

Investment Incentives and Electricity Market Design: the British Experience

FABIEN A. ROQUES *

Judge Institute of Management, University of Cambridge, England¹

DAVID M. NEWBERY

Department of Applied Economics, University of Cambridge, England

WILLIAM J. NUTTALL

Judge Institute of Management, University of Cambridge, England

Abstract

There is no academic consensus on which electricity market design provides the least distorting investment incentives. Theory suggests that “energy-only markets” can allow capacity cost recovery by generators. However, separate payments for capacity or reserve obligations do not need to rely on infrequent price spikes to remunerate reserve capacity. Three years after the controversial change from the compulsory British Electricity Pool with capacity payments to the decentralised energy-only New Electricity Trading Arrangements (NETA), we contrast the two market designs in terms of investment incentives, analyse NETA’s balancing market failures, and review the case for regulatory support for investment.

1 Introduction

Industries that are capital intensive and produce homogenous products or services frequently exhibit investment booms and busts. Where the product is traded on competitive markets, prices can be extremely volatile. Aluminium provides a good example, as do computer chips, and more recently, airlines. The electricity industry is very capital intensive, with variable costs typically half average costs, and electrons are completely

* Contact author: Judge Institute of Management, University of Cambridge, Trumpington Street, Cambridge, CB2 1AG, England. Email: f.roques@jims.cam.ac.uk The authors would like to thank three anonymous referees and the participants of the 3rd International Conference on Applied Infrastructure Research, “Network Economics: Financing, Regulation and Capacity Allocation in Infrastructure Sectors”, Berlin, 9 October 2004 for their helpful comments. We would also like to thank BNFL for their support in investigating issues of security of supply, although the views expressed here are solely those of the authors and not necessarily those of BNFL.

¹ Research support from Cambridge-MIT Institute under the project 045/P “Promoting Innovation and Productivity in Electricity Markets” is gratefully acknowledged, as well as Heren Energy for providing the data on electricity and gas prices.

homogenous. Like airline seats, but unlike aluminium, electricity is essentially non-storable, so that competitive spot markets are likely to be even more volatile over short periods of time. Under the old-style vertically integrated regulated (or state-owned) franchise monopoly structure, investment booms and busts were not eliminated, but prices were regulated and were typically very stable. In well-managed regimes, and notably under cost-of-service regulation, prices were set to cover total costs and hence enable the utility to finance investment.

The recent wave of electricity liberalisation has dramatically changed this structure, and in many countries the vision is that generation is a potentially competitive industry whose prices should be set on an unregulated wholesale market, confining regulation to the natural monopoly transmission and distribution businesses. Similarly, investment in generation should be left to market participants, responding to current and expected market prices. The case for such liberalisation is that competitive markets provide better incentives to deliver efficiency than is possible under regulation or state-ownership. The evidence supports this claim, at least where the wholesale market has been reasonably well designed. Labour productivity has increased, other operating costs have been reduced, and plant operated more efficiently (Newbery and Pollitt, 1997; Markiewicz and Wolfram, 2004). However, there is as yet insufficient experience to assess the long-term benefits from liberalising the electricity industry. The key issue is whether a liberalised electricity industry can deliver *sustainably* competitive outcomes; that is whether there will be both adequate and timely investment, built at least cost and delivering competitive electricity prices. The objective of this paper is to shed light upon the interactions between market design, market structure and investment incentives by reviewing the British experience of its two successive and radically different market frameworks. In so doing, we hope to illuminate the determinants of investment incentives in capital-intensive industries more generally, and show how a range of supporting market institutions (various kinds of contracts, spot and forward markets, as well as regulatory interventions) influence the outcomes.

In the early years of liberalisation, the focus of academic research and regulatory scrutiny concentrated mainly on short-term market efficiency and competitiveness. As the first territories to liberalise – among which are included England and Wales and some U.S. states – have now reached the end of their first investment cycle, much attention is being paid to assessing the long-term dynamic performance of the liberalised electricity industry. The initial signs in England and Wales were promising: new capacity of about one quarter of the existing (adequate) capacity was added in the first decade after liberalisation. This initial investment wave was prompted by an initial market structure that sustained prices at above entry levels. Excess entry and subsequent over-capacity followed by (regulator-encouraged) market restructuring created intense competition that squeezed price-cost margins and created financial difficulties for many generating companies, pressure from their creditors, and a subsequent reluctance to invest in new plant in these markets. Elsewhere, other electricity markets are just beginning to approach their first major investment cycle as surplus capacity is absorbed (or withdrawn), so that the decisive test for liberalised electricity markets remains ahead. This is the case in the more mature European continental markets such as The Netherlands, Germany, and the Nordic countries.

The reliability of electricity supply has been one of the overriding concerns guiding the restructuring of the electric power industry. The policy imperative of "keeping the lights

on" has been the principal motivation for many technical and economic constraints imposed on market designs. In Britain, there is considerable Government concern whether generating capacity will continue to be adequate to ensure security of electricity supply, and whether investment signals are timely and properly reflect the social profitability of investment, given the turbulent experience of recent investments. The market now appears considerably more risky than it did a few years ago. This raises the hurdle rate of return, and may bias investments against capital-intensive base-load plant and risky peaking capacity, undermining the quest for efficient and secure supply. Other market failures such as investor risk aversion, regulatory uncertainty, and lack of long-term contracting could also distort investment incentives. Do these market failures constitute a strong enough case for regulatory intervention to support investment? Is there any "best-practice" market design that can ensure generation adequacy in the long run at least cost while minimising regulatory interference with the market?

There is no consensus among academics on which market design provides the least distorting long-term investment incentives. Theory suggests that energy-only markets with spot prices that are allowed (and expected) to fully reflect scarcity rents will generate sufficient income to allow capacity cost recovery by generators. Hence as far as supply adequacy is concerned, a well-functioning energy-only market can provide the correct incentives for generation adequacy. The critical role of electricity in the economy and the political ramifications of widespread electricity shortages or price spikes have prompted many regulators around the world to take steps above and beyond reliance on market forces in order to ensure generation adequacy. Different market designs, with separate payments for capacity or reserve obligations have the advantage of not relying on infrequent price spikes (of possibly uncomfortably long duration) to remunerate reserve capacity. Three years after the controversial change of the British market design from a compulsory gross Pool with capacity payments to the New Electricity Trading Arrangements (NETA) with its decentralised pay-as-bid energy-only market framework, this paper contrasts these two "extreme" market designs in terms of investment incentives. It also highlights the market design changes and regulatory actions that appear necessary to maintain power generation investment and long-term generation adequacy.

The first section introduces marginal pricing theory in electricity markets, and reviews the practical issues that render "energy-only" markets more risky than other markets design remunerating generators through a capacity mechanism on top of the energy sales. The second and third sections illustrate how market design issues impact on investment incentives by concentrating on the British experience: the former centralised Pool market and the current decentralised NETA market arrangements are compared as regard to investment incentives. The fourth section concentrates on the main market failure of NETA, namely the imbalance pricing mechanism, which does not reflect capacity scarcity appropriately. The fifth section studies the impact of the lack of long-term contracting on investment decisions of power investors, showing that vertical and horizontal concentration provide an organic hedge at the possible expense of market power. The last part reviews the regulatory actions that might reduce the investment distortions identified in the previous sections, considering in particular whether regulatory intervention to encourage investment is justified, and if so which form it should take.

2 Investment incentives and electricity market design

In an “energy-only” market, there is no separate payment for energy and capacity; the primary income source for recovery of capital costs is the difference between the market clearing price, or the contract price in a pay-as-bid system such as NETA, and the generators’ marginal costs. When ancillary services are procured separately by the system operator, generators can earn additional revenue by selling ancillary services, such as regulation and spinning reserve capacity. Energy-only electricity markets have been adopted in the original (defunct) California design, in Nordpool, the Australian Victoria pool (although with an *ex post* Value of Lost Load, or VOLL) and in the British NETA design.²

The theory of spot pricing claims that electricity spot markets can provide efficient outcomes both in the short and in the long run (see Caramanis et al., 1982 and Caramanis, 1982). The conditions for this to hold are strong – the usual General Equilibrium assumptions of a complete set of spot and forward markets or perfect foresight, price-taking behaviour by producers and consumers, risk neutrality (or adequate risk-sharing contracts), and convex production possibilities. In such a world, the spot prices should result in an efficient dispatch and allocation of available resources and, together with forward prices and rational expectations, should also signal the need for additional generating capacity. The optimal capacity stock should be such that scarcity payments to the marginal generators when demand exceeds supply exactly cover the capacity cost of these generators. A shortage of capacity will increase scarcity rents, producing profits in excess of what is needed to cover the amortised capacity cost. Such profits will attract generation expansion. On the other hand, excess generation capacity will eliminate scarcity rents driving prices to marginal cost. When this occurs, generators on the margin will not be able to cover their investment cost. Unless such generators receive extra revenues through some form of capacity payments this will result in early retirement or mothballing of plants which will reduce capacity and drive prices back to their long-run equilibrium level.

Of course, these ideal conditions never occur, but the same could be said of all other markets, where (apart from legitimate competition concerns) we are normally prepared to accept the outcome as workably competitive and feasibly efficient, and not requiring specific policy interventions. The question is not whether electricity markets will deliver perfect outcomes, but whether the specific characteristics of electricity introduces systematic biases in market behaviour that require more complex market designs and regulatory requirements.

The special features of electricity have indeed been invoked to argue against relying on energy prices to cover capacity costs through scarcity rents. The non-storability of electricity, demand and supply uncertainty, inelastic demand and the steepness of the supply curve at its high end all contribute to high price volatility when reserve margins are low. In turn, this volatility might distort investment signals, by biasing investment against risky peak-load plants and capital-intensive technologies. Reserve capacity is provided by peaking units running only at times of high prices. As there is no separate payment for capacity, reserve generation is paid for by infrequent price spikes in periods of scarcity, as shown by *figure 1*. In theory, the scarcity rent that these peaking units are able to earn at

² The value of lost load (VOLL) is defined as the marginal value of unserved load, and represents in practice the value that an average consumer puts on the last kWh of electricity before being disconnected.

times when generation is scarce should allow them to recover both their fixed and operating costs.

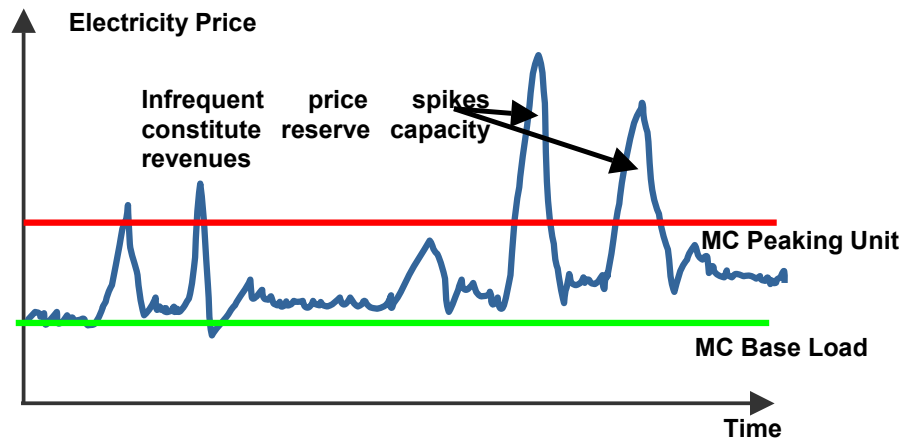


Figure 1: Energy-only markets rely on infrequent price spikes to reward reserves

Relying on periodic price spikes to signal the need for peaking capacity has some significant weaknesses. Strbac and Kirschen (2000) argue that the uncertainty of the returns leads risk-averse investors to under-invest in peaking units. Neuhoff and de Vries (2004) model the impact of risk aversion on investor investment choices and consumer contracting decisions. They show that risk-averse investors who fund their generation investment with spot market returns will provide less generating capacity than risk-neutral investors. The logical solution is for generators to offer one-sided contracts for differences with a high strike price. This allows consumers to hedge against the price spikes (as their maximum payment is the strike price) while the revenue from the sale of these contracts provides the capacity payment needed to provide power in the peak periods.

Further, Grimston (2004) claims that unbundled electricity markets discriminate against capital-intensive technologies such as nuclear and renewables. Indeed, while the various risks facing an investor in liberalised electricity markets affect all generation technologies, technologies which have a higher specific investment for capacity (for example, nuclear or wind) might appear to be more exposed to electricity price risk.³ A firm relying on a capital-intensive technology may be competitive in terms of short-run marginal cost, but will be more exposed to electricity price risk in the long-term to cover the cost of capital employed.⁴

The most convincing argument against energy-only markets lies in the difficulties for regulatory authorities to distinguish between the exercise of market power and legitimate scarcity rent (Fraser, 2003). While some temporary high prices reflect legitimate economic signals that are needed in order to attract investment and reduce demand, they might be

³ In particular, in electricity markets where open-cycle gas turbines set the spot market clearing price, strong correlation between gas and electricity prices is observed, reducing gas-fired plants exposure to gas and electricity price risks, while increasing other technologies' exposure to volatile electricity prices.

⁴ Note that the lower is the variable cost, the higher the mean profit will be, and the lower the variance of profits will be. This is because profits will be bounded below by zero more often the higher is variable cost, so the risk is lower for running such technologies. The problem is that the capital at risk is higher.

politically unacceptable if it is impossible to differentiate between legitimate scarcity rents and high prices resulting from market power abuse or gaming flawed market rules. In response, regulators frequently suppress high energy prices by designing market power mitigation schemes, such as bidding price caps. These price caps may in turn create a revenue deficiency for peaking generators that may cause insufficient investment in generation capacity. The perceived threat of regulatory intervention to curb scarcity rents alone may be sufficient to inhibit capital intensive investments and raise the capital cost for investment in generation capacity, leading to higher average prices but no greater security. Thus, regulatory support for investment can be viewed as a remedy to offset energy price caps (or perceived threat of caps) to ensure adequate investment in reserves.

2.1 Alternative designs for supporting investment

Regulators have been concerned that energy prices in some restructured systems are not high enough to prompt adequate investment, and so have implemented various support schemes. Arriaga et. al. (2002) review the different capacity mechanisms available to regulators to encourage investments in generation. Capacity mechanisms pay generators in exchange for the undertaking to supply electricity if required. In one version of capacity mechanisms, the regulator sets a price for capacity and lets the market determine the amount of capacity available. In the other version, the regulator sets the amount of capacity that has to be available and lets the market determine its price. These are known, respectively, as capacity payments and capacity requirements. Britain adopted capacity payments under the Pool, as did Spain and several Latin American countries. The eastern pools in the U.S. including PJM, NYPP, and New England ensure generation adequacy by imposed installed capacity requirements (ICAP) on load serving entities (LSEs).

The concept of capacity payment is rooted in the theory of peak load pricing, whose application to electricity was pioneered by Boiteux (1949). According to this theory, generation of electricity requires two factors of production, capacity and energy, where the amount of energy that can be produced in any given time period is constrained by the available capacity. Energy is efficiently priced at marginal cost and an adequate charge is imposed on the peak-period energy users.

Subsequent developments of peak load pricing theory focused on two important aspects of electricity supply: uncertainty and technology mix (Chao, 1983). The effect of uncertainty leads to redefining the basic ingredients of electricity service as energy and reliability, where reliability is manifested by the loss of load probability (LOLP) calculation as a function of available capacity relative to load. The distinction between peak and off-peak then becomes a matter of degree (and can be determined endogenously). This perspective rationalises levying a time-varying capacity charge on all consumption, and its payment to available generation capacity whether or not dispatched, as such capacity provides added reliability.

Under optimal capacity planning the marginal cost of incremental capacity equals the marginal cost of unserved load, which can be approximated by the marginal Value of Lost Load (VOLL) times the probability or fraction of time that load must be curtailed due to insufficient capacity (LOLP). Under the British Pool, available generators received such a payment. The alternative method for calculating capacity payments (which, in theory, is equivalent to VOLL times LOLP under optimal capacity configuration) is to base the payment on the cost of peaking technology (for example, a combustion turbine) (see also Joskow and Tirole, 2004b).

The motivation for installed capacity requirements (ICAP), sometimes also called capacity obligations, is similar to that for capacity payments. The resulting capacity markets provide generators with the opportunity to collect extra revenue for their unutilised reserve generation capacity and provide incentives for the building of adequate reserves. The supply of adequate capacity can be determined either through prices (capacity payments) or through quantities (capacity obligations). The case for quantity control can be supported by the classic prices vs. quantities argument of Weitzman (1974). The basic argument is that the demand function for capacity is nearly vertical while the supply function is flat. Thus a small error in price will result in a large error in quantity, so that direct quantity control is superior (Oren, 2003).

2.2 The drawbacks of existing capacity mechanisms

Capacity mechanisms have been criticised for being based on engineering calculations. Either the price (VOLL) or the quantity (the suppliers' capacity obligations) have to be administratively determined by the system operator (Arriaga et al., 2002). Chuang and Wu (2000) suggest that VOLL figures based on demand side bidding would provide better estimates. However, the fundamental issue lies in the fact that VOLL varies among consumer categories, and with the duration of the supply interruption as well as whether the interruption is scheduled and consumers forewarned or not. In Britain, demand-side bids were accepted up to VOLL, but the schedule was very non-linear. An efficient capacity payment would have to vary by consumer-type and perhaps the expected duration of outage, leading to very complex calculations. Choosing a second-best average VOLL is the most that is likely to be feasible, coupled with demand-side bidding and/or interruptible contracts for large consumers.

Capacity requirements imposed on suppliers and ICAP markets suffer a similar problem: the allocation of capacity among the different suppliers is done *ex ante* by the system operator according to engineering simulations. However, many electricity dispatch constraints only manifest themselves in real time, and thus an *ex post* revision of the capacity requirement of each supplier would be required. Besides, Oren (2003) argues that one of the fundamental problems with capacity markets is their disconnectedness from energy markets. In the long run, the expected social cost of unserved energy as reflected by the energy-only market prices should equal the marginal cost of incremental capacity. However, the separate capacity markets created for trading reserve capacity requirement set through engineering based methods may produce prices that are not in equilibrium with the energy market prices. For instance, overestimating the expected cost of lost load will create artificially inflated demand for capacity and result in high capacity prices, which in turn will lead to overcapacity that results in suppressed energy prices and socially inefficient production and consumption.

Another problem that capacity payments and obligations share lies in the LOLP calculation. Graves et al. (1998) and Hirst and Hadley (1999) point to the fact that the LOLP calculations often employ simplistic models of probabilistic failure (for example, Poisson arrivals) and do not account for more complex phenomena such as the incentives of operators to keep plants running during peak price periods. Both the arbitrariness in the VOLL and the approximate nature of the LOLP calculation are likely to result in a mismatch between energy market prices and capacity values set directly or via a capacity market induced by capacity obligations (Newbery, 1998b). Furthermore, the predictability

of calculated capacity payments can lead to gaming and manipulation of the payments in concentrated markets (Newbery, 1995).

Most importantly, capacity payments or mechanisms that attempt to mitigate price volatility can be thought of as targeting the symptoms rather than the cause of electricity mechanism deficiencies. Fraser (2003) argues that the lack of demand response in electricity markets is the heart of the problem. A bidding price cap coupled with a capacity mechanism may mitigate the exercise of market power and decrease market volatility, but may also, as a side effect, further reduce demand responsiveness as consumers no longer face high prices. Any regulatory intervention requires caution, since it may introduce secondary side-effects or perverse bidding incentives making further regulatory intervention self-perpetuating. This has clearly been the case in Argentina, where large capacity payments that are paid on the basis of generated energy induce generators to bid below marginal cost so as to increase production and capacity payment revenues.

The British liberalisation experience reviewed next is particularly relevant to this theoretical debate, as England and Wales have experimented successively two “extreme” market designs, namely the Pool market with a capacity payment from 1990-2001, and the NETA market without capacity payment after 2001.

3 Power investments in the English and Welsh Pool (1990-2001)

The restructuring and privatisation of the electricity industry in England and Wales in March 1990 profoundly modified institutional and regulatory structures (Newbery and Pollitt, 1997). The Central Electricity Generating Board (CEGB) was split into three generation companies and the National Grid Company (NGC), which owns and operates the transmission network. Formal responsibility for generation adequacy was with the twelve Regional Electricity Companies (REC). They were deemed to meet it as long as they bought from the Pool (at prices that could rise as high as the value of lost load).

3.1 Investment incentives and the Pool design

The Pool (meaning the Electricity Pool of England and Wales), a mandatory auction spot market, was established during 1990. The Pool operated a day-ahead market. The least-cost unconstrained schedule then determined the system marginal price (SMP) as the most expensive generating set (genset) required to operate in each half hour, assuming that there were no transmission constraints. Newbery (1998a, b) provides a detailed explanation of the market rules.

Capacity payments were made to each genset declared available to operate in each half hour, and were set equal to the loss of load probability (LOLP) multiplied by the excess of the value of lost load (VOLL)⁵ over the station's bid price (if not dispatched) or the SMP (if dispatched).⁶ The sum of capacity payments and the SMP made up the Pool Purchase Prices. The Pool Selling Price included the cost of ancillary services and constraint costs.

The Pool design facilitated entry by independent power producers, and to that extent contributed to the decrease in market concentration (Newbery, 2004). The opportunity to

⁵ VOLL was set administratively at £2,000/MWh in 1990 and was then increased annually by the RPI – in 2000, it stood at £2,816/MWh (data from NGT).

⁶ More detailed information is available on the Pool website still maintained by Elexon (<http://www.elecpool.com/index.html>).

sell all generation in a compulsory gross Pool, which was therefore a highly liquid market, provided a transparent and simple price signal on which to develop contracts for differences and other hedging financial instruments. Moreover, the franchise on which Regional Electricity Companies (REC) could write long-term Power Purchase Agreements (PPA) did much to stimulate entry and improve competition. Hunt (2002) considers that a compulsory gross pool is the “clear market design winner”. Similarly, the Council of Australian Governments report studied a number of electricity markets (Nordpool, PJM, NETA) and concluded that there were good reasons for retaining the single price gross pool design in Australia (CoAG, 2002).

The main problem of the Pool was the sustained market power of the incumbents. Newbery (1995) and Sweeting (2001) provide a review of market power leverage under the Pool. Wolfram (1999) studies the years from 1992 until 1994 and finds that the two dominant generators did not apply as high mark-ups as theoretical models of oligopolies would predict. Sweeting (2004) shows that generators exercised increasing market power in the second half of the 1990s despite declining market concentration, consistent with tacit collusion.

The relationship between market power and the capacity payment design is particularly interesting. Did the Pool’s capacity payment design facilitate the exercise of market power and gaming? If so, would this capacity payment design have worked better if the market were more competitive? Unfortunately, the test was never taken, as the Pool was abolished just as the market was becoming more competitive,

The capacity mechanism was widely criticised for being prone to manipulation through capacity withholding (Newbery, 1995). *Figure 2* shows the percentage of time in various years of the Pool during which capacity payments exceeded the Y-axis value. The last years of the Pool witnessed the highest capacity payments. Green (2004) uses a Cournot model to assess strategic withholding of capacity in order to raise capacity payments. He observes that such behaviour was not a significant feature of the early years of the Pool. The last year of the Pool recorded the highest capacity payments, while the capacity margin for 2000/01 was as high as 25.3 % (NGT, 2000).

Green (2004) suggests that these abnormal payments were not the result of strategic capacity withdrawal, but rather were caused by the anomalous way in which the availability factor of new plants was calculated. One of the major criticisms of the Pool was that it was designedly difficult to change the Pool rules, and hence hard to deal with such anomalies.

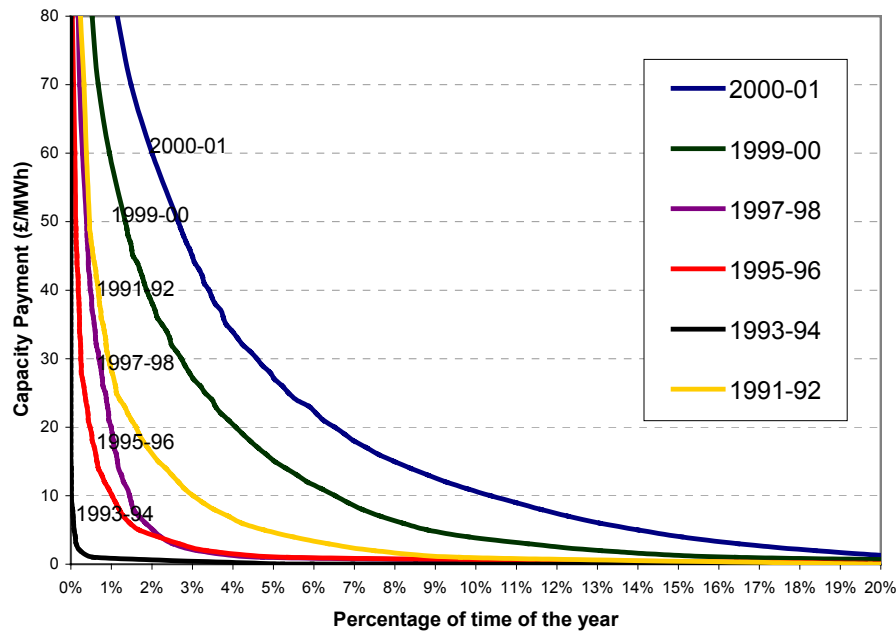


Figure 2: Cost duration curve of capacity payments (£/MWh) (data from Elexon)

3.2 The flawed design of the capacity payment

Newbery (1998b) argues that the computer program used to calculate the LOLP overestimated the chance of a power failure. The LOLP relied on the standard error in the demand forecast, and probabilities of “disappearance” of each generating unit between the date from which availability was deduced and the time of the forecast.

The way the disappearance ratios were calculated systematically underestimated the amount of capacity actually available and despatched at peak times, and made it easy for generators to game. The calculation used average availability, ignoring the fact that some plants are fully available at peak times but less available off peak. Besides, the disappearance ratio represented the probability of a plant not being available on any random day of the year, given that it was available in the previous week: it made no allowance for the various reasons why plant was not available. An unplanned outage at a relatively new plant in one month could significantly increase the LOLP and hence capacity payments in the following month, despite high reliability of the plant on a day-to-day basis. Lastly, demand reduction offered by large customers received capacity payments, but this demand reduction was anomalously not included within the VOLL calculation, thus increasing capacity payments.

The second element of the capacity payment, the administratively set VOLL, was also subject to criticism. Newbery (1998b) found indirect evidence using price elasticities that the VOLL figure was probably underestimated. VOLL estimates are rendered difficult by the lack of direct measurement of the value consumers put on electricity security of supply, as power outages are exceptional phenomena. Current estimates in use in electricity markets around the world are often arbitrarily determined, and a better approach would consist in deriving them from consumer willingness to reduce load at times of scarcity.

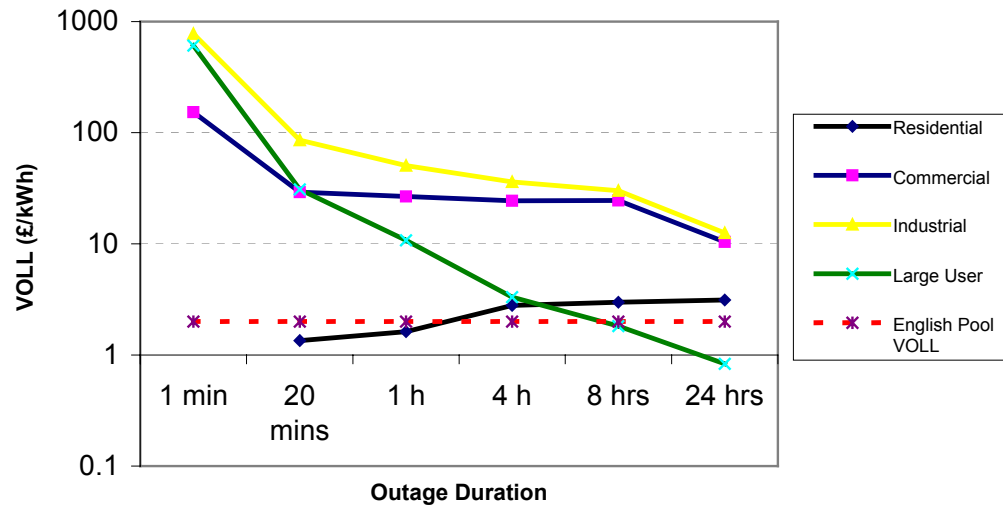


Figure 3: VOLL for different interruption durations (data from UMIST)

As exemplified by *figure 3*, the problem is that there is no such thing as a generic VOLL for all consumers. VOLL varies greatly among consumers categories, large industrial and commercial consumers having a much higher VOLL which can justify the installation of back-up systems, while individual consumers have a VOLL close to the £2,000/MWh chosen administratively in 1990. Moreover, VOLL varies with the duration of the power shortage: while VOLL decreases with time for large industrial users, which have the possibility to start back-up systems or to adapt their activity, VOLL increases with time for domestic consumers whose activity crucially depends on electricity (for example, freezer loss after a certain number of hours).

As the level of the capacity payment was determined by the product of LOLP and (VOLL-SMP), it might be that the overestimation of LOLP was partly compensated by the underestimation of VOLL. Newbery (1998b) pointed out that although each calculation of VOLL and LOLP might seem somewhat arbitrary, it is their combination that mattered to determine the level of the capacity payment and thus the investment incentives. However, the relative weight of each component actually influenced the riskiness of investing in reserve generation capacity. If VOLL had been increased, and LOLP estimate reduced so that their product remained unchanged, the capacity payment would have been higher but paid less often to generators. In other words, reserve generation plants (that is, peaking units) would have been run less often, but would have earned the same profit on average as they would have been paid higher capacity payments. Moreover, the generators would have adapted their bidding and investment strategies to the increased risk in the market for reserve capacity. Both the portfolio of plant technologies (peaking and base units) and the operational strategies would have been modified.

In the light of this example, the over-estimated LOLP combined with a relatively low VOLL might have been a “politically” strategic choice: by providing a relatively constant flow of revenues, the so defined capacity payment made investment in power plant easier than if generators would have had to rely on higher but rarer and thus more uncertain cash flows. The choice of this lower risk option for the capacity payment appears designed to encourage entry and make the market more competitive, while guaranteeing that the lights would not go out and take the gleam off privatisation. In this regard, it is interesting to

compare the British and Australian experience of determining capacity payments. In Australia VOLL is paid whenever the system has inadequate capacity to maintain supply, so that the LOLP is the actual *ex post* probability of loss of load. The VOLL was set at the British level, but several actual black-outs lead to a reconsideration of the VOLL, which had to be substantially increased to make peaking capacity commercial.

4 The New Electricity Trading Arrangements and investment incentives

In October 1997 the Government launched a *Pool Review*, and the resulting New Electricity Trading Arrangements (NETA) went live on 27 March 2001. Under the Pool all generation was centrally dispatched while under NETA each plant is self-dispatched. The obligation to balance output with contracted demand is now placed on each generator (and suppliers are similarly required to match contracted with actual demand), with the System Operator (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 National Grid Transco (NGT), the successor company to 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. Finally, capacity payments have been abandoned and there is no separate remuneration of reserve capacity: NETA is an “energy-only” market.

4.1 Investments indicators evolution since privatisation

Figure 4 shows that since privatisation in 1990, generation connections and disconnections have remained at high levels during the Pool but decreased dramatically since the implementation of NETA. It is noticeable that more than 6 GW of CCGT power stations have been given regulatory consent since 2000, but only one of them was under construction by November 2004 (JESS, 2004).

Similarly, the capacity margin, defined as the percentage of installed capacity in excess of peak demand over a given period, has fallen dramatically after the implementation of NETA in 2001 (see *figure 5*). Under the Pool, capacity payments and relatively high wholesale electricity prices resulted in large investments in generation during the second half of the 1990s, characterised by the so-called “dash for gas”, leading to excess capacity and, after 2000, falling wholesale prices. That contributed to more “mothballing” of electricity generating plant, to the postponement of construction on a number of new power stations which had received planning permission (JESS, 2004).

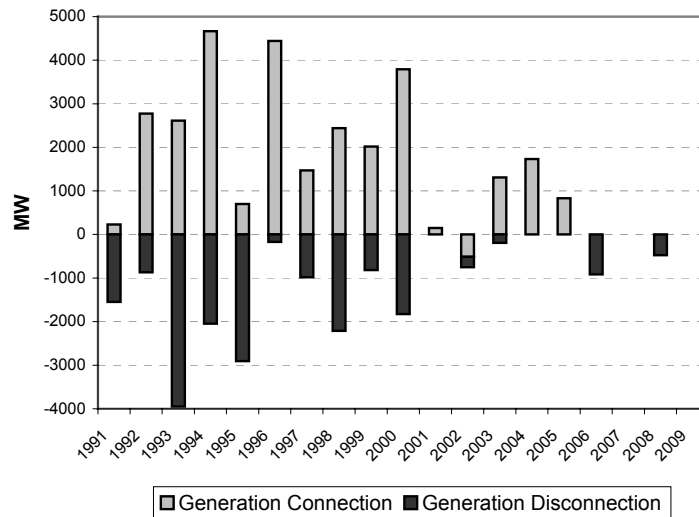


Figure 4: Past and declared future Generation Capacity Changes (NGT SYS 2004)

There is no single figure that defines an optimal reserve margin, because the reserve requirement depends greatly on the characteristics of the electricity system, namely on transmission and distribution capacities, the size of the system, pumped storage capacity, and the portfolio of generating plants. According to the International Energy Agency (IEA, 2002), “reserve levels in the range of 18 to 25% of total generating capacity are often considered appropriate”. The former national-owned CEBG operated a 24 % planning capacity margin (in the years when the time to complete a large base-load station was 7-10 years), while the current operating target of NGT is a 20 % capacity margin (noting that CCGT can be built in 2-3 years). *Figure 5* shows that under some future scenarios the capacity margin might fall below NGT’s 20 % planning target.

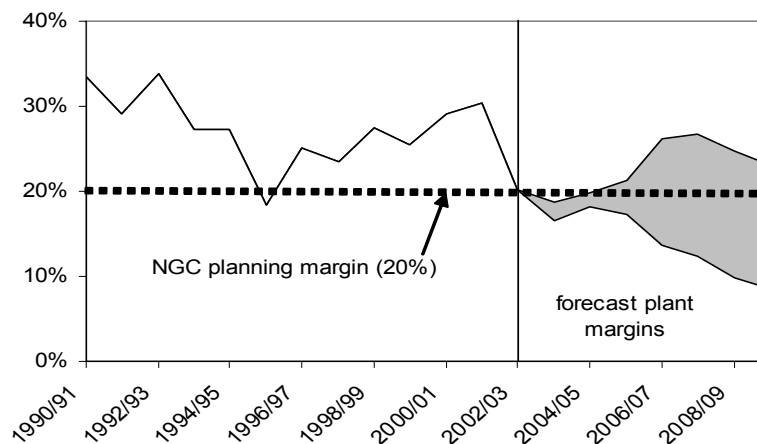


Figure 5: UK electricity system capacity margin (NGT SYS 2003)

4.2 The winter 2003/04: the first stress test of NETA?

During the year 2003, the NETA arrangements were confronted with their first “stress test”, as British capacity margin forecasts for the winter 2003/2004 dramatically fell during the spring and summer of 2003 to a low point of 16.2% in May 2003. This was the result of the closure or mothballing of a number of plants in 2002 and early 2003, following the historically low levels of wholesale prices observed during 2002. *Figure 6* shows the evolution during the year 2003 of the capacity margin forecast for the winter 2003/2004. The capacity margin for December 2003 (engineering weeks 43 to 47 – the engineering year starts on March 1) was forecast to be 8200 MW in January 2003, but only 5200 MW in May 2003. At this point NGT was concerned that such a low capacity margin forecast might not allow it to operate the system in extreme conditions during the winter 2003/04.

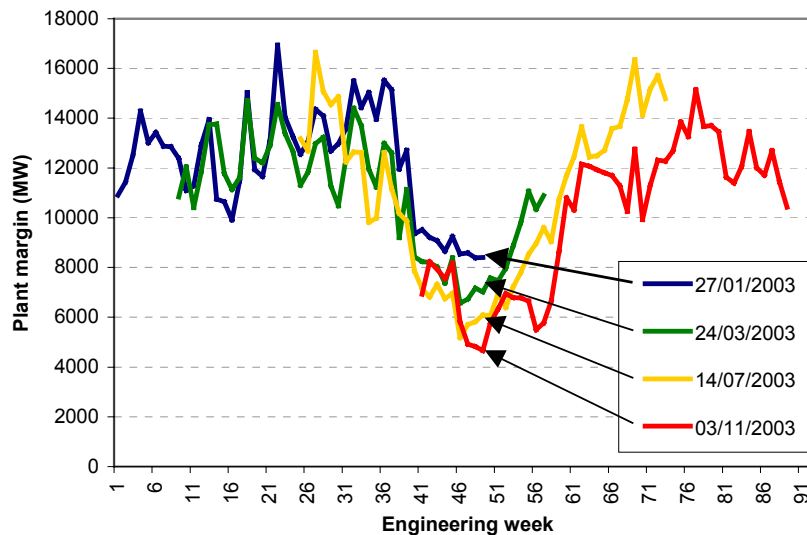


Figure 6: Evolution of Capacity Margin Forecasts 52 weeks ahead (NGT SYS 2003)

During the summer 2003, winter 2003/04 forward prices increased substantially, reaching a peak in October 2003. *Figure 7* shows the evolution of the forward base-load and peak-load electricity prices for November 2003, December 2003, and January 2004. Base load prices increased from 23 £/MWh to 33 £/MWh between July and October 2003 (43% rise), and peak load prices from 35 £/MWh in July 2003 to 55 £/MWh in October 2003 (57 % increase).

Before interpreting these electricity price trends as a market response to the forecast capacity shortage, one needs to assess the impact of the other determinants of forward electricity prices. Forward gas prices for delivery during the winter increased slightly during 2003, so that the increase in forward prices was not completely translated into increased profitability. *Figure 8* shows the forecast peak-load winter 2003/2004 spark spread (the spread between the price of electricity and the price of the gas required to generate that electricity) evolution during 2003.⁷ The forecast spark-spread for December

⁷ Data sources: Electricity and NBP gas forward prices from HEREN, assuming a plant heat rate of 7000 BTU/KWh and a thermal efficiency of 46.8% corresponding to the 2002 UK average efficiency of gas turbines according to NGT.

2003 rose from -4 £/MWh to 11 £/MWh, confirming that most of the forward electricity prices increase was due to the forecast capacity shortage.

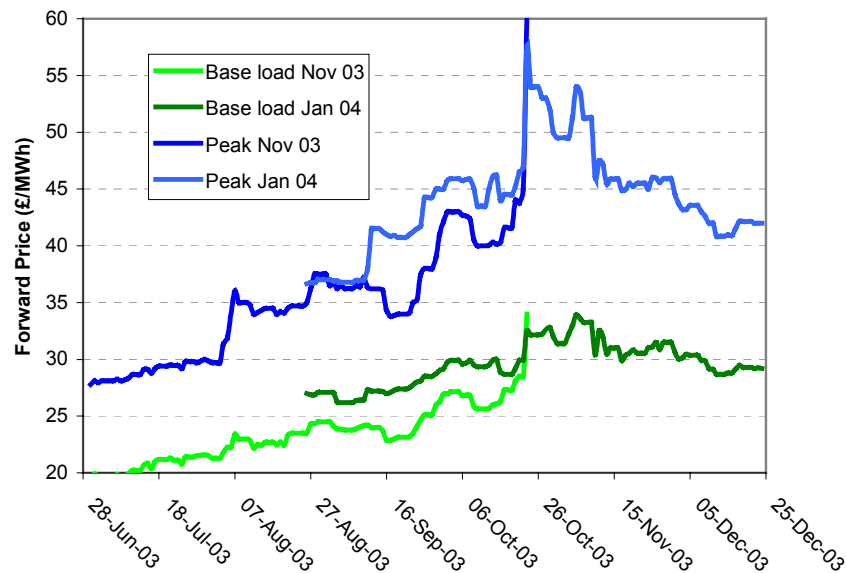


Figure 7: Winter 2003/04 forward electricity prices (£/MWh) (data from Heren)

These price increases during the summer 2003 motivated some generators to bring back on stream some mothballed plants for the winter 2003 between August and December 2003. This increased the capacity margin of the system to more secure levels: in January 2004, NGT reported a capacity margin at 21.6%, as compared to the forecast 16.2% in May 2003. *Figure 7* shows that base load and peak-load prices subsequently settled down respectively around 28 £/MWh and 42 £/MWh during the winter. OFGEM concluded in its retroactive winter 2003 report that the actions taken by firms during the second half of 2003 were to be expected in a competitive market whereby the availability of capacity responds to prices, which reflect underlying supply and demand conditions (OFGEM, 2004a).

One needs, however, to be careful when interpreting the quick response of the market to the forecast capacity shortage during the winter 2003/04. When looking more closely at the role played by the system operator (NGT), it seems hard to share OFGEM's assessment that this "stress test" demonstrated the robustness of NETA to capacity shortages. Throughout the summer 2003, NGT warned generators during industry meetings that under some cold weather scenarios the Operational Planning Margin Requirement (OPMR) would not be met in a number of weeks during the coming winter. Most importantly, following pressures from OFGEM and analysis of its likely reserve requirements for winter 2003/04, NGT initiated a Supplemental Standing Reserve Tender (SSRT) on 14 October 2003 to increase its reserve capacity. The tender closed on 27 October 2003, and NGT procured a total of 852MW of Supplemental Standing Reserve (SSR) at a total cost of £18.87 million or £22/kW (OFGEM, 2004a). The majority of this volume was provided from plant that had previously been mothballed and a significant amount came from the demand side.

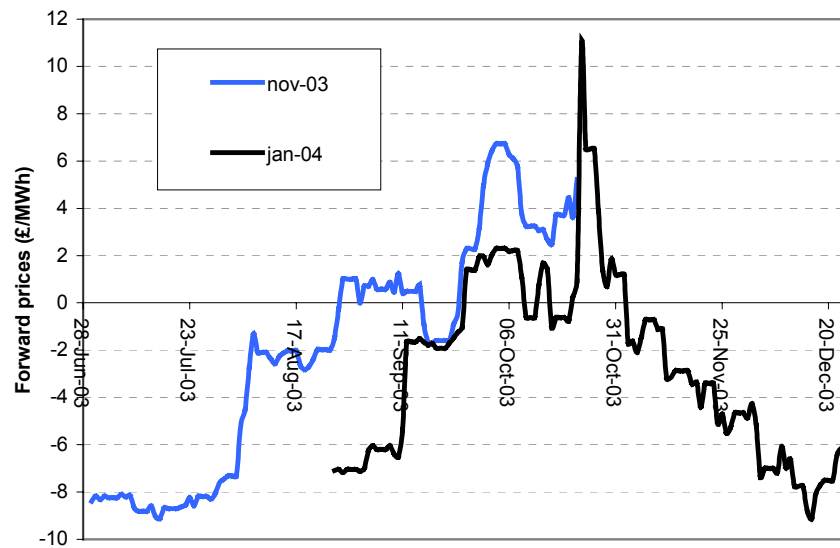


Figure 8: Forecast Winter 2003/04 Peak-load Spark Spread (data from Heren)

The role that this supplementary tender for reserve played in signalling NGT concerns and in increasing forward market prices is demonstrated by the two previous figures, which reveal that the period of NGT tender between the 14 and 27 October 2003 coincided with the forward electricity and spark-spread price spikes. In the light of this, it cannot be concluded that the winter 2003/04 was a successful market test for NETA, as NGT's intervention to procure operating reserves well ahead of real time, and in larger quantity than it usually does, was the main driver of the forward price increase.

5 NETA's flawed balancing mechanism design

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 (BM) is run as a discriminatory (pay-as-bid) auction. Elexon determines two cash-out prices: the weighted average of accepted offers (to increase output or reduce demand) determines the System Buy Price (SBP) and that of bids (to reduce output or increase demand) the System Sell Price (SSP). The imbalance pricing scheme has been subject to much criticism since the introduction of NETA (Cornwall, n.d.; Newbery and McDaniel, 2003). It is based on innovative principles, namely decentralised despatch, dual-imbalance pricing and a volume-weighted average calculation of the imbalance charge. These characteristics of the Balancing Mechanism could be changed independently of the overall NETA design concept.

5.1 The increased cost of imbalances

Balancing prices are considerably more volatile and unpredictable than the Pool prices that served as a more liquid balancing market. *Figure 9* shows 7-day moving averages of the

buy and sell prices, as well as the underlying spot price (UKPX). To give a sense of the risk in the SBP, it also shows one standard deviation (SD) of the 7-day half-hourly buy prices.

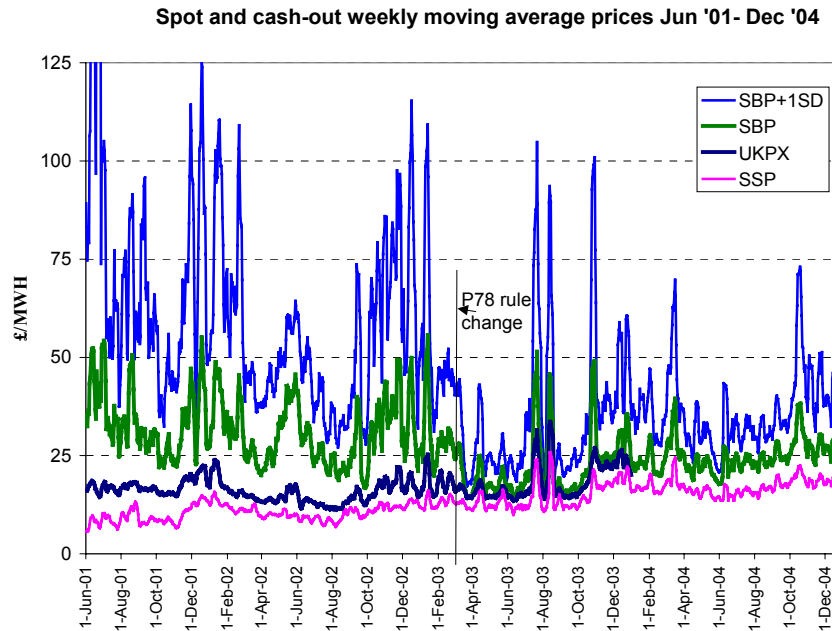


Figure 9: Spot and cash-out weekly moving average prices June 2001-April 2004 (data from UKPX and Elexon)

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 9*) made the reverse balancing price (that is, the price facing parties who are in the opposite position to the overall market, for example, long when the market was short, and hence aiding balance) revert to the spot price, in order not to penalise those helping balance the system relative to their selling in the spot market (Elexon, 2002). Its effect was quite dramatic, lowering the volatility of the balancing price. The standard deviation of the excess of the SBP over the spot price for each half-hour fell from £32/MWH in the previous 9 months to £21 in the following 9 months, and to £12 for all half-hours in 2004.

The critical feature of having two different prices ($SBP \geq SSP$) for being long or short is that it penalises each party's imbalances, whether or not they amplify or reduce the system imbalance as a whole. Combining two different imbalance prices with an average imbalance calculation makes it 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. This has led to claims that the Balancing Mechanism increases the costs of balancing to the detriment of non-portfolio generators (that is, new entrants and British Energy) and intermittent suppliers like wind (Henney, 2002).⁸

⁸ 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.

Figure 10 illustrates 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 and is unable or unwilling to re-contract before gate-closure, the cost will be the 24-hour average of the SBP from that moment.⁹ In the year before the P78 rule change, 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.

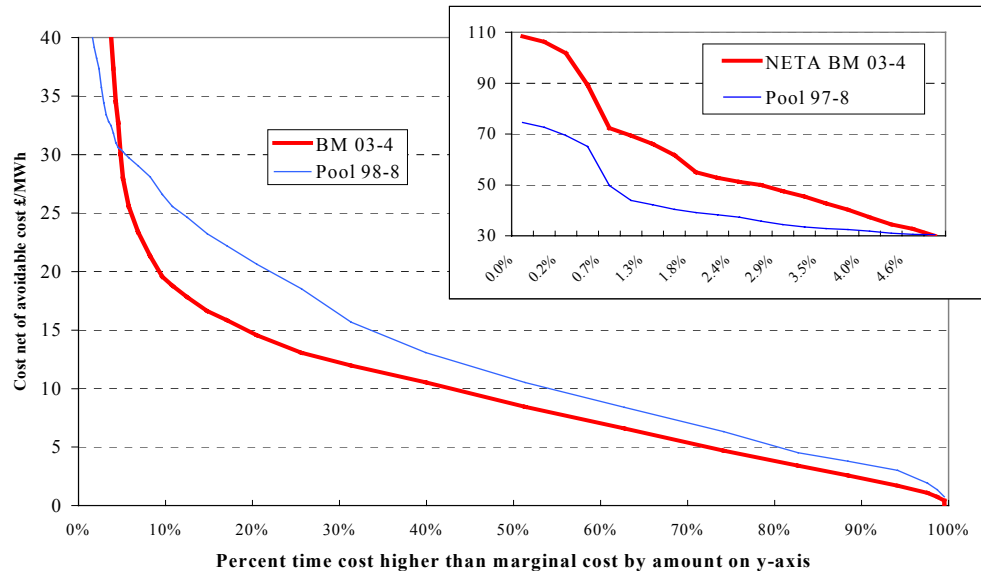


Figure 10: Cost of 24-hour failure under the Pool and NETA (Newbery, 2004)

One should however interpret these findings 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.

5.2 Decentralised despatch increases system balancing costs

Another argument that is often levelled against NETA is its inefficient decentralisation of despatch (“nominations”). While under the Pool NGT, as System operator, centrally coordinated the despatch of the generators, under NETA generators are individually responsible to self despatch. Hunt (2002) argues that the reliance on self despatch actually increases total system balancing costs, because of the loss of system multiplexing, whereby

⁹ The assumption of not being able to re-contract is conservative, tending to exaggerate the risk involved.

some imbalances cancel each other, and because the system-wide forecasting accuracy is better than the one of individual generators. Similarly, the CoAG review of different market designs concluded that NETA's incentives for individual balancing created "a significant inefficiency that adds cost to the system" (CoAG, 2002).

The fundamental issue is whether the burden of providing – and paying – for reserve capacity should be decentralised and be borne individually by the players of the market, or centrally optimised by the system operator, and paid collectively by the market players through the balancing mechanism costs. A consensual starting point is that the costs of balancing the system and providing for reserve should be minimised. However, this is not equivalent to NGT minimising the costs of the balancing mechanism, as part of the system security costs are borne by the generators through their individual hedging strategies. An optimal design of the market arrangements would thus minimise the sum of the costs of NGT balancing operations and the players' hedging costs.

The ideal balancing mechanism design lies somewhere in-between the two extreme cases of central dispatch of reserves and individual balancing. It would optimise incentives for market players to balance their positions, and the collective efficiency gains of a centralised despatch. It is thus more likely to be based on a single cash-out price to take advantage of system multiplexing, with some reasonable penalty charge for being in imbalance if the system marginal price is considered not to give sufficient incentives to contract. In addition, the system operator should be given efficient incentives to minimise aggregate balancing costs, unlike the present incentive scheme that over-encourages generators to self-balance (OFGEM, 2004c).

5.3 The lack of liquidity of the balancing mechanism

One of the arguments for implementing dual imbalance pricing was that making the spot market a poor guide to the system balancing price would encourage contracting. Liquidity would increase through bilateral trading. As expected, "over-the-counter" or OTC power trades have strongly increased under NETA (though from a very low level), as demonstrated by *figure 11* (Brown, 2003). The diversification of contract types confirms this evolution, as demonstrated by *table 1*. Several price reporters have entered the market since 2001 (that is, Platts, Heren, Energy Argus and Anderson Spectron Power Index) and the number of registered contracts types has tripled.

On the other hand, the lack of liquidity in the balancing mechanism is a real source of concern under NETA. One of the advantages of the compulsory participation into the old English Pool is that it created *de facto* liquidity. While the forward contracts under NETA appear to be quite liquid, the volume of NGT's balancing actions is comprised between 4 and 6% of the total demand out-turn. Henney (2002) argues that this lack of liquidity undermines the quality of the price signals of the balancing mechanism, which do not reflect the underlying fundamentals. Stoft (2003) points out that in a multi-market framework such as NETA, investment signals depend on the ability of the price signals to feed in to various markets without distortion. Under NETA the price relationship between the balancing and contract markets is distorted by the lack of liquidity of the balancing mechanism, and the reliance on average imbalance pricing.

Financial Year	Number of different contract types	Number of reported trades
2000/2001	137	8,351
2001/2002	322	26,538
2002/2003	336	81,794

Table 1: Diversity and volume of forward contracts (data from Heren)

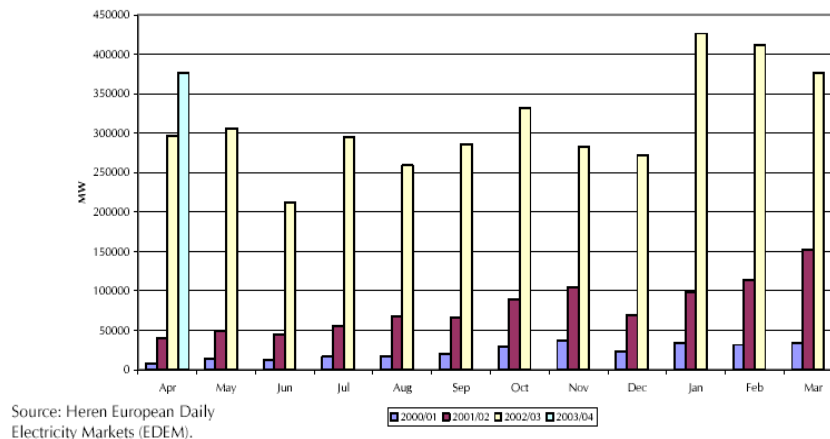


Figure 11: Evolution of reported UK OTC Power Trades Volumes (OFGEM 2003)

The events of 10 December 2002 illustrate these problems. On that day the system demand was the highest then recorded, and it exceeded the level forecast by NGT. Whilst the price in the day-ahead market showed only a slight increase over the system peak to around £30/MWh, similarly to the average System Buy Price in the Balancing Mechanism which increased only to £71.6/MWh, the System Operator accepted offers in the Balancing Mechanism at £9,999/MWh for the marginal System Buy Price (Ball, 2003). The insulation of the energy market from the costs of short-term balancing, which is both due to the average pricing formula and to the lack of liquidity of the balancing mechanism, is most significant at times of scarcity, but creates the risk that the market will fail to deliver appropriate price signals for long-term investments. A price that might warn of impending shortage may indeed not materialise until the market is under severe stress, and the delay in the price signals might undermine timely investment decisions.

5.4 Marginal pricing would provide better scarcity signals

Market fundamentals dictate that during times of shortage, electricity prices should rise to the marginal cost of generation required to meet demand. The balancing mechanism mutes scarcity signals by paying generators their bid price and not the marginal price (in order to mitigate market power and possibly reduce volatility). The current arrangements result in imbalance prices which can be significantly lower than the marginal energy balancing price, particularly at times of shortage.

Figure 12 shows a comparison of prices calculated by the current and marginal methodologies, aggregated into weekly averages across the financial year 2003/04 (NGT,

2004).¹⁰ Using a marginal pricing regime would have increased mean SBP by £2.78/MWh and decreased SSP by £1.80/MWh. The standard deviation of SBP would have increased by £14.39/MWh indicating an increase in volatility, whilst the increase in standard deviation of SSP was much smaller at £1.23/MWh.¹¹

Using average rather than marginal prices gives incorrect imbalance signals, as the imbalance charge depends on the relative contracting position of the generator – short or long – as compared to the market global imbalance. From a security of supply perspective, the crucial issue is that the signals are right at times of capacity shortage. Marginal pricing can be expected to deliver better reliability incentives in times of scarcity than average pricing as in times of scarcity the whole system can be expected to be short, and thus all generators being short in the same direction as the system would face the “real cost” of their imbalance.

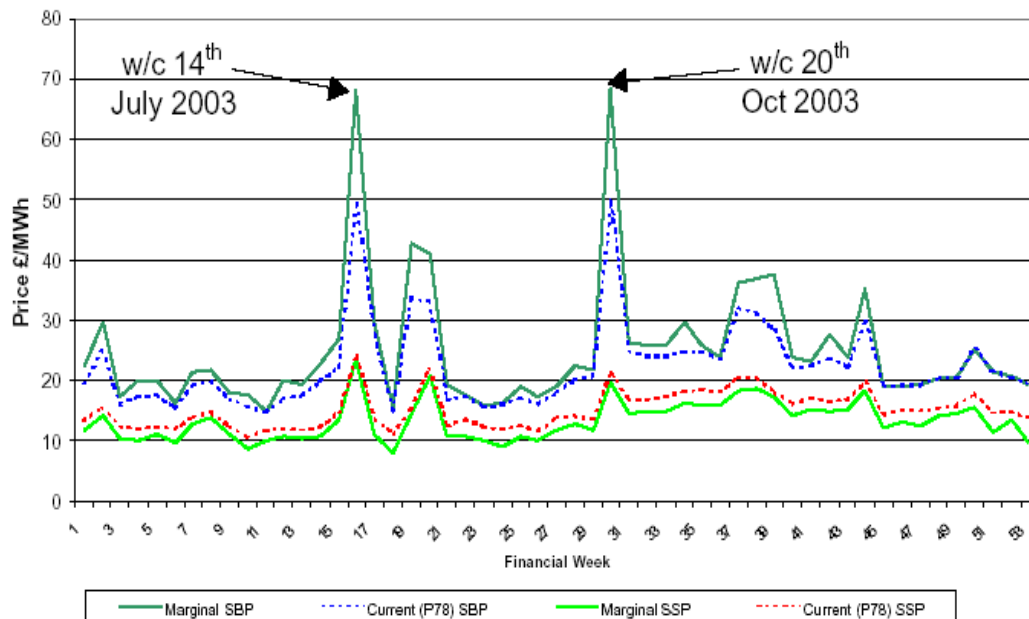


Figure 12: Comparison of Average and Marginal Pricing Methodologies 2003/04 (NGT 2004)

The main concern is that the energy price prevailing in the forward market will be artificially capped at the cost of being in imbalance, calculated as the *volume weighted average* price of providing balancing energy, rather than the price at the margin. This could lead to inappropriate contracting decisions as energy from plant with a marginal cost greater than this may not be purchased in the forward markets as it becomes economic not to buy sufficient energy in the forward market and face paying potentially lower imbalance prices. In the longer term this could lead to closure of generation with higher marginal

¹⁰ NGT analysis replicates the current tagging methodologies, but contains the following assumptions: No Arbitrage actions were taken and BPA has been added to marginal prices for the purposes of easier comparison with the current methodology.

¹¹ NGT simulations results should be interpreted with care as they rely on the strong assumption that generators and suppliers would not have changed their bidding behaviour into the market; there is a large literature demonstrating the impact of market rules on market players bidding strategies.

costs, and in the absence of a specific mechanism to ensure sufficient capacity is available to meet demand, may lead to an erosion in the levels of security of supply enjoyed to date.

NGT has proposed that imbalance prices should be calculated using a marginal methodology, but these modification proposals have been rejected by OFGEM (NGT, 2004).¹² OFGEM expressed two specific concerns regarding the proposal (OFGEM, 2004b). The first concern was that a full marginal cash-out pricing methodology would increase the incentives on participants to game the market; the second concern was that marginal cash-out prices could create distortions because they could be set based on a very small volume of energy accepted by the SO. As the market appears today both on the generation and supply-side much more competitive than it was when the review of the Pool was launched, OFGEM's concerns about market gaming could be relaxed. Besides, NGT's simulation shows that marginal prices are not likely to be set by very small trading volumes (NGT 2003). A way to address such concerns would be to use the volume weighted average of a pre-defined volume of actions (a marginal "chunk").

A more compelling argument against the change to marginal imbalance prices is that this is likely to increase volatility and the spread between the SBP and SSP to levels comparable to those experienced at the beginning of NETA, prior to the P78 modification. This would increase the over-contracting incentive for generators to reduce their imbalance exposure, and discriminate against non-portfolio generators. The increased volatility and spread between the SSB and SSP might also worsen some of the previously mentioned problems of the current balancing mechanism, namely the lack of liquidity and the despatch inefficiency of over-contracted generators.

To conclude, a marginal methodology for the calculation of imbalance prices would provide more appropriate signals to the forward markets than the current methodology, especially at times of scarcity. However, the increased volatility and spread between the SSB and SSP would worsen some other well-known problems of NETA. Therefore, the only workable solution seems to come back to marginal pricing with a *single* imbalance price. Given the change of market structure since the creation of NETA, the concerns about market-gaming that led OFGEM to advocate pay-as-bid dual imbalance pricing could be appropriately relaxed.

6 Financing issues

The long-run nature of investments in the electricity industry makes them heavily dependent on the ability to raise capital or finance. Very few power stations have been commissioned in England since 2001, or are under development. This contrasts with the large number of projects that have been given building permission, but are not at present proceeding (JESS, 2004). Raising funds appears to be the major obstacle to building new plants; bankers and lenders are extremely reluctant to engage in merchant power projects (that is, those without a guaranteed buyer of the output). De Luze (2003) explains bankers' current lack of confidence in the power sector around the world by the confluence of many seemingly unrelated events, including the California crisis, the fraud and bankruptcy of ENRON in late 2001, the questioning of the credibility of deregulation and of a power sector business model relying heavily on financial trading.

¹² The current version of the Balancing and Settlement Code (the "Code") can be found at http://www.elexon.co.uk/ta/bscrl_docs/bsc_code.html.

In the British market, additional factors have also played a role, particularly the collapse of power prices after 2000, the near total U.S. withdrawal from the market, and the financial problems of TXU Europe and British Energy. Rating agencies have lowered their credit rating of most power companies in the past years; Monnier (2003) from Fitch Rating points out that “investment-grade rating is currently extremely difficult to achieve for power projects” and that 40% of the U.K. market capacity is controlled by operators in financial distress, either insolvent or in restructuring. OFGEM does not appear to be worried by potential bankruptcies. It says that “the capital market does not appear to be a barrier to the market delivering secure supplies” (OFGEM, 2004a). Bower (2003) has an opposite position and asserts that OFGEM and the Financing Services Authority (FSA) have undermined security of supply by failing to request generation and supply companies to have sufficient capital to withstand losses arising from financial and credit risk exposures. But that should not of itself precipitate a sudden capacity shortage if a company goes into liquidation, because the receivers will continue to operate and sell power, as happened with TXU.

De Luze (2003) believes that it will take time for the confidence in the power industry to rebuild. PriceWaterHouse Coopers (2002b) recognises that power companies need to question their practices and business model profoundly to regain investors’ confidence. There is no question that there will be profitable investment opportunities in the future, but changes in corporate governance, improved risk management strategies, more integrity and transparency in financing and in financial statements will be required. A survey of the major world utilities confirms that utility companies do not seem to be on the same wavelength as investors (PriceWaterHouse Coopers, 2002a). There is indeed a huge gap between the rating of the importance of key indicators by companies and the ratings given to these indicators by investors.

6.1 The lack of forward contracting makes investment risky

Contracts hedges between generators and consumers provide a basis for hedging the risks associated with developing new power plants. However, Littlechild (2002) points out that large consumers continue to rely mainly on one to three-year contracts rather than contracts of longer term. One can expect that there will be little interest from end-consumers for entering into long-term contracts before they experience some high price spikes.

In most instances, however, generators do not sell their electricity directly to final consumers but to supply (or retail) companies which act as intermediaries. In this perspective, long-term contracting by retailers might be the missing link between the investors’ desire for a long-term contract and some consumers’ tendency to rely on short-term markets and spot prices. The possibility for long-term contracting with an electricity supplier with a large and relatively stable base of customers could be a way forward for investors searching for a long-term hedge. British supplier Centrica has in the past years entered into long-term agreements with Intergen for a UK gas-fired power plant in 2002, and with British Energy for a four-year contract in 2003 (British Energy, 2003). Somerset (2002) details the example of a new co-generation power plant in the Netherlands financed by a 15-year power purchase agreement between the project developer (Intergen) and the supplier (Nuon).

However, such long-term arrangements are to date unusual in liberalised markets.¹³ Woo et al. (2003) point out that the lack of liquidity of forward markets is a general feature of electricity markets world-wide. MacKerron and Shuttleworth (2002) suggest that removing some of the industry restrictions in Britain on the use of long-term contracts could benefit investment. Such regulatory restrictions, such as the adoption of annual tariffs for transmission access and retail sales, were indeed motivated by preventing incumbent players from tying up their customer base and using “foreclosure” to stop other companies from competing with them. The incumbents’ loss of market share makes such restrictions less necessary today, and possibly damaging to investment prospects.

6.2 The role of forward contracts

Does this lack of forward contracting more than a year ahead have an adverse impact given that construction lead times are much longer, usually between two (CCGT) and six years (Nuclear)? The economic literature looking at the factors influencing the duration of contracts and the degree of vertical integration emphasises the importance of hold-up and related opportunism considerations (Joskow, 1987).

De Vries and Hakvoort (2002) argue that the lack of long-term contracts represents an important market failure, increasing investment risk. However, in other commodity industries with a comparable lack of liquid longer term forward markets, such as oil for instance, investments are made on the basis of price expectations and not thanks to binding forward contracts. Investors and traders use forward curves representing expectations of the near-term future prices to guide their short- to medium-term decisions, but rely on market fundamentals to predict longer-run future prices and returns. Thus, the lack of forward contracting as such should not represent an important impediment for long-term generation investment, provided that investors develop an appropriate understanding of the risks and returns at stake. Indeed, a survey of the major world utilities reveals that although utilities recognise that market short-termism is a problem, the majority of the utilities surveyed claim that it does not inhibit investment (Price WaterHouseCoopers, 2002a).

But if investors ground their long-term prices estimates in market fundamentals, one might wonder to what extent investment decisions are taken independently of short-term price signals. An extreme case would see investors rely on their own forecasts and business strategies for long-term investments, while market signals would only be taken into account for their hedging and operational strategies. In this perspective, plants demothballing (as for instance in Britain during the Autumn 2003) would be a short-term response to increasing forward electricity prices, but new investment decisions would rely on much longer-term price and demand estimates. Such disconnection between short-term market signals and long-term investment decisions could lead to *herd behaviour* and hysteresis cycles in electricity market investments, as modelled by Ford (1999). Ford’s model suggests that electricity markets are likely to be characterised by boom and bust construction cycles, as observed in other capital-intensive industries with long building times and inelastic demand. On the other hand, considerable investment cycles were also a feature of the post oil-shock world under the old vertically integrated model, and may be inevitable given that the future is inherently uncertain.

¹³ In many developing and transition countries, private generator investment invariably requires a long-term power purchase agreement, often with sovereign risk guarantees.

6.3 Does retail competition hamper long-term contracting?

Full liberalisation includes retail liberalisation that may affect the incentives for retailers to enter into long-term power purchase agreements with generators. It can be expected that retailers will be less interested in signing long-term power purchase agreements if consumers can switch suppliers. Neuhoﬀ and de Vries (2004) argue that retail liberalisation undermines long-term contracting between generators and retailers, because of the free riding opportunities it creates. Retail companies may lose their customers to new retail companies in times when their long-term contract price exceeds the spot price. In periods with average wholesale prices and retail prices above long-term contract prices, retail companies benefit and generators lose from their long-term contracts. In exchange, generators would expect to win from long-term contracts in periods with low wholesale prices. In such periods, new retail companies may enter the market and offer cheap retail electricity. Thus retail companies with existing long-term contracts would incur losses, and some might go bankrupt and not honour their contracts. Green (2003) estimates the size of this effect, combining models of electricity retailing and of competition in the wholesale market. Given enough volatility and an otherwise competitive long-term market, he finds that retail competition might raise wholesale prices by up to nineteen per cent.

Littlechild (2002) argues against this belief that retail competition precludes the signing of long-term contracts. He notices that “if the contract is really worth signing, the utility could match any price reductions to customers and still come out ahead. A consequence of retail competition is that suppliers who wish to sign long-term contracts have to back their own judgement rather than pass the risks to customers; this is likely to improve the quality of decision-making”. Perhaps more to the point, banks and building societies make fixed-interest mortgages available while short-run rates vary, exposing them to similar risks which are typically handled by demanding a pre-agreed exit or renegotiation fee for ending a contract that is out-of-the-money for the borrower. In other markets there are penalties for early ending a contract that is out of the money, and that would be the logical solution to this perceived problem.

Neuhoﬀ and de Vries (2004) suggest that one way to solve the problem would be to depart from retail competition and create a credible counterparty with which generators could sign long-term contracts (see also Joskow and Tirole, 2004a). If, for example, retail companies held regional monopolies, consumers would not have the option to switch. The most direct way to maintain generation adequacy would therefore be to retain the consumer franchise (Newbery, 2002). Littlechild (2002) notices however that a disadvantage of retail monopoly is that utilities and regulators, who do not have to test their judgements in the market, are typically not well placed to judge the costs and risks of long-term contracts. They can nevertheless force customers to enter such contracts and to bear the resulting costs and risks. Some of this can be addressed by yardstick regulation of the contract costs that can be passed through to captive consumers. Most importantly, generators and suppliers have alternatives to long-term contracts such as vertical integration to provide them with greater revenues security.

6.4 Vertical and horizontal re-integration

Vertical integration between suppliers and generators constitutes an organisational hedge improving the prospects of stable cash flows and finance for large, capital intensive investments. *Figure 13* shows the consolidation of British supply market since 1995; five generation companies now provide power to most of the domestic UK consumers.

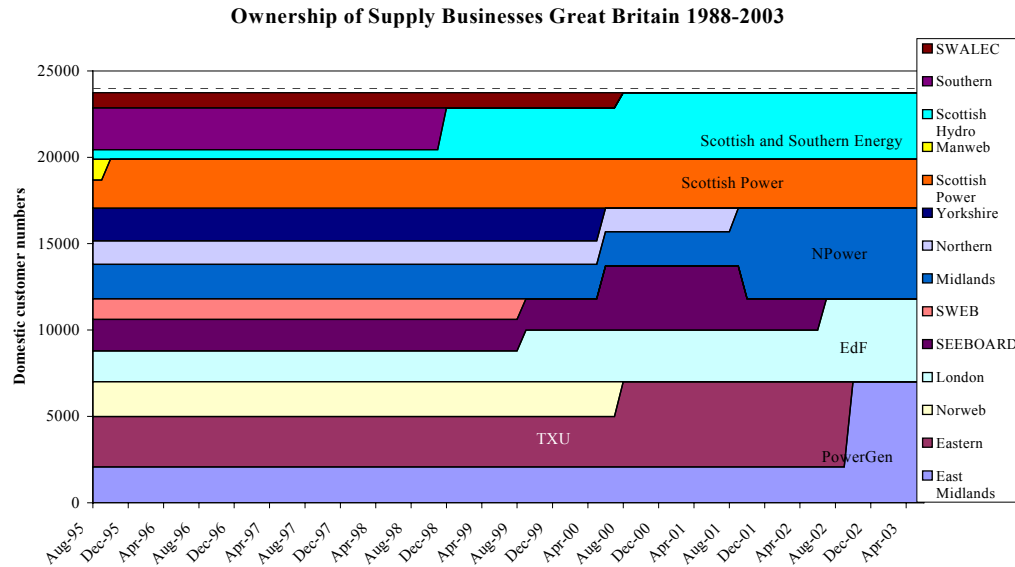


Figure 13: Ownership of supply businesses in the UK (data from J Bower)

Horizontal integration between gas and electricity businesses is another hedging strategy for companies making a significant investment in gas-fired generation. The risk of volatile gas prices can be hedged by acquiring companies with upstream natural gas assets. The last decade has seen several gas-electricity cross-sectoral mergers and acquisitions in the USA and in Europe. Meritet (2001) analyses the convergence of natural gas and electric power industries in the United States, and the consequences of mergers and acquisitions between electric power firms and natural gas (principally gas distributors). Gas-electricity convergence mergers may be attractive to investors. The ability to generate electricity from gas, or to stop generating and sell gas, if that is more profitable, helps the merged company to manage the price volatility in natural gas and electricity in the absence of long-term contracts or liquid forward markets. Besides, higher equity financing is a way to cover more risky investments, but comes, however, at the price of higher financing costs. Indeed, the growth in power firms through mergers is, for this reason, not surprising.

The recent mergers and acquisitions in the European electricity industry have raised concerns about concentration of the industry. Vertical re-integration or horizontal concentration can create market power that can be abused to reduce competition. Mergers between electricity and gas companies are under close scrutiny of the competition regulatory authorities. Convergence mergers of gas suppliers with electric utilities raise competitive concerns if they result in market power over the supply of fuel to a supplier. The prospect of utility assets being concentrated into the hands of a select group of companies may undermine EU plans for an open European energy market. European and national Competition authorities have intervened in several instances to prevent mergers and acquisitions that might lead to excessive market power. A better strategy would be to allow companies in one sector (gas or electricity) to move into the other sector through investment rather than acquisition, thus increasing the number of competitors in each market.

For regulators, there are important trade-offs to be made as regards horizontal and vertical integration and investment incentives. On the one hand, financing long-lived

investments such as power plants in a liberalised industry may require companies to be above some critical size. On the other hand, regulators must be careful that these companies do not acquire too much market power. These conflicting objectives require co-ordination between the electricity market regulators and the competition authorities and a fuller understanding of the nature of market power in electricity markets.

7 Market design changes to improve market signals

Many of the problems of electricity markets stem from the low elasticity of demand and the inability to transmit short-run scarcity price signals to the large number of consumers without real time meters. In Britain about half of total consumption is half-hourly metered and receives price information a day ahead. Consumers can also bid into the balancing mechanism. In other countries the fraction of consumption that can respond to prices may be considerably lower, and the problems of market responsiveness more acute. Even in Britain, balancing supply and demand relies mainly on supply side flexibility. Power cuts are very infrequent and most customers cannot signal the value they place on reliability. This creates a bias in the perception of the commercial value of electricity. Electricity provision and reliability are today perceived as a right, which is reasonable as the value of providing the reliability is clearly high relative to current prices, but there is a lack of connection between this value and the price paid.

If access to electricity supply were to be regarded as a right, security and quality of electricity supply need not be. Rochlin (2002) points out that as long as security of supply continues to be treated as a right, “customers and politicians should know it involves significant stranded costs in both generation and transmission”. Patrick and Wolak (1997) showed in their econometric study of the inter-temporal substitution of electricity consumption that even a small increase in demand elasticity can have a big impact on power prices. A stronger demand response would help mitigate market power in electricity markets and provide potential investors with more predictable energy (and ancillary service) prices, and therefore decrease investment risks.

Providing each individual consumer with half-hourly meters does not currently seem cost effective, although ENEL is investing heavily in south Italy to install individual consumer meters, arguing that the poor payment rates of the area make it worth the investment. Borenstein (2004) concludes from a simulation exercise that the efficiency gains of real time pricing (RTP) for the largest customers are likely to far outweigh the costs, but that the incremental benefits of putting more customers on RTP are likely to decline as the share of demand on RTP grows, so that the gains from putting smaller customers on RTP might not justify the costs. However, it may be cost-effective to replace existing meters with RTP meters whenever they need replacement for ageing reasons, and then to offer such customers additional services that such meters make possible.

Even if the extreme solution of RTP meters proves too expensive, a mix of technological improvements and financial contracts for delivery insurance can potentially greatly improve the demand responsiveness. Doorman (2003) describes a capacity subscription scheme, based on the theoretical concepts of priority service and self-rationing, whereby demand is limited to a pre-subscribed capacity through a Load Limiting Fuse activated in times of scarcity by the System Operator. Such technological device would be much less expensive to install than individual meters. Another improvement

would be the development by suppliers of standardised insurance contracts which would compensate for loss of load on the basis of the reliability level chosen and paid for by the customer. These contracts would simply be a systematisation of the current case by case compensation of power cuts (NGT, 2001).¹⁴ Their time-varying and consumer-type depending attributes provide a basis on which to elaborate. Insurance contracts for loss of load would allow customers to choose a cheap or more expensive insurance contract on the basis of the reliability level s/he would like to be compensated for in case a shortage occurs. Such insurance contracts need no technical change and could provide valuable signals on consumers' valuation of security of supply.

7.1 Reducing regulatory uncertainty

Regulatory uncertainty is one of the major sources of risk for investors in a fast transforming regulatory framework (de Vries and Hakvoort, 2002). One straightforward way to improve investment incentives is for the government to adopt sustainable market rules, and carefully review the necessity of any change. There are two main concerns that may trigger government intervention in the UK electricity market. First, the Government's commitment to promote renewables and low-carbon generation options introduces considerable uncertainty. Uncertain carbon prices and changing government stance towards renewables lead investors to prefer cheaper and quicker-build plant such as CCGT rather than capital-intensive nuclear plant. The second source of uncertainty is related to security of supply. It is far from certain that any Government would be able to resist political pressures to intervene during periods of high prices and enforced supply reductions. In Britain, many parliamentarians question the Government's reluctance to intervene in energy markets. Since electricity is so vital, the Government should not restrict its role to just monitoring the electricity market (Trade and Industry Committee, 2002).

There is a delicate trade-off to be found between governance adaptability and flexibility in a relatively new electricity markets and investors' need for predictability. As the market matures, adaptability and reactivity of the regulatory framework is critical for maintaining security of supply. The lack of flexibility of the governance arrangements was one of the criticisms levied at the Pool. NETA's governance structure seems much more flexible: during the first two years of NETA, 146 modifications were made to the Balancing and Settlement Code, and 64 to the Connection and Use of System Code (Brown, 2003).

Parliamentary sovereignty in Britain prevents any government from limiting the decision-making powers of future governments. Thus the required assurance must emerge from the accumulation of accepted customs and practices, backed up by law or statutory instrument where possible. Licences are the standard form of providing such reassurance, as these are commercially enforceable contracts that can only be changed by consent or due process, and go a considerable way towards providing regulatory stability. To further reduce regulatory uncertainty, MacKerron and Shuttleworth (2002) propose that British Governments recognise and apply at all time the revenue standard, which would prevent or compensate any shortfall of cost recovery resulting from the introduction of a new

¹⁴ Under the Guaranteed Standard on Supply Restoration, domestic customers are entitled to £50 compensation (business customers £100) if their electricity is not restored within 18 hours. Customers are entitled to an additional payment of £25 for each subsequent period of 12 hours that supply is not restored.

regulation. They suggest that some legally binding clarification of regulators' statutory duties with regard to the revenue standard might help to reinforce investment incentives.

7.2 Re-introducing a capacity mechanism?

Should a capacity payment or obligation be reintroduced in Britain in order to maintain security of supply? Would this be more efficient or cost effective than relying on market signals supplemented by (a better incentivised) NGT reserve procurement? The most compelling argument in favour of a "well-designed" capacity mechanism coupled with a price cap is that this reduces consumer's exposure to price volatility, while providing steadier revenues for generators.

This should be determined on the costs and benefits of such a mechanism (taking due account of the intrinsic benefit of maintaining the status quo to reduce regulatory risk). The UK Government commissioned a study by NERA in 2002 to review the case for a capacity payment to be reintroduced (NERA, 2002). NERA estimated that introducing a capacity payment instrument could increase costs to consumers by some £151 million per year. On the basis of this study, the 2003 Energy White Paper concluded that "the case has not been made for such an instrument in the UK" (DTI, 2003, page 86 at 6.43). NERA's calculation appears, however, to be highly debatable and seem to greatly overestimate the cost of such a capacity instrument. The calculation was over-simplified in that it just estimated the increase in customers' electricity bills if the system reserve was to be maintained artificially at higher levels than the assumed market reserve level. This did not take into account the benefits to customers of maintaining a higher capacity margin, namely a lower LOLP, that is, less frequent power shortages. A robust cost-benefit analysis would have included the value consumers placed on better reliability. Besides, the assumptions on which NERA based its calculation greatly over-estimated the reserve difference in the two scenarios. The assumed 8% market-determined capacity margin seems very low as compared to the capacity margins observed in the market, whose lowest point was expected to be 16.2 % in 2003. A capacity payment instrument would therefore probably increase costs to consumers by much less than the £151 million NERA estimates.

Ultimately, the choice between a well working energy-only system and a market supplemented by capacity payments or obligations is a policy choice as regard to the level of volatility that policy makers are willing to see in the market. One could indeed consider the introduction of capacity payments as a proactive measure in the form of a mandatory hedge or insurance that will assure that prices stay within a socially acceptable range, which indeed bears a cost. Besser et. al. (2002) make such an argument to defend the requirement of a capacity obligation in FERC's SMD, by saying that the volatility of energy-only markets has a socially and politically unacceptable cost. Thus, while electricity markets may be delivering adequate levels of investment, price spikes are testing government commitment to allow markets to sort things out.

In Britain, however, there are two factors that should help give the Government greater confidence to rely on market mechanisms to resolve any capacity crisis. First, there has been a long history of open electricity markets, which have led to a better utilisation of generating capacity and lower electricity prices over several years. The fact that customers have already enjoyed several years of falling prices should increase confidence that liberalised electricity markets can create consumer benefits. Second, both the industry and the regulator have gained considerable experience with electricity markets. The maturity of the market players, combined with relative flexibility of the NETA markets arrangements, which have proven much easier to amend than the Pool, should survive sustained price

spikes periods without the need for regulatory intervention. Norway already experienced an extreme example of this in the winter of 2002/3, when consumer prices trebled for a sustained period without requiring regulatory intervention (von der Fehr et al, 2004).

If the Government or regulator is forced to intervene to supplement market signals, the choice should not be limited just to capacity payments or requirements, or centralised reserve procurement by NGT, as was the case during the winter 2003/04. Capacity options (see Arriaga et al. (2002), Chao and Wilson (2003), and de Vries (2004)) and capacity subscriptions (Doorman, 2000) are two promising alternatives to capacity mechanisms that await a market test. A capacity option is a one-sided Contract-for-Differences (CfD) that limits the price paid by the holder to the strike price. If the strike price is set at a suitably high level, the price of the option will be modest, and this certain payment reduces the risk to the issuer, and can be used to pay for the reserve capacity which would otherwise wait for the price spikes for its income. It might be a requirement that any supplier offering fixed price contracts either hold such CfDs, or provide sufficient collateral to ensure that they do not fail when spot prices are high. The spot market could then be left to signal true scarcity without the fear of political fall-out. Such insurance instruments are likely to provide better investment incentives and price volatility mitigation at lower cost than capacity payments or requirements and price caps.

8 Conclusions

While there is compelling evidence to suggest that the electricity industry liberalisation has, in general, improved productive efficiency, and reduced operating and maintenance costs (Newbery and Pollitt, 1997; Wolfram, 2004), there is as yet insufficient experience to assess the long-term benefits from liberalising the electricity industry. One key issue lies in the ability of a liberalised electricity industry to deliver appropriate investment incentives and maintain generation adequacy in the long-term. As the first countries to liberalise – among which are included Britain and some U.S. states – have now reached the end of their first investment cycle, much attention is being paid to assessing the long-term dynamic performance of the liberalised electricity industry.

There is as yet no clear academic consensus on which market design provides the least distorting long-term investment incentives. In liberalised markets investment must be profit motivated, and under current EU Directives capacity choices are left to the market, except if there is a potential threat of shortage (at which point tenders can be accepted for new plant). The reliability of electricity supply has been the principal motivation for many technical and economic regulations imposed on market designs by regulators. The main issue is whether a separate payment for reserve capacity is required on top of energy sales. Energy-only market designs can in theory deliver appropriate investment incentives, but they rely on high prices during periods of capacity scarcity to remunerate peaking units (although these can be hedged by contracts). This has led some regulators to implement price caps coupled with capacity mechanisms to protect consumers from extreme price volatility, and provide investors with more stable revenues.

Britain has successively experimented with two extreme market designs, the Pool which included a capacity payment, and the New Electricity and Trading Arrangements (NETA), which is a decentralised pay-as-bid energy-only market. The Pool design had many advantages as regard to investment incentives: its compulsory nature created

liquidity and a reference marginal system price on which to devise hedging contracts. In addition, capacity payments provided an additional source of revenue that facilitated entry by Independent Power Producers. However, excessive market concentration lent itself to gaming by generators, and the inflexible governance structure made incremental reforms difficult. It was abolished just as the market was becoming workably competitive, after excess entry and industrial restructuring traded horizontal for vertical integration (Newbery, 2004).

Three years after NETA, there is current concern that the investment drought might endanger generation adequacy. As the end of the Pool coincided with a massive decrease in generation concentration, before and after comparisons risk conflating changes in market design with changes in market structure. NETA went through its first stress-test during 2003 without power cuts thanks to the massive procurement by NGT of reserve capacity ahead of real time, which triggered a substantial increase in forward prices and the reconnection of previously mothballed plants.

The main issue lies in NETA's balancing mechanism. Its dual imbalance pricing methodology uses an averaging calculation that mutes scarcity signals. If NETA is to deliver appropriate investment incentives, then the balancing mechanism should be revised. Calculating two marginal imbalance prices does not constitute a satisfactory solution, as it would raise the spread between the two prices and thus would worsen the current tendency of market players to over-contract. The only workable solution seems to be for NETA to return to a single marginal imbalance price.

Turning to the current hedging and financing difficulties of power projects, the absence of liquid forward markets and corresponding contracts of more than a few years duration makes merchant investment in electricity more risky. Ending the domestic franchise in 1999 removed the ability of the RECs to write long-term contracts with generators that could be passed through to the price-regulated consumers. Financing capital intensive investments will therefore require alternative underwriting. Vertical integration provides many of the equivalents of long-term contracting, particularly, as seems the case, if domestic customers are "sticky".

The trend to vertical integration in most markets, and horizontal integration on the Continent, suggests that larger companies are actively reducing risk exposures. Horizontal integration may confer sufficient market power to increase the price-cost margin to the point that investment becomes attractive (to incumbents and entrants). Size itself and the ability to operate in different markets also reduces the risk of individual investments. The regulator and competition authority should however carefully monitor what could be defined as a *competition versus generation adequacy* trade-off: a critical size is required for generators to be able to withstand the risks of investment in a liberalised market, but horizontal reintegration raises market power concerns.

How could investment incentives under NETA be improved? The priority should be to reform the balancing market, further encourage demand responsiveness and reduce regulatory uncertainty. Is a capacity mechanism necessary to maintain generation adequacy? If a reform of the Balancing Mechanism to limit price signal distortions is not undertaken, and if the market is to remain unconcentrated, as seems desirable, such a mechanism to support investment may indeed be required to maintain generation adequacy. Re-introducing a capacity payment could lower market volatility and improve investment incentives, but if poorly designed could entail substantial costs and negative side-effects. Ultimately, the choice between a well-working energy-only system and a

market supplemented by capacity payments or obligations (including capacity options) is a policy choice regarding the level of volatility that policy makers are prepared to see in the market.

If the Government or regulator is to make the choice to supplement market signals with some form of regulatory intervention, then it should not restrict its choice to existing mechanisms such as capacity payments or requirements, or centralised reserve procurement. Capacity options appear more promising (if still untested) solutions that are likely to provide better investment incentives and risk reduction at a lower cost.

8 References

Arriaga, I., C. Vasquez, and M. Rivier (2002) "A Market Approach to Long Term Security of Supply", *IEEE Transactions on Power Systems*, 17: 349-357.

Ball, R. (2003) "A Day in the Life ... – An Operator's View of Tuesday 10th December 2002", National Grid Company.

Besser, J., J. Farr, and S. Tierney (2002) "The Political Economy of Long-Term Generation Adequacy: Why an ICAP Mechanism is Needed as Part of Standard Market Design", *The Electricity Journal*, 15: 53-62.

Boiteux, M. (1949) "La tarification des demandes en pointe: Application de la théorie de la vente au coût marginal," *Revue générale de l'électricité*.

Borenstein S. (2004) "The Long-Run Effects of Real-Time Electricity Pricing," CSEM Working Paper 133. University of California Energy Institute.

Bower, J. (2003) *Why OFGEM?* Oxford Institute for Energy Studies. www.oxfordenergy.org/pdfs/Sep03_Why%20Ofgem.pdf.

British Energy (2003) "British Energy Signs Major Supply Deal with Centrica," *British Energy Press Release*, 7 February.

Brown, S. (2003) "England & Wales Wholesale Market – 2 Years On", www.ofgem.gov.uk/temp/ofgem/cache/cmsattach/3654_neta_sonia.pdf.

Caramanis, M., R. Bohn, and F. Schweppe (1982) "Optimal Spot Pricing: Practice and Theory," *IEEE, Transactions on Power Apparatus and Systems*, PAS-101: 3234-3245.

Caramanis, M.C. (1982) "Investment Decisions and Long-term Planning under Electricity Spot Pricing," *IEEE, Transactions on Power Apparatus and Systems* 101: 4640-4648.

Chao, H. and R. Wilson (2003) "Resource Adequacy and Market Power Mitigation via Option Contracts," EPRI Working Paper.

Chao, H. (1983) "Peak Load Pricing and Capacity Planning with Demand and Supply Uncertainty," *Bell Journal of Economics*, 14: 179-190.

Chuang A. and F. Wu (2000) "Capacity Payments and Pricing of Reliability in Competitive Generation Markets," *Proceedings of the 33rd Hawaii International System Science Conference*, January.

CoAG (2002) Council of Australian Governments, "Towards a Truly National and Efficient Energy Market," (The Parer Report), Canberra, 20 December. <http://www.industry.gov.au/assets/documents/itrinternet/FinalReport20December200220040213110039.pdf?CFID=242389&CFTOKEN=11377123>.

Cornwall, N. (n.d.) "The Need for Balance in NETA Imbalance Pricing," www.ksg.harvard.edu/hepg/Papers/Cornwall%20NETA%203-02%20DWAP%20view.pdf.

Department of Trade and Industry (2003) "Our Energy Future: Creating a Low Carbon Economy," Energy White Paper, February.

Doorman G. (2003) "Capacity Subscription and Security of Supply in Deregulated Electricity Markets," *SINTEF Energy Research*, August.

Ellexon (2002) "Modification Report, Modification Proposal P78" – Revised Definition of System Buy Price and System Sell Price, prepared by ELEXON on behalf of the Balancing and Settlement Code Panel, 16 August, Document Reference P078RR. www.ellexon.co.uk/.

Ford, A. (1999) "Cycles in Competitive Electricity Markets: A Simulation Study of the Western United States," *Energy Policy*, 27: 637-658.

Fraser, H. (2003) "A Critique of the Resource Adequacy Requirement Proposed in FERC's SMD NOPR", *The Electricity Journal*, 16: 23-28.

Fraser, P. (2003) "Power Generation Investment in Electricity Markets," OECD/IEA Publication.

Graves F., P. Hanser, and R. Earl (1998) "The Death and Resurrection of Electric Capacity Pricing," Working Paper. The Brattle Group, October.

Green, R. (2003) "Retail Competition and Electricity Contracts," CMI Electricity Project Working Paper 33, December.

Green, R. (2004) "Did English Generators Play Cournot? Capacity Withholding in the Electricity Pool," CMI Working Paper 41, March.

Grimston M. (2004) "Generating Profits? Experience of Liberalised Electricity Markets," NICAB Working Paper.

Henney, A. (2002) "An Independent Review of NETA," EEE Ltd. Publication, 21 November.

Hirst, E. and S. Hadley, (1999) "Generation Adequacy: Who Decides?" *The Electricity Journal*, 12: 11-21.

- Hunt, S. (2002) *Making Competition Work in Electricity*. New York: Wiley.
- International Energy Agency (2002) “Security of Supply in Electricity Markets, Evidence and Policy Issues,” OECD/IEA Publication.
- JESS (2004) Joint Energy Security of Supply Working Group DTI/OFGEM, Fifth Report, November.
- Joskow, P. (1987) “Contract Duration and Relation-specific Investments,” *The American Economic Review*, 77: 168-185.
- Joskow, P. and J. Tirole (2004a) “Retail Electricity Competition”, CMI Electricity Project Working Paper 44, April.
- Joskow, P. and J. Tirole (2004b) “Reliability and Competitive Electricity Markets,” CMI Electricity Project Working Paper 53, May.
- Littlechild, S. (2002) “Competition in Retail Electricity Supply,” CMI Electricity Project Working Paper 9.
- de Luze, G. (2003) “Investment in Power Generation: A Banker’s Perspective,” 25th March 2003 IEA/NEA Conference, Paris.
- MacKerron, G. and G. Shuttleworth (2002) “Guidance and Commitment: Persuading the Private Sector to Meet the Aims of Energy Policy,” NERA report for Powergen.
- Markiewicz, K., and C. Wolfram (2004) “Has Restructuring Improved Operating Efficiency at US Electricity Generating Plants?” Center for the Study of Energy Markets Working Paper 135. Berkeley.
- Meritet S. (2001) “The Convergence Of Electric Power And Natural Gas Industries in the United States and In Europe,” CGEMP Research Paper. Université Paris Dauphine.
- Monnier, L. (2003) “The Battle for Drax: Are UK Generators Back in Power?” *Fitch Rating Special Reports*, October. www.fitchratings.com.
- National Grid Transco (2004) *Response to OFGEM Consultation, Electricity and Gas Cashout Review*, Executive Summary. www.ofgem.gov.uk/.
- National Grid Transco (2003) “Proposed Revisions to the BSAD Methodology Statement required to facilitate Marginal Imbalance Pricing as proposed by BSC Modification Proposals P136 and P137,” 25 November. www.nationalgrid.com/uk/indinfo/balancing/mn_archive_consultations.html#3.
- National Grid Transco (2003) Seven Year Statement. <http://www.nationalgrid.com/uk/library>.
- Neuhoff K. and L. de Vries (2004) “Insufficient Incentives for Investment in Generation,” CMI Working Paper 42, March.

- NERA (2002) "Electricity Markets and Capacity Obligations," A Report for the DTI, 13 December. www.dti.gov.uk/energy/whitepaper/cap_study.pdf.
- Newbery, D. (1995) "Power Markets and Market Power," *Energy Journal*, 16: 41-66.
- Newbery, D. (1998a) "The Regulator's Review of the English Electricity Pool," *Utilities Policy*, 7: 129-41.
- Newbery, D. (1998b) "Pool Reform and Competition in Electricity," in M. Beesley (ed.) *Regulating Utilities: Understanding the Issues*. London: Institute of Economic Affairs.
- Newbery, D. and M. Pollitt (1997) "The Restructuring and Privatisation of the CEGB: Was it Worth it," *Journal of Industrial Economics*, XLV: 269-303.
- Newbery, D. and T. McDaniel (2003) "Auctions and Trading In Energy Markets: An Economic Analysis," in P. Vass (ed.) *CRI Regulatory Review 2002/ 2003*. Bath: Centre for the Study of Regulated Industries.
- Newbery, D. (2004) "Electricity Liberalisation in Britain: The Quest for a Satisfactory Wholesale Market Design," *Energy Journal*, forthcoming.
- OFGEM (2004a) "Security of Supply April to September 2003, Six Month Retrospective Report," February.
- OFGEM (2004b) "Decision Letter in Relation to BSC Modification Proposals P136 and P137," 30 March.
- OFGEM (2004c) "NGT System Operator Incentive Scheme from April 2004, Proposals and Statutory Licence Consultation Office of Gas and Electricity Markets," 3 February.
- Oren S. (2003) "Ensuring Generation Adequacy in Competitive Electricity Markets," Energy Policy and Economics Working Paper Series. University of California Energy Institute's (UCEI).
- Patrick, R. and F. Wolak (1997) "Estimating the Customer-Level Demand for Electricity Under Real-Time Market Prices," mimeo, August.
- PriceWaterHouseCoopers (2002a) "Competing for Capital: Recognising Utilities' Worth", 2002. www.pwcglobal.com/energy/trading.
- PriceWaterHouseCoopers (2002b) "Energy Trading: Re-establishing Sound Foundations," 2002. www.pwcglobal.com/energy/trading.
- Rochlin C. (2002) "Is Electricity a Right?" *The Electricity Journal*, 15: 31-36.
- Somerset, M. (2002) "Lessons Learned from an IPP," *Converging Gas and Power Markets*, December, Amsterdam.
- Stoft S. (2002) "Power Systems Economics – Designing Markets for Electricity," IEEE Press.

Stoft S. (2003) "The Demand for Operating Reserves: Key to Price Spikes and Investment," *IEEE Transactions on Power Systems* 18: 470-478.

Strbac, G. and D. Kirschen (2000) "Who Should Pay for Reserve?" *The Electricity Journal*, 13: 32-37.

Sweeting A. (2001) "The Effect of Falling Market Concentration on Prices, Generator Behaviour and Productive Efficiency in the England and Wales Electricity Market," Working Paper. May, MIT.

Sweeting, A. (2004) "Market Power in the England and Wales Wholesale Electricity Market 1995-2000," CMI Electricity Project Working Paper 55, August.

Trade and Industry Committee (2002) "Energy Supply: How Secure are We?" *House of Commons*, Fourteenth Report, 12 February.

Von Der Fehr, N-H., M. Amundsen and L. Bergman (2004) "The Nordic Market: Signs Of Stress?" *Energy Journal*, forthcoming.

de Vries, L. and R. Hakvoort (2002) "The Question of Generation Adequacy in Liberalised Electricity Markets," Working paper. Delft University.

de Vries (2004) "Securing the Public Interest in Electricity Generation Markets: The Myths of the Invisible Hand and the Copper Plate," PhD thesis, TU Delft, Netherlands.

Weitzman, M. (1974) "Prices vs. Quantities," *Review of Economic Studies*, 41: 477-91.

Wolfram, C. (1999) "Measuring Duopoly Power in the British Spot Market," *The American Economic Review*, 89: 805-826.

Wolfram, C. (2004) "The Efficiency of Electricity Generation in the US after Restructuring," in J. Griffen and S. Puller (eds.) *Electricity Deregulation: Where to from Here?* Chicago: University of Chicago Press.

Woo, C., D. Lloyd, and A. Tishler (2003) "Electricity Market Reform Failures: UK, Norway, Alberta and California," *Energy Policy*, 31: 1103-1115.