Did English Generators play Cournot? Capacity withholding in the Electricity Pool', CMI WP41 (available at <http://www.econ.cam.ac.uk/electricity/publications/wp/index.htm>)
- Henney, A. (1994) *A study of the privatisation of the Electricity Supply Industry in England and Wales*, London EEE Ltd.
- Henney, A. (2001) 'The illusory politics and imaginary economics of Neta', *Power UK*, March, 16-26

- House of Commons (1993) *British Energy Policy and the Market for Coal: Report HC 237*. London: HMSO
- Klemperer, P. D. and M.A. Meyer, (1989), 'Supply Function Equilibria in Oligopoly under Uncertainty', *Econometrica*, 57(6) Nov. 1243-1277.
- MMC (1996) *National Power plc and Southern Electric plc: A Report on the proposed merger*, Monopolies and Mergers Commission Cm 3230, London: HMSO
- Newbery, D. M. (1992) 'Supply function equilibria', mimeo, Cambridge
- Newbery, D. M. (1995) 'Power Markets and Market Power', *Energy Journal* 16(3), 41-66
- Newbery, D. M. (1998a) 'Competition, Contracts and Entry in the Electricity Spot Market', *RAND Journal of Economics*, 29(4), Winter, 726-49. (DAE Working Paper 9707)
- Newbery, D. M., (1998b) 'Pool Reform and Competition in Electricity', ch.5, pp117-166 in M. Beesley (ed.) *Regulating Utilities: Understanding the Issues*, London: Institute of Economic Affairs
- Newbery, D. M. (1998c) 'The Regulator's Review of the English Electricity Pool', *Utilities Policy*, 7(3), 129-41
- Newbery, D.M. and T. McDaniel (2003) 'Auctions and Trading In Energy Markets - An Economic Analysis', Ch 10, pp195-234 in *CRI Regulatory Review 2002/ 2003* P. Vass (ed.) Bath: Centre for the Study of Regulated Industries.
- Newbery, D. M. and M. G. Pollitt (1997) 'The Restructuring and Privatisation of the CEGB- Was it worth it', *Journal of Industrial Economics*, XLV (3), 269-303
- NGC (1997), *Report on the Tender for the 1997/98 Standing Reserve Service*, National Grid Company plc.
- NGC (2004) 'National Grid Company Consultation on Options for Allocating GB Transmission Access Rights under BETTA', 21 May 2004, Ofgem web-site
- Offer (1992). *Review of Pool Prices*, Office of Electricity Regulation, Birmingham, December 1992.
- Offer (1998a) *Review of Electricity Trading Arrangements: Background paper 1 - Electricity Trading Arrangements in England and Wales*, February, Birmingham: Office of Electricity Regulation
- Offer (1998b) *Review of Electricity Trading Arrangements: Background Paper 2 - Electricity Trading Arrangements in other Countries*, February, Birmingham: Office of Electricity Regulation
- Offer (1998c) *Review of Electricity Trading Arrangements: Working Paper on Trading Inside and Outside the Pool*, March, Birmingham: Office of Electricity Regulation
- Offer (1998d) *Review of Electricity Trading Arrangements - Interim Conclusions*, June, Birmingham: Office of Electricity Regulation
- Offer (1998e) *Review of Electricity Trading Arrangements*, July, Birmingham: Office of Electricity Regulation
- Offer (1999) *The new electricity trading arrangements (Volume 1)*, July, Birmingham: Office of Electricity Regulation
- Ofgem (2002) *Annual Report* London: Office of Gas and Electricity Markets
- Patrick, R.H. and F.A. Wolak (1997), 'Estimating the customer-level demand for electricity under real-time pricing', paper presented to the POWER conference, March, Berkeley: University of California Energy Institute
- Pollitt, M.G. (1999), 'The Restructuring and Privatisation of the Electricity Supply Industry in Scotland' Cambridge: mimeo (available at <http://www.econ.cam.ac.uk/electricity/people/pollitt/esiscotland.pdf>)
- Robinson, G. (2001) *The Unconventional Minister: My life inside New Labour*, Penguin
- Shuttleworth, G. (1999) 'The Electricity Industry 1997-98', ch. 2, pp 19-34 in *Regulatory Review 1998/9*, ed. P. Vass, CRI, University of Bath
- Sweeting, A. (2001) 'Market outcomes and generator behaviour in the England and Wales wholesale electricity market 1995-2000', mimeo, MIT