Markets for Electricity in Europe

Richard Green*

University of Hull and CEPR

Department of Economics
University of Hull
Hull HU6 7RX

Tel: 01482 465720
Fax: 01482 466216
Email: R.J.Green@hull.ac.uk

Draft for *Oxford Review of Economic Policy*, Autumn 2001
September 2001

Competition in the electricity industry requires properly functioning wholesale markets. This paper discusses the better-established electricity markets of Europe, in the Nordic countries, England and Wales, and Spain. The different market designs can all produce efficient outcomes, if not hindered by market power. Long-term contracts provide some defence against market power, and give price stability to promote investment. Competition in electricity retailing may discourage retailers from signing long-term contracts in future.

I. INTRODUCTION

The countries of the European Union are liberalising their electricity industries. The 1996 Directive on the internal market in electricity required member states to allow new generators access to the transmission system, and to allow their largest electricity consumers a choice of electricity retailer. The Directive's Single Buyer Model might have created a single wholesaler to buy power from generators, but in practice, this model has been

* I would like to thank John Flemming, Claude Crampes, Dieter Helm, Jean-Jacques-Laffont, Pedro Marin, and participants at the Oxford Review of Economic Policy seminar for helpful comments, and Ignacio Pérez-Arriaga for sharing his knowledge of the Spanish electricity market. Any mistakes are mine.

discarded in favour of providing direct Third Party Access to the transmission system. Regulation or negotiation sets the terms of access, allowing generators to trade with retailers and customers. The nature of electricity, however, requires specialised markets to allow such trading to take place.

Electricity is a difficult commodity to deliver. Every power station connected to a given AC network must be synchronised, running at exactly the same speed. Electricity must be generated at the instant it is consumed, and so some capacity must always be held in reserve, able to increase output instantaneously in response to an increase in demand or the failure of another plant. Other capacity must be kept on stand-by, ready to replace the first tier of reserves within a few minutes of being called into action.

Power flows from generators to consumers cannot be directed, but will be distributed along every line in the network, according to physical laws. If too much power attempts to pass along a given line, or through a particular transformer, that component of the network will fail. Following a failure in the network, the power flows will instantaneously redistribute themselves across the remaining circuits. If any of these are now overloaded, they in turn will fail. Millions of consumers can be blacked out in seconds. To minimise the risk of this, the grid controllers must run the system in such a way that power flows will be within safe limits, not just given the present state of the network, but if any link in it suddenly fails. This implies leaving a margin of spare capacity on every part of the network, with the greatest margins on links liable to receive additional flows after failures.

Given these requirements, it is hardly surprising that markets for electricity have to be designed with some care. At one level, it is possible to write simple contracts, in which a power trader buys an even flow of 50 MW throughout the year from a generator, and sells this power to a large industrial consumer, at agreed prices. When it comes to delivery, however, the generator's operations must be co-ordinated with the rest of the industry. In practice, it is unlikely that the generator will actually produce the even flow of power envisaged in the contract, and even less likely that the consumer will take it. The market should ensure that generators are paid for all the power they actually produce, and that consumers pay for all the power they consume.

All power markets therefore contain some kind of short-term balancing mechanism, in which actual operations are paid for. Prices in these short-term markets are driven by unpredictable events, and are very volatile. Most generators and consumers do not wish to face this level of volatility on their overall bill for electricity, and so most markets provide a way of hedging short-term volatility, allowing participants to make most trades in advance.

In many cases, trades made in advance are for physical delivery, and are settled by deducting the amount agreed in the contract from the generator's actual output when determining the amount it has sold in the short-term market. An identical effect can be obtained with a financial contract, however, in which case the generator's entire output is sold in the short-term market. The contract specifies a side payment, inversely related to the price in that market, that should fix the total amount (short-term market price plus side payment) paid for the volume in the contract.

Either system implies that the generator will receive the price in the short-term market if it increases its output level, however. If it reduces its output level, compared to the amount in its earlier contracts, it will have to pay the short-term price. If the short-term price equals the marginal cost of providing power from other stations, this would give each generator in a competitive market the correct incentives. It should increase output if its own marginal cost is below that of the power it would replace, and reduce output if its own marginal cost exceeds that of other stations with spare capacity. The marginal cost of electricity depends upon its location in the network, however, and this must be taken into account if the market is to produce a fully efficient result.

We can summarise this discussion by listing the main objectives for a set of electricity trading arrangements. The arrangements must ensure that the grid remains electrically stable in operation. They must ensure that all power is paid for. They should encourage the efficient operation of the system, using the cheapest power stations available to meet demand. As far as possible, agents should have the choice of making most transactions at stable prices, agreed in advance.

The next section of the paper shows how marginal cost pricing can give efficient results in an electricity market, and ensure that generators cover their costs. Section III begins the discussion of the European experience with electricity markets, with an account of the Electricity Pool of England and Wales. Section IV discusses NordPool and the other markets in Scandinavia, and Section V the electricity market in Spain. Section VI examines the New Electricity Trading Arrangements recently adopted in England and Wales. Section VII discusses the experience with competition in retail markets, and Section VIII concludes.

II. THE MARGINAL COST OF ELECTRICITY

In general, prices should equal marginal costs in competitive markets. What is the marginal cost of electricity? It is easiest to start our discussion with the traditional model of peak-load pricing for a public utility. At off-peak times, the price should be set equal to variable cost (v

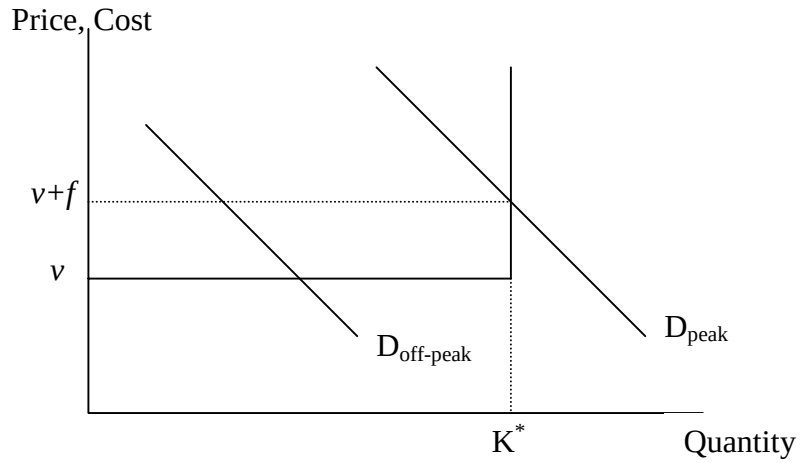


Figure 1

in figure 1). At peak times, the price should be set equal to variable cost plus fixed costs ($v+f$ in figure 1). Capacity (K^*) should be chosen so that it just equals demand at this price. If there are constant returns to scale, revenues will exactly equal costs. We could generalise this to any number of sub-periods, and to "off-peak" demands which would exceed the available capacity if priced at variable cost, by raising the price above variable cost whenever it is necessary to reduce demand to capacity. The level of capacity should be chosen so that the sum of these payments (above variable cost) per unit over the period as a whole should equal the per-unit cost of capacity.

In reality, demand is random, since it depends on the weather. If it were still possible to ration demand by price, the optimal level of capacity would now be at the point where the expected price is sufficient to cover fixed costs. Revenues will now vary with the level of demand, and a low level of demand would imply that revenues were insufficient to cover costs. In most electricity markets, however, most sales take place at prices fixed in advance, based upon the expected level of demand, reducing the impact of the out-turn level of demand on generators' revenues.

The level of available capacity is also a random variable, since it depends upon the unpredictable number of plant outages. Once again, the welfare-maximising level of capacity is such that the expected price, given available capacity and demand, is sufficient to cover the expected costs of the marginal power station. The question is how this price is set. If the market has an active demand side, with enough consumers willing to adjust their demand in real time in response to price signals, then it will be possible to set a rationing price at the intersection of demand and available capacity, as in figure 1. If demand at normal prices is too great, relative to available capacity, however, then the market will run

out of consumers willing to reduce demand voluntarily, and involuntary disconnections will be required. In this case, there is no market-clearing price, and administered arrangements will be needed. In Australia, the real-time price is set at an estimate of the cost of involuntary disconnections (originally A$ 5,000 per MWh) if such disconnections actually take place. In the Pool in England and Wales, the spot price was set a day in advance, and included the expected cost of involuntary disconnections – the Value of Lost Load multiplied by the Loss of Load Probability. Both rules can produce the same expected revenues, if they use the same cost of disconnections, and if the Loss of Load Probability used to set prices is an accurate measure of the real risk of power cuts.

Whatever system is used, the intention is that marginal plants will expect to earn an amount that represents consumers' willingness to pay for having their capacity available. If prices are set by supply and demand, this is measured directly as the premium over the variable cost of electricity that consumers are willing to pay for power at peak times. If prices are administered, it reflects the value of the power cuts that an extra unit of capacity can prevent. If the market has the correct amount of capacity, the marginal unit will just cover its costs. If the market is over-supplied with capacity, revenues will be low, investment will be unattractive, and owners of high-cost, typically old, stations will consider closing them. If capacity is short, revenues will be high, and the owners of old stations will have an incentive to defer their retirement. Even though short-term market prices are set from day to day, and would not provide a safe signal for long-term decisions, the hedging contracts available in most markets allow generators and retailers the security of locking themselves in to prices for a year or more.

The level of variable costs requires a short discussion. In a thermal system,[1] fuel costs are by far the greatest component of variable costs. Stations that are intended to operate for most of the year should have relatively low fuel costs, perhaps at the expense of higher fixed costs, while for peaking stations, only running at times of maximum demand, it is acceptable to have higher fuel costs, in return for lower fixed costs. The upper panel of figure 2 shows a possible trade-off between fixed and variable costs for two types of plant. Peaking stations should be built to meet demands lasting for t^* hours or less, while base load stations should be used to meet longer demands. The lower panel shows the resulting prices. At off-peak times, the price is equal to the variable cost of the base load stations ($V_{base\ load}$),

[1] That is, a system where most power stations are powered by fossil fuels or nuclear power. Norway has an almost entirely hydro-based system, and Finland, Spain and Sweden have a high proportion of hydroelectricity in their mixed systems. Other European countries have mainly thermal power systems.

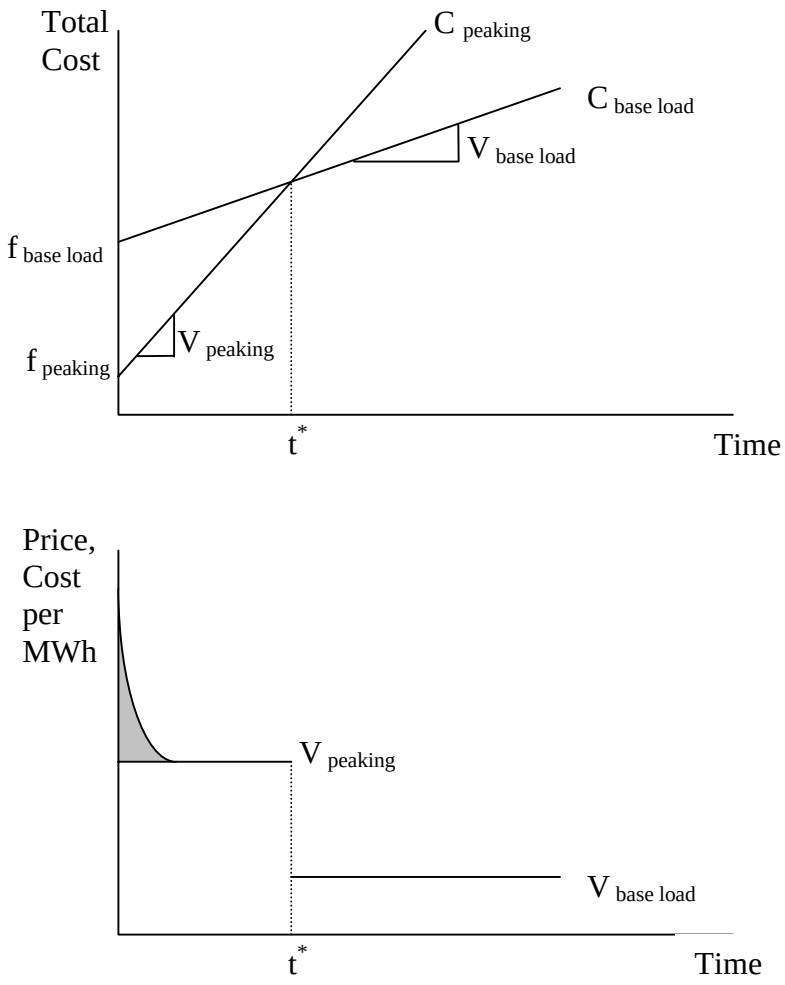


Figure 2

which are marginal at those times. As demand rises, the peak stations become marginal, and set the price at the level of their variable cost ($V_{peaking}$). At peak times, the price rises above variable cost to pay for capacity, as discussed above. The amount raised, given by the shaded area, should just equal the fixed cost of the peaking plant ($f_{peaking}$ in the upper panel) in equilibrium. This allows the peaking plant to recover its costs of operating for up to t^* hours. By the definition of t^*, this is also enough to cover the costs of operating a base load plant for this length of time (a mix of fixed and variable costs), and the off-peak price, at its variable cost, suffices for the extra cost of longer operating hours.

We should also discuss the special case of a hydro system. It is extremely unusual to be able to build a reservoir that is large enough to allow the associated hydroelectric power station to run throughout the year. In order to store enough water to meet the system's energy needs, so many power stations will be needed that there will always be enough generating capacity to meet the peak demand. The system is energy-limited rather than

capacity-limited, as a thermal system is. In a pure hydro system, the marginal cost of energy is the shadow value of water – the price that should be set to ration demand to the expected amount of energy available, given rainfall patterns and storage capacity. This price will be almost constant over short time periods, but will rise in periods of low rainfall, and fall when water is abundant. This contrasts with prices in a thermal system, which will very with the level of demand over the day and across the seasons, but should not necessarily change from year to year unless fuel prices change. In a mixed hydro-thermal system, both types of plant may be marginal at different times, and the shadow value of water will equal the variable cost of the thermal stations that would increase their generation in response to a reduction in the available water supply.

Finally, we should acknowledge the spatial nature of electricity, and the costs of transmission. As these are covered in the paper by Crampes and Laffont, we can be brief. Schweppe *et al* (1988) proposed a system of optimal spot pricing, in which the price of power at any node on a network would equal the marginal cost of electricity at that point. This might equal the cost of generating the power elsewhere and transmitting it to the node, incurring electrical losses, and having to cope with constraints. If it is necessary to use an expensive, but local, generator instead of increasing output from a cheap generator at the wrong end of a heavily loaded transmission line, this should be reflected in the pattern of prices. At a node where a power station is running, its marginal cost should set the price. With an optimal dispatch, the two formulations (the cost of a local station, and the cost of generating the power elsewhere and moving it to the node) will be equivalent. In general, the marginal cost of power will be lower in an exporting area where generation exceeds demand, and higher where demand exceeds the local generation. Because electricity flows cannot be directed, however, a constraint on one part of the network will affect the pattern of transmission costs throughout the country, and counter-intuitive effects can sometimes be seen.

III. THE POOL IN ENGLAND AND WALES

England and Wales was the first power system in Europe to introduce market-based reforms, separating generation from transmission, distribution and retailing.[2] Since the industry believed itself to be a pioneer (possibly unaware of the successful spot market in Chile), the

[2] Retailing is known as "supply" in the UK, but since this paper concentrates on wholesale markets, in which "suppliers" form the *demand* side, it is perhaps sensible to use the generic "retailers". Distribution and retailing

system operators felt that they were heading into the unknown, and were worried that the changes would make the transmission system hard to control. It took a long time to decide how the wholesale market would work, and the eventual choice had a lot in common with the operating systems used before privatisation. The Central Electricity Generating Board had used a computer program called GOAL to perform a least-cost dispatch of all its stations. At privatisation, this planning process was transformed into a market known as the Pool, through which practically all power trades had to pass. Generators, and a few customers, submitted daily bids to the National Grid Company (NGC). GOAL was still used to perform the dispatch, and so the bids consisted of the same information as before, but with cost parameters converted into prices.

The System Marginal Price was based on the five price components bid by the most expensive station in normal operation – it can be described as the average cost of output from the marginal plant. A Capacity Payment was added to this, based on the calculated Loss of Load Probability, multiplied by the Value of Lost Load (set by the government at privatisation and increased annually for inflation). As described above, this should measure the economic value of capacity (as long as the formulae are correct) and give generators the correct incentives to make capacity available. The sum of SMP and the Capacity Payment gave the Pool Purchase Price, paid for generation in the "day-ahead unconstrained schedule". Actual generation differed from this schedule, and a number of other payments were made, ensuring that every generator recovered at least its own costs, whatever its actual pattern of generation.[3] These additional payments were recovered through a charge known as Uplift, added to the Pool Purchase Price to give the Pool Selling Price. Retailers did not have to bid quantities in advance, but paid PSP for all their demand, whether predicted or not.

Scheduling all plant in this way was intended to secure the electrical stability of the system. It should also produce an efficient dispatch, as long as the scheduling software is appropriate, and generators' bids reflect the true costs and physical characteristics of their

became separate activities, but the Regional Electricity Companies (RECs), the distributors, inherited a monopoly over retailing to small consumers.

[3] The payments ensured that the station would recover its costs, if these were equal to its bid prices. Since prices could exceed the station's costs, some stations made large profits from these additional payments. The Pool could not differentiate between changes in generation instructed by NGC, for which such compensation is appropriate, and those that were not. In this sense, the generators' bids were not "firm". Most industries have separate real-time markets, which automatically pay a good price for changes in generation that help the system, and a poor price for those that do not. When power is short, the real-time price will rise, and generators that fail will have to pay a lot to replace their output. In the Pool, the failing generator simply had to buy back its power at its own bid price, almost certainly below its replacement cost.

plant. If the market is not sufficiently competitive, however, bids are unlikely to be cost-reflective, and the dispatch may not be efficient.

Generators could obtain price stability by trading contracts for differences (bilateral, and often tailored to the parties' needs) or electricity forward agreements (also bilateral, but standardised, and traded on an electronic exchange). One of the rules of the system was that practically all electricity had to be traded through the Pool, and so these contracts were financial hedges, rather than contracts for physical delivery. A standard 100-MW contract with a strike price of, say, £25/MWh, would commit the generator to pay £100×(PPP – 25) during each hour[4] for which the contract was valid. If the generator also produced 100 MW, its revenues were fixed at £2,500/hour. This gave it stability, but it still received the Pool Purchase Price for any extra output (and lost PPP for each MWh by which its output fell below the contracted quantity), giving it appropriate incentives to change its output level. The contracts were not quite as helpful for retailers, who had to pay the Pool Selling Price, since they were generally settled against the Pool Purchase Price. This left retailers exposed to the level of Uplift, and they complained bitterly during the early years of the market's operation, when Uplift rose significantly.

Uplift represented the costs of keeping the system stable, and the Pool performed well on this count. Pool prices were not a good signal of marginal costs, however. The industry was privatised with two dominant generators (Green and Newbery, 1992, 1997) and their market power allowed them to raise prices well above marginal costs. Since they in fact raised prices well above the entry level, a large amount of new capacity was built, somewhat in advance of need, and to the detriment of the British coal industry.

[4] Pool prices were actually set for each half-hour, but it is easier to work in hours for this example.

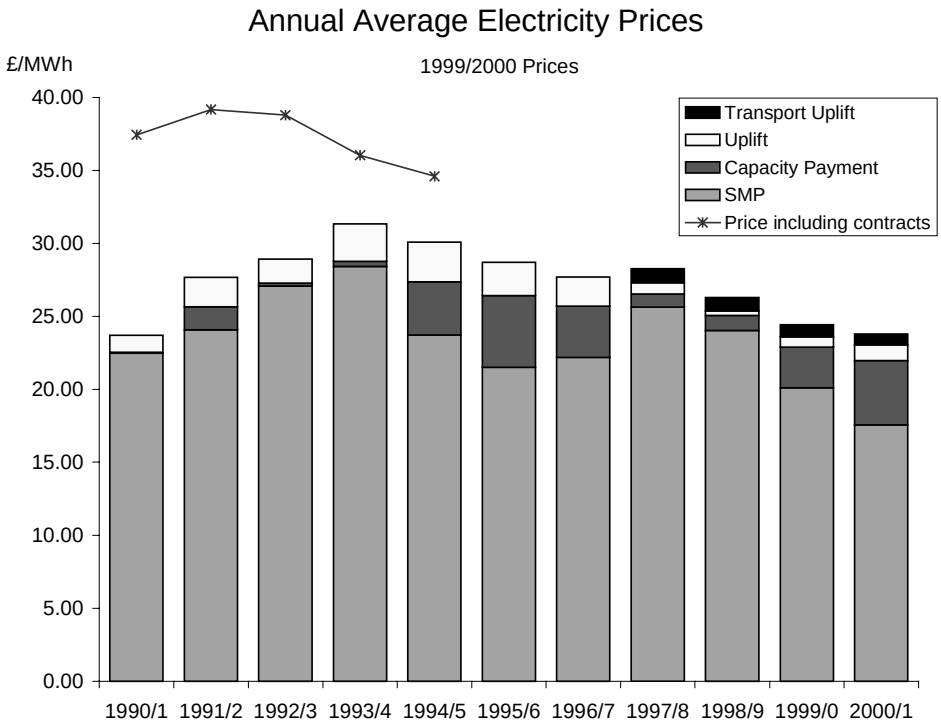


Figure 3

The price stability that comes from contracts was important in allowing a successful privatisation. In 1990, the restructuring had been built around three-year contracts that committed the generators to take practically all of their coal requirements from British Coal, at prices above the cost of imports, and linked electricity contracts that allowed them to pass the excess cost of this on to the Regional Electricity Companies, the distributors. In turn, the RECs were allowed to pass on the cost to their captive smaller customers.[5] The package allowed final prices to these consumers to be held constant in real terms (they had been increased just before the privatisation), but the contract prices were well above the predicted level of Pool prices. These were expected to follow the generators' marginal costs, based on the price of imported fuel. At first, Pool prices were even lower than expected, but they rose rapidly, as shown in figure 3, provoking warning comments from the regulator. Even so, they had little impact on the generators' profitability, which was largely determined by the coal contracts.

In 1993, a second set of five-year contracts was negotiated. The price and volume of coal to be taken fell, leading to pit closures and provoking a massive (if short-lived) political

[5] The nature of the RECs' regulation also encouraged them to invest in new generation, since they could pass the costs, and their profits, straight on to their captive customers, subject only to an "economic purchasing" requirement that never found an uneconomic purchase.

row, but entry by gas-fired plants, and rising nuclear output, meant that there was no room in the market for any more coal burn. In the associated electricity contracts, the major generators passed on the reduction in the cost of coal, but their profit margins stayed high. In 1994, the regulator determined that Pool prices had become too high, and forced the generators to agree to hold them down, and to divest about 15% of their capacity. This restrained prices, but the regulator still believed them to be above the level of entrants' costs, and new entry certainly continued apace.

The generators also gained from the complexity of the Pool Rules, which they could "game", fine-tuning their bids to exploit idiosyncrasies of the scheduling algorithm. An early example was PowerGen's strategy of declaring plant unavailable one day in advance, reducing the apparent capacity margin and raising the level of capacity payment, and subsequently re-declaring it to be available. The capacity payment was not re-calculated, and so the company received the higher payment for the plant whose "absence" had raised the price in the first place! The Pool Rules were soon changed to close this particular loophole, and the regulator imposed reporting requirements, covering their stations' planned and actual availability, on the major generators. Nevertheless, the episode shows the difficulty caused when generators are able to search for loopholes to exploit, and the authorities can only react to problems once they have occurred. The complexity was a disadvantage of determining an entire day's prices from a single set of bids – markets that allow different bids for each period (even if they are all submitted at the same time) can be much simpler.

The Pool was also prone to price spikes, when a high bid set the system marginal price. At times, this is an appropriate price signal for an electricity market, but the Pool also suffered from glitches in its rules that made it more susceptible to spikes than was strictly necessary. Once such a problem was identified, the next step might seem obvious: change the rules to get rid of it. Unfortunately, the Pool's decision-making procedures were designed to protect minority interests against exploitation by a majority. Since most potential rule changes will be against the interests of at least some companies, those affected have been able to delay changes, sometimes for years. European Electricity Directives have been known to suffer a similar fate! Joskow's paper reports the delays in taking action to resolve problems during California's energy crisis.

One example of the Pool's slowness concerns the treatment of transmission losses and constraints. The Pool started on a non-geographical basis, because it had been impossible to reach agreement on geographical charging in the time available to create the

market. All consumers' metered demands were raised (in proportion) until they equalled the level of metered generation, so that all faced the same payment for transmission losses. Transmission constraints were dealt with by "counter-trading" – NGC bought power from generators in import-constrained areas, and sold an equivalent amount back to generators who were on the exporting side of a constraint and were unable to sell their output. As discussed by Crampes and Laffont, these trades inevitably involve buying at high prices and selling low, and the net cost was recovered from all consumers through Uplift. The Pool never succeeded in introducing more cost-reflective charges - an attempt to scale metered demands by loss factors which varied from zone to zone was delayed for years and eventually ended up in the law courts. By the late 1990s, most observers were convinced that the Pool was incapable of reforming itself, leaving the market in a vulnerable position, should change be required of it.

IV. THE NORDIC ELECTRICITY MARKETS

Norway was the second country in Europe to implement market-based reforms in electricity. Practically all the generation in the country is hydroelectric, and there were (and are) a large number of vertically integrated utilities. There was a market for occasional power in which companies could trade small volumes in order to manage water levels in their reservoirs, but the distribution utilities (generally municipally owned) were required to own, or have long-term contracts, for their expected power requirements. Every utility kept spare capacity to cover the risk of low rain- and snowfall, giving the potential for excess capacity in the system as a whole. In 1990, the government introduced a market-based reform, intended to create competition in generation and retailing. All companies were allowed open access to the grid system, and distribution utilities were no longer required to have long-term contracts for all of their requirements.

The existing market for occasional power evolved into a fully-fledged spot market, with participation from generators, retailers and end-users. The market was run by a subsidiary of the grid company, Statnett marked. Participants bid straightforward supply or demand functions, giving the number of MWh they wish to buy or sell at each possible price in each hour. The market-clearing price, where demand equals supply, is determined by the market operator. Because firms bid separately for each hour, there is no need for complicated scheduling procedures or market rules to determine many periods' prices and quantities from a single set of bids. Accepted trades are "firm" on the participants, and are settled by physical delivery or consumption. Taking part in the spot market has always been

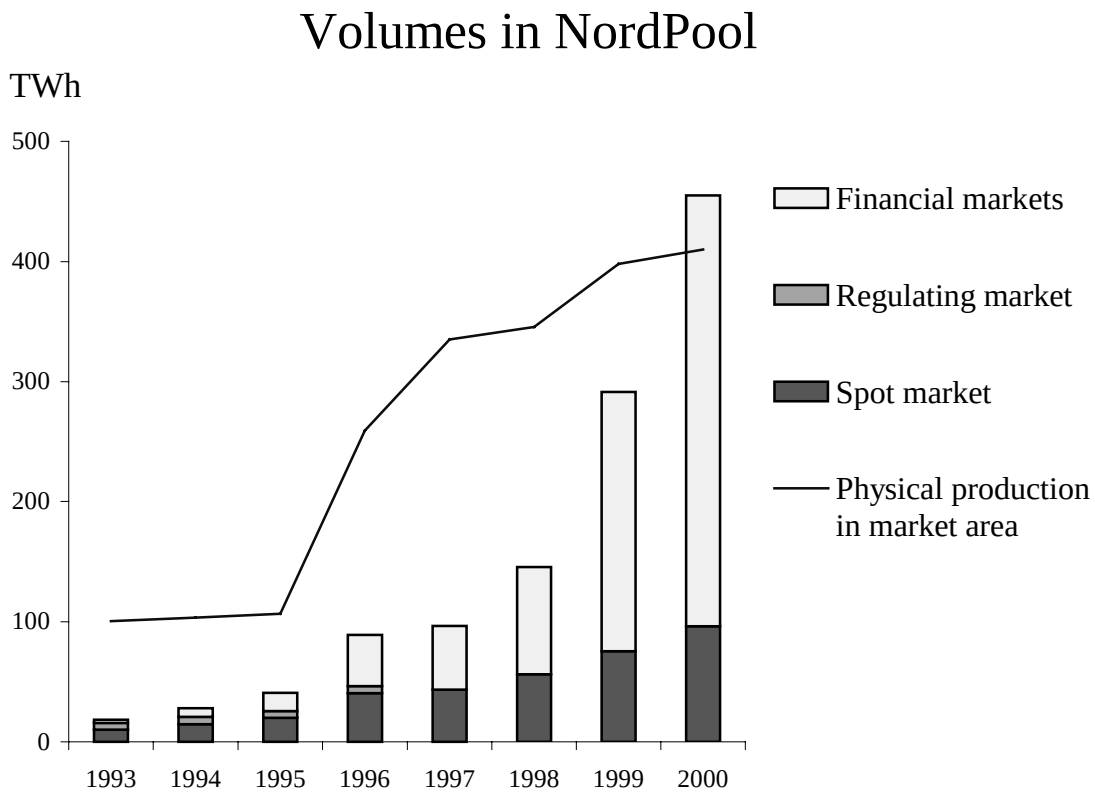


Figure 4

entirely voluntary, however, and utilities also use self-generation and bilateral trades to meet their demands. Before the reform, the prices of bilateral contracts were generally based upon the participants' production costs, while after the reform, they depended upon supply and demand. There was a real-time regulating[6] market to balance the system, and to resolve any imbalances between utilities' contractual and physical positions.

In January 1996, Sweden joined the market, which was renamed NordPool. Generation in Sweden is dominated by Vattenfall, with more than 50% of output, and four companies account for 85% of generation in the country. If Sweden had liberalised the electricity industry on a national basis, problems with market power would have been severe (Andersson and Bergman, 1995). Joining NordPool, however, meant that Vattenfall's share of the combined market fell to just under 30%, reducing its potential market power (Bergman and von der Fehr, 1999). Finland joined the market in October 1996, the eastern part of Denmark in 1999, and the western part (the two are not directly connected) in 2000.

[6] The term has nothing to do with economic regulation, for the market is designed to regulate the frequency on the system.

Figure 4 shows the increasing volume of trade in NordPool, relative to production in the member countries.[7] The spot market accounts for a rising proportion of the countries' output, but it is still only a minority. Most electricity is traded bilaterally, or distributed by the company that generated it, which allows price stability. NordPool runs an increasingly successful futures market, allowing participants to trade electricity up to three years in advance, but this is for financial settlement, not physical delivery. A contract commits its parties to make payments based on the NordPool spot price, as with a CfD in the UK, rather than to generate or consume electricity. Some of this trading is to hedge physical positions and reduce risks, but there is also a lot of speculation. Volumes in NordPool's futures market are now approaching the physical level of consumption in the market area, and far exceed those in the underlying spot market, which is a sign of a mature commodity market. The spot price is believed to follow market fundamentals, so that traders are willing to take positions against it, and the rising level of liquidity in the market has created a virtuous circle.

Furthermore, NordPool's futures market is only a small part of the electricity trading system – there are over-the-counter markets for bilateral trades. NordPool operates a clearing system that allows parties to these contracts to settle up with NordPool, rather than with their original counter-party, reducing credit risk and allowing opposing trades in opposing directions to be netted off. In 2000, NordPool settled contracts, originally traded on other markets, for 1,160 TWh of electricity – roughly three times the volume of physical production, and twelve times the volume traded in the spot market.

The NordPool system clearly works for financial traders – how well does it handle the task of trading physical electricity across borders? The main spot market runs each day, to schedule trading for the following day. Initially, demand and supply bids are ranked in single stacks, and hourly market-clearing prices for the entire market area are obtained. If the resulting trades would breach transmission constraints, however, the bids are reorganised into separate stacks for each market area separated by a binding constraint, and new prices obtained for each area, taking the constraints into account. The power that does flow across the constraint will have been sold (by generators) for a low price, and bought (by retailers or large consumers) at a higher price, giving a surplus that can accrue to the grid operator. Companies involved in bilateral trades across the constraint must make an equivalent payment.

[7] The regulating market in Norway was initially operated by the company that became NordPool, but it was transferred to the Norwegian grid operator in 1996.

This is equivalent to the system described in the paper by Crampes and Laffont (box 1), and should lead to an efficient result. The one problem with separating markets in this way is that market power is no longer restrained in importing areas once the lines are constrained. Bergman and von der Fehr (1999) cite figures showing that Sweden has its own price 60% of the time, and that demand cannot be met without production from Vattenfall, giving that company significant market power, 25% of the time.[8] The Danish electricity industry, where generation is controlled by two regional associations, is also highly concentrated, and monthly market prices during 2000 were generally well above those in the rest of the NordPool area.

Real-time system operation is not dealt with in NordPool, but is the responsibility of national system operators. In Sweden and Finland, a short-term market called Elbas allows companies to adjust their positions up to two hours before real time. Each country has a regulating market for real-time adjustments, and cashing out imbalances between a company's contracts and its physical position. In Sweden and Finland, local transmission constraints are dealt with by counter-trading on the part of the grid operator, as with the Pool in England and Wales. In Norway, the method for handling most intra-national constraints is the same as that for dealing with inter-national constraints in NordPool, with the country being separated into several market areas. Their boundaries are not fixed, but announced a week at a time in response to the changing conditions on the grid. This increases their effectiveness at handling constraints, but makes hedging harder for companies that cannot predict which grid area they may find themselves in.

V. THE SPANISH ELECTRICITY MARKET

The Spanish electricity market was liberalised under a law of 1997, which opened the market from 1 January 1998. Spain has a mix of hydroelectric and thermal power stations (in 1996, 23% of output was hydroelectric, 35% nuclear, and 32% coal-fired, with most of the remainder coming from industrial self-generation), and the rules for the spot market combine the approaches of the UK and of NordPool.

The day-ahead market is voluntary, and bilateral physical trading is also allowed, although volumes have been very small. Generators can submit "simple" price-quantity bids for every hour, as in NordPool, and the system price is obtained by the intersection of the bid

[8] That is, demand in Sweden exceeds the capacity of the other local generators, plus that of the cross-border transmission links.

curve with a demand curve. However, generators can also submit "complex" bids, which recognise the constraints upon thermal power stations, compared to hydro stations. Complex bids can include a minimum level of generation, so that if the unit's lowest price-quantity bid is accepted, it must be accepted in full, and a maximum rate of change of generation from hour to hour, reflecting technical constraints on the station. Generators could also specify a minimum level of income that the station must receive over the course of the day if it is to be scheduled. This ensures that the station will not have to bear the costs of starting up, only to produce a small amount of output. If the initial schedule implied by treating all bids as "simple" breaches any of these constraints, generation from the affected unit is adjusted, and the process iterates until all the constraints are met (Gómez and Vázquez, 1998). This is a somewhat simpler approach to the problems of scheduling thermal units than the one adopted by the Pool in England and Wales, which explicitly included start-up costs in the stations' bids.

Generators and retailers can adjust their positions in an intra-daily market, which uses the same rules as the daily market. This market first opened in April 1998 with two trading sessions per day, gradually rising to six sessions by the time of writing (August 2001). Transmission congestion is handled by the system operator, which makes counter-trades, as in all the other markets described in this paper, apart from Norway. The system operator also runs a real-time market to fine-tune generation to the level of demand and cash out imbalances. Spain has therefore adopted a "standard" approach to the problems of ensuring system stability and making sure that all electricity is paid for.

The Spanish market includes capacity payments to generators, but the amount to be collected is determined by a flat rate charge of 0.8 PTA/kWh, on all final demand. Payments to individual generators are linked to their availability, although this is not targeted to their performance at peak times. Large customers, who have hourly meters, face time of day capacity payments, varying from 1.3 PTA/kWh for the 6 peak hours each weekday between November and February, to zero at night-time, weekends, and throughout August. The rest of the revenue needed is collected from distribution companies (and hence small consumers) as a flat rate per kWh throughout the year. For the many small consumers who face a flat rate tariff, due to the costs of metering for anything more sophisticated, the complete absence of time-related price signals in this part of the wholesale market may not matter. Other small consumers could have limited time-of-day metering, and it is unfortunate that there is no attempt to send them an appropriate signal. The fact that a flat rate per kWh is used to calculate the total revenue required, with no relationship to the balance between demand and

capacity, makes the payment look more like a way of increasing generators' revenues than a market signal.

Suspicions about the capacity payment may be reinforced by the historically close relationship between the Spanish government and the major generators. In 1995, the publicly-owned Endesa contributed 37% of generation, and the privately-owned Iberdrola another 28%. This was very close to the level of concentration that caused such problems in the UK. Rather than attempting to break up Endesa, however, the government allowed it to take over two smaller generators, raising the group's market share of generation to more than 50% (Regibeau, 1999). The industry was also vertically integrated, with both the Endesa group and Iberdrola responsible for about 40% of distribution.

This vertical integration will have been one motive for the generators to avoid abusing their wholesale market power, since retail tariffs for most customers were fixed by the government, after receiving advice from the sectoral regulatory commission. A further motive was given by the system of stranded cost recovery. To allow the generators to recover the costs of past investments, the difference between the regulated final price (net of transmission and distribution costs) and the cost of power in the wholesale market was paid into a competitive transition fund, and shared amongst the generators. The maximum that could be paid to the generators over a ten-year period was fixed at just under PTA 2,000 billion, but this amount would be reduced to the extent that the average cost of generation exceeded 6 PTA/kWh (CNSE, 1999). In 1998 and 1999, the average cost was below this level.

In September 1998, however, the government and the major generators agreed to "securitise" these payments, fixing their level, and changing the recovery method to a constant 4.5% surcharge on electricity bills. This was against the advice of the regulator, which believed that the payments would be excessive, and regretted the loss of flexibility. This agreement removed one incentive for keeping wholesale prices down, and the final price was 12% higher in 2000 than in 1999. In February 2001, the system was changed back to link the payments to the difference between tariff receipts and the wholesale price. This means that the industry again has a "natural hedge" for wholesale prices, which should aid price stability. This is one advantage of this kind of scheme (compared to the fixed surcharge). A second is that the European Commission (2001) clearly favours schemes that relate stranded cost payments to the difference between market revenues and costs incurred before the introduction of market reforms, and might reject fixed compensation schemes as illegal state aid. The third advantage is that the payments reduce generators' incentives to

raise prices, in the same way as long-term contracts, and prices in March to July 2001 were 8% below their level of the same months in the previous year.[9]

VI. BRITAIN'S NEW ELECTRICITY TRADING ARRANGEMENTS

The last electricity market that we will consider is the newest. Section III ended with the observation that many people had long been growing impatient with the Pool's inability to change its rules and reform its procedures. In May 1997, a Labour government was elected, with historical ties to the UK's dwindling number of coal miners. Since privatisation, the demand for coal had nearly halved, as coal-fired stations were replaced by Combined Cycle Gas Turbines. There were accusations that this had only happened because the nature of competition in the Pool was unfair, due to its uniform-price rule. Gas-fired stations could submit a bid of zero, but still receive the Pool Purchase Price, set by coal-fired stations that had to submit higher bids, and therefore lost market share.[10] The fact that all stations received the price set by the marginal bidder was also believed to make the Pool more vulnerable to the exercise of market power. In October 1997, the government announced a review of the Pool. The following year, an Energy White Paper announced that the Pool would be replaced with New Electricity Trading Arrangements (NETA), which started, after some delay, in March 2001.

NETA has replaced the Pool's centralised market with as much freedom to contract as possible. There is no central market until 3½ hours before real time, and most electricity is being traded via bilateral contracts or in futures markets. Three companies have set up power exchanges to facilitate trading in the last day or so before real time operation, left entirely to their own devices by the government and the regulator. In the last 3½ hours before real time operation, the grid company, NGC, runs a balancing mechanism, keeping the system stable by accepting bids and offers to raise or reduce generation and consumption. Each accepted trade is paid its own bid or offer. The balancing mechanism is also used to resolve transmission constraints, by counter-trading. The regulator has proposed a system of tradable access rights to deal with congestion (Ofgem, 2001b). Generators would have to pay for rights in an exporting area, but could be paid for them in an importing zone. Adding the price of these rights to a (still uniform) energy price would give overall prices equilvalent to those described by Crampes and Laffont (box 1). The proposals may

[9] This price fall, and the earlier price rise, could of course have been caused by other factors, but their directions are compatible with the view that securitising the stranded cost payments removed a mechanism that mitigated market power.

not be adopted, as they would be expensive to implement, and congestion costs are currently low.

Bilateral contracts, signed well in advance, give traders the option of price stability, although it is also possible to sign contracts linked to prices in a short-term market, or to delay trading until those markets open. The 3½ hour window for the balancing mechanism has been chosen as the minimum that NGC believes is necessary for it to guarantee system security – the regulator hopes to reduce the length of this period in the future.

Traders have to notify Elexon (formerly described as the Balancing and Settlement Company) of their electricity contracts, which are compared with their metered demand or output. Companies with imbalances between their contracted positions and their metered quantities have to cash them out at unattractive prices. Companies needing to buy power (generators who have under-generated, and retailers with consumers who have bought more than the retailer contracted for) have to pay the System Buy Price (SBP), the average of the prices that NGC paid to buy power in the Balancing Mechanism.[11] Companies needing to sell power get the System Sell Price (SSP), the average of the prices that NGC received for selling power in the Balancing Mechanism. The intention is to give companies a strong incentive to contract accurately in advance, for the SBP is normally expected to be greater than the cost of power bought in advance, while the SSP is expected to be lower. In the first summer of NETA's operation, the median difference between the SBP and SSP was £16/MWh, and the mean was £30/MWh. This implies a very skewed distribution, for while the SSP can be negative (and was in one-seventh of the half-hours from June until August), it is not as volatile as the SBP, which exceeded £100/MWh, nearly six times the mean price on the UKPX's day-ahead market, in 6% of half-hours. In five half-hours, the SBP was more than £1000/MWh greater than the SSP.

It is very early to assess NETA's performance. Volatile balancing market prices were to some extent a design objective, to encourage early contracting. Generators who cannot predict their output in advance, such as wind generators, are likely to make losses because of this volatility (Bathurst and Strbac, 2001). Responding to a government request to review their position, the regulator conceded that wind generators' position had been worsened (although other renewable generators had not been particularly disadvantaged), but argued that this should not be used as a reason to change NETA (Ofgem, 2001a). If the

[10] In other words, the incumbents kept prices high, encouraging entry, and therefore losing market share.
[11] The most expensive trades are excluded from this calculation, on the basis that these will have been required to deal with transmission constraints, rather than to balance the overall level of demand and generation.

government wants to promote a particular level of wind generation, more direct support might be required.

For consumers, however, the overall price of power will be a more important measure of NETA's success or failure. The regulator has complained that Pool prices were as high at the end of the 1990s as at the start of the decade, despite significant reductions in generating costs. Figure 3 showed that this is true, and that it is misleading. The effective price for most electricity trades are set by contracts, and average contract prices were well above Pool prices in 1990/1. Over the next few years, however, the differential between contract prices and Pool prices fell.[12] This means that the effective price of electricity *has* declined significantly over the decade, even though it may still be above the level of a new entrant's costs.

The government and the regulator have argued that moving from uniform pricing to bilateral trading will reduce the scope for generators to exercise their market power. Studies in auction markets for government securities (quoted by Newbery, 1999) imply that moving from uniform pricing to a discriminatory auction will make prices more competitive for a given number of traders, but may discourage entry, reducing the number of traders, which could make prices less competitive overall. Bower and Bunn (2000) used a simulation model to predict that moving from the Pool to a discriminatory pricing auction would increase prices. The regulator and the government have pointed to the reduction in Pool prices at the end of the 1990s as evidence that the market was anticipating NETA.

It is hard to see how an expected reduction in the future price of a commodity can affect its current price, when the commodity in question is non-storable. The normal arbitrage conditions via storage fail to apply. That means that the price reductions preceding NETA have more to do with a reduction in market power, as National Power and PowerGen sold a number of their stations to other companies, than the future change in market rules. The companies' incentive to use their market power may also have been reduced by a fear that the new regulator would become less tolerant in the run-up to NETA.

To find the true impact of NETA, we must await a study of pricing behaviour, linked to market structure, once the market has been operating for a reasonable period. At present, we can only say that although the Pool's governance structure proved inadequate, it might

[12] The line in the figure takes actual revenues per kWh generated for National Power, PowerGen and Nuclear Electric, and assumes all other sales were made at the Pool Purchase Price. From 1992/3 onwards, entrants linked to the RECs generated an increasing proportion of electricity. Their contract prices were well above Pool prices, and calculations based on the major generators' selling prices would have given an underestimate of average prices from the mid-90s onwards.

have been reformed for a lot less than the £500 million that NETA has cost to implement. NETA will need effective arbitrage between a number of separate markets if it is to produce an efficient dispatch.

VII. COMPETITION IN RETAILING

All of these countries allow competition in the retailing of electricity to final consumers. The UK was the first to open part of its market, allowing consumers with a maximum demand of more than 1MW to choose their retailer from April 1990. Most of these consumers saw significant price reductions at this time, since they no longer had to pay a share of the cost of subsidising British coal. This cost was collected through contracts between the major generators and the RECs, while Pool prices were based on international coal prices. The large consumers had the option of buying from a retailer purchasing through the Pool, and so could not be forced to pay the higher prices in the coal contracts. Even so, the major generators were generally seen as offering better prices than the incumbents distributors, and it did not take long before a majority of consumers were buying from a second-tier supplier – that is, not their local incumbent. A few very large consumers, who had enjoyed special terms from the nationalised industry, did see their prices rise for a few years, until offset by reductions in regulated charges for transmission and distribution. This is a general implication of competition in retailing – it is hard to maintain cross-subsidies.

The second wave of consumers to get a choice, those taking between 100kW and 1 MW of electricity, were able to choose their retailer from 1994. Unfortunately, the industry was not ready for them, and many consumers were allowed to change retailer before they had the appropriate half-hourly meter, or before it was linked to the Pool's data systems. They suffered from months of chaos and estimated bills, and the insult of having to pay the electricity industry's costs of putting things right. After the teething problems, however, these customers also saw price reductions, and a significant number also switched retailer.

The problems in 1994 made the regulator (and the government) determined to avoid a repeat when the rest of the market was opened up to competition, due in 1998. It would be impossibly expensive to provide all domestic consumers with meters that could record their consumption on a half-hourly basis, or to transfer the data to the Pool every day. It was decided to use a system of profiles, which could allocate demand between retailers on the basis of the typical consumption patterns of their consumers. The IT needed to do this, and to keep track of consumers as they changed retailer, cost £¾ billion, and was not ready by

the time the market was due to open. To avoid chaos, the start of competition was postponed, and the market was opened in (geographical) stages.

To date, over 25% of customers have switched retailer, for savings of up to 10% of their bill. The incumbents' tariffs were regulated, but at levels that allowed them to recover the cost of existing contracts at above-market prices. Competing firms were able to buy power at lower prices and undercut the incumbents – although most of them were incumbents asking higher prices in other regions! In late 1999, the regulator deliberately set price limits above his previous estimate of the companies' costs, in order to increase the "headroom" facing competitors. Such tactics are likely to increase the number of consumers changing retailer, and hence the apparent success of competition, but are bad for the majority of customers who have so far stayed with their local retailer. Given this, and the cost of setting up the systems required, it is possible to argue that competition in domestic electricity retailing in the UK has actually reduced welfare (Green and McDaniel, 1998).

Competition has been more successful in the Nordic countries. Norway opened its market to all consumers in 1991, but as they all needed an hourly meter, and had to pay a switching fee, the market was only theoretically open to small consumers. The regulator gradually reduced this switching fee, finally abolishing it in 1996, and allowed the use of profiles rather than meters to calculate the cost of power. There is much less intra-day price variation in the Nordic market than in England and Wales, due to the high proportion of hydroelectric power. Within each grid area, all the customers without hourly meters therefore use the same profile. Effectively, each consumer has a constant share, based on their annual consumption, of the non-hourly metered consumption in each hour. The amount to be apportioned to non-hourly metered customers is derived by subtracting the hourly metered consumption and estimated losses from the total input to the grid in each hour (Jonassen, 1998). The relative simplicity of this system means that the costs of introducing full competition were far lower in Norway than in the UK.

Around 13% of Norwegian customers have now switched retailer (Bowitz *et al*, 2000). Bergman and von der Fehr (1999) show that there was a wide range of prices in 1998, but Bowitz *et al* and NVE (2000) give evidence that margins, and price dispersion, appear to have fallen, implying that competition is becoming an effective discipline upon retailers.

The experience in Sweden and Finland follows the same pattern as in Norway. The markets were formally opened in 1996, and some switching took place among large consumers, who enjoyed some price reductions, but few small consumers were willing to

buy an hourly meter. The requirement to buy a meter was lifted in 1998 (Finland) and 1999 (Sweden), and the number of domestic customers changing retailer has risen, although lagging the proportion in Norway.

Spain is liberalising its market in stages. The largest consumers, taking more than 15 GWh a year, were allowed to choose their retailer from January 1998, and the ceiling was lowered to 1 GWh a year in stages by October 1999. By the end of that year, 7,200 consumers taking 27 TWh were on market-based tariffs, representing about 17% of total demand. Regibeau (1999) reports that some of these customers had obtained price reductions of 25-30% at first, but also suggests that the incumbent firms were soon attempting to raise prices. All consumers are due to be given the choice of retailer by 2003, in line with the trend in most European countries.

Newbery (2001) points out that this trend is not without dangers. If large numbers of customers learn to switch retailer on a frequent basis, searching out the lowest prices, then retailers may become reluctant to buy electricity on long-term contracts, as they will be exposed to losses if prices subsequently fall. If there is little demand for long-term contracts, generators will find it hard to guarantee stable prices for their output. This could reduce investment, raising prices, and in extreme cases jeopardising the security of supply. Furthermore, since contracts have the effect of mitigating market power, generators who are unable to sell their power in advance are likely to set higher prices than would otherwise be the case. High prices driven by market power are likely to encourage investment, of course, but the fact remains that an electricity market with too little long-term contracting is likely to be a volatile place. Joskow's paper shows how many of the problems that California faced came from relying too heavily on short-term markets.

VIII. CONCLUSIONS

All of the markets studied here have passed the most important test – the lights have stayed on. They started with adequate capacity margins, and were able to obtain sufficient investment to meet demand growth. The markets through which the system operator provides short-term electrical balance have succeeded in this task. Long-term financial stability has been provided by contracts and bilateral trading in NordPool and England and Wales, while the stranded cost payments in Spain appear to have had an equivalent impact on generators' receipts and incentives.

NordPool is probably the most successful among them, at least in terms of creating the confidence needed to sustain active secondary trading – volumes in the financial markets

surrounding it are many times those in the spot market itself. The market has benefited from the relative ease of hydro scheduling, and the apparent absence of market power. In England and Wales, the Pool was plagued by market power, which worsened the problems caused by its complex and sometimes inappropriate rules. The vast majority of Pool sales were hedged with bilateral contracts, but there was little secondary trading in the financial markets until the last years of the Pool's life. Financial traders clearly lacked confidence in the Pool, but as the electricity companies had hedged most of their physical positions with long-term contracts, they had little need to trade actively until these expired in 1998. NETA has led to more trading, but this in itself does not necessarily mean that it is a better system. The Spanish electricity industry has a well-designed spot market, but its concentrated structure may reduce the gains from liberalisation.

Other countries are now following these pioneers, with a variety of market designs. What are the lessons from experience to date? There is no "one size fits all" design that should be universally adopted. The very different designs described in this paper have all been reasonably successful – problems have generally come from outside the market design. At the same time, all markets are likely to face some "teething problems", and need to change some of their rules in response to unforeseen issues. It is essential that their governance structures create effective procedures for changing the market rules.

No market design will work well unless market power is curbed. Some European countries have relatively unconcentrated electricity sectors at present, while others are dominated by a small number of companies. When Sweden joined NordPool, the potential market power of its largest company was diluted because of the larger market size. A similar effect could be helpful as trade in electricity across European borders increases. Some of the larger European electricity companies are busily merging with smaller companies in other countries, however. If this trend continues, joining electricity markets together might do little to reduce effective concentration. National competition authorities and the European Commission should take a tough position on mergers that threaten to enhance market power across borders. Competition authorities should also consider taking action against companies with high shares within individual markets.

If trade across European borders can reduce problems with market power, it is important that markets are designed in a way that does not hinder it. This does not mean that every market needs to have identical rules. It does mean that they must have clear procedures for handling cross-border trades, and Crampes and Laffont have already stressed the need to ensure that transmission charges are not "pancaked", making such trades

uneconomic. The regulatory interface between markets is important - Joskow's paper points out that Californian generators could avoid some of the constraints that regulators attempted to impose, simply by selling in a different market. Efficient trading between markets can reduce market power, smooth out peaks and troughs in prices through arbitrage, and reduce the amount of reserve capacity that any one market needs to carry.

It is clear that most companies want to be able to hedge their trades in advance, rather than pricing all of their output in the spot market, and so bilateral trading, or contracts for differences, are a vital component of a market design. Forcing all trades through an unhedged spot market means that volatile wholesale prices must either be passed straight through to consumers (politically unattractive), or utilities will run the risk of a California-type disaster. Long-term contracts promote investment in generation, for only large companies are likely to be willing to stake the cost of a new station upon prices in the spot market. A concentrated market, where market power acts as a kind of insurance, might see adequate investment in the absence of long-term contracts, but consumers would suffer from higher prices.

Electricity retailers are likely to be wary of contracting if intense competition in this part of the market means that they may not be able to make consumers bear the cost of contracts that turn out to be "out of the market", however. There are grounds for believing that using "yardstick competition", which links regulated retail rates to a weighted average of a company's own purchase costs and those of its rivals, and hence reduces the cost of being out of the market, would promote more contracting. At the same time, simply passing through the cost of all contracts, as seemed to happen in the UK, can create incentives for excessive investment. National governments considering whether to go beyond the degree of market opening required by the European Directives have a fine line to tread.

Bibliography

Andersson, B., and L.Bergman (1995) "Market Structure and the Price of Electricity: An *Ex Ante* Analysis of the Deregulated Swedish Electricity Market" *The Energy Journal* vol 16 no 2, pp 97-109

Bathurst and Strbac (2001) *The value of intermittent renewable sources in the first week of NETA*, Tyndall Briefing Note No 2, Tyndall Centre for Climate Change Research

Bergman, L and N-H von der Fehr (1999) "The Nordic Experience: Diluting Market Power by Integrating Markets" in (ed) R.Vaitilingam *A European Market for Electricity?*, London: Centre for Economic Policy Research

Bower, J. and D.W.Bunn (2000) “Model-based comparisons of Pool and Bilateral Markets for Electricity” *The Energy Journal*, vol 21 no 3 pp 1-29

Bowitz, E., T. Bye, O. Rosnes and H.Vennemo (2000) *The Nordic Electricity Reform: Economic and Environmental Consequences*, working paper 3/00, ECON Centre for Economic Analysis

CNSE (1999) *Most frequent asked questions on the Spanish Electric Power System*, Madrid: Comisión Nacional del Sistema Eléctrico

Crew, A.M. and P.R.Kleindorfer, (1976) `Peak Load Pricing with a Diverse Technology', *The Bell Journal of Economics*, Vol.7, No.1, Spring, pp.207-31

European Commission, (2001) *Commission Communication relating to the methodology for analysing State aid linked to stranded costs*, August 10, 2001, available from http:/europa.eu.int/comm/competition/state_aid/legislation/stranded_costs/en.pdf

Gómez, T. and C.Vázquez, (1998) *The Spanish Day Ahead Electricity Market*, mimeo, Universidad Pontifica Comillas

Green, R.J. and T.McDaniel (1998) “Competition in Electricity Supply: Will "1998" be Worth It?” *Fiscal Studies*, Vol 19, No 3, pp. 273-293

Green R.J., and D.M.Newbery (1992) `Competition in the British Electricity Spot Market' *Journal of Political Economy* 100 (5), October, pp 929-53

Green, R.J. and D.M.Newbery (1997) `Competition in the Electricity Industry in England and Wales' *Oxford Review of Economic Policy* Vol 13 No 1, pp 27-46

Hogan, W.W. (2001) *Electricity Market Restructuring: Reforms of Reforms* mimeo, Harvard University

Jonassen, T. (1998) Opening of the Power Market to End Users in Norway 1991 – 1999, report to NVE

Newbery, D.M. (1999) *Evidence on Discriminatory Auctions*, mimeo, Department of Applied Economics, University of Cambridge

Newbery, D.M. (2001) *Problems of liberalising the electricity industry*, presented at the 16th Annual Congress of the European Economic Association, Lausanne, 1st September, 2001

NVE (2000) *Market report 1998-1999*, Oslo: Norwegian Water Resources and Energy Directorate

Ofgem (2001a) *Report to the DTI on the Review of the Initial Impact of NETA on Smaller Generators, August 2001, 52/01*, London: Office of Gas and Electricity Markets

Ofgem (2001b) *Transmission Access and Losses Under NETA: Consultation Document May 2001, 37/01*, London: Office of Gas and Electricity Markets

Regibeau, P. (1999) "Regulatory Reform in the Spanish Electricity Industry: Same As It Ever Was?" in (ed) R.Vaitilingam *A European Market for Electricity?*, London: Centre for Economic Policy Research

Schweppe, F.C, M.C.Caramanis, R.D.Tabors, and R.E.Bohn (1988) *Spot Pricing of Electricity* Boston, Kluwer