ELECTRICITY MARKETS: CHALLENGES FOR ECONOMIC RESEARCH

Richard Green
University of Hull Business School, Hull, HU6 7RX, UK

September 2003

Abstract

This paper discusses a number of challenges facing researchers who study electricity markets. We know the basic ingredients for a successful market, but not how to fine-tune the designs. We know that market power is often a problem, but could do with more accurate models of strategic firms. It is not clear that retail competition offers the best protection for small consumers, and there are unanswered questions about security of supply. Finding effective mechanisms to link markets together, and for regulating the transmission system, will become increasingly important as deregulation spreads. And we should not forget that since a major aim of our research is to find ways in which the industry’s performance can be improved, that won’t happen unless we pay enough attention to ensuring that our advice can be adopted.

1. INTRODUCTION

Electricity markets are a relatively new phenomenon. It is only twenty years since Paul Joskow and Richard Schmalensee (1983) published “Markets for Power”, which arguably launched the idea that electricity should be traded on wholesale markets like other commodities. It is fifteen years since the British government published its White Paper “Privatising Electricity” (Department of Energy, 1988), which launched the first fully-functioning electricity market. It is ten years since Victoria’s Electricity Act 1993, which led to the creation of a market in Victoria, later expanded to other states. The most obvious candidate to continue the sequence is California’s ill-fated experiment, which became operational on 1 April 1998. That is hardly an inspiring example, of course, but it does not seem to have stemmed the flow of countries liberalising their electricity industries. Europe and Latin America are in the lead (and we should not forget that Chile and Norway adopted elements of the electricity market model some time before England and Wales did so), but the move towards electricity markets appears to be world-wide.

Not surprisingly, the trend to adopt electricity markets has encouraged economists to focus their research on these markets. A search on Econlit for “electricity” reveals 42 papers published in 1988. Eight of these were demand studies, seven historical pieces, and five apiece dealt with pricing (most doing so in the context of a regulated industry), and the forthcoming privatisation of the British industry. Going on the titles alone, but taking a broad definition, I classified just five papers under “research on electricity markets”. A similar search for 2000 reveals 163 papers, and more than 50 of them appear to be on electricity markets.

Given all the research that has been done over the last two decades, do electricity markets still pose any challenges for economists? There are some issues where we have reached a consensus, but I believe that there are still enough important unresolved questions to keep a large number of economists busy for some time to come. In each research area, I will start by discussing...
the questions that I believe no longer pose a research challenge, and then move on to the challenges that remain.

2. MARKET DESIGN

We now have a number of market designs that appear to work well, and have unfortunately discovered some that turned out to work badly. The caveat about appearing to work well is important here, since California’s market design seemed to be performing reasonably well for its first two years, albeit with a number of teething troubles that were being resolved. The point is that some of the markets that appear to be working well today may yet succumb to unanticipated problems.

In general, however, we know the main ingredients for a successful electricity market design. Most trading should be conducted some time in advance, and could be either physical, or financial. The choice between physical and financial trading depends upon the design of the real-time, or near to real-time, market. A centrally organised mechanism is needed to ensure that all the electricity produced is paid for, and that the system is kept stable in electrical terms. There has to be some way of setting a price for last minute adjustments to generation, and for unanticipated deviations from expected generation and demand. The price (or prices) set in this way will only apply to the small quantities that had not previously been traded. Most markets have an additional central mechanism, typically operating one day ahead, and the great majority of power will have been traded by the time that this market closes. This day-ahead market could be based upon either a gross Pool or a net Pool model. The gross Pool model deems that all electricity passes through the central market, and so forward trades have to be on the basis of financial contracts for differences, settled against the price in the day-ahead market. In the net Pool model, generators and suppliers can simply inform the market operator of their bilateral transactions, and only residual quantities are scheduled and traded in the central market. In this model, forward trading is for physical delivery. The two models should be equivalent in operation, and the illusion of greater freedom associated with physical contracting probably gives this model the edge in political terms.

In theory, the optimal design for the real-time market is the nodal spot pricing approach used in the PJM Interconnector in the US. This requires the system operator to run a “bid-based, security-constrained, optimal dispatch”, and to set the price at each point on the system equal to the marginal cost of providing more power to that point (Hogan, 1992). Those marginal costs are automatically revealed as a by-product of the calculations required to perform the dispatch. Market participants who inject power at a node are paid the price at that node, while those who withdraw power must pay the nodal price. A bilateral contract that both injects and withdraws power exposes the participants to the difference in the two nodal prices, which is the cost of moving power across the grid. If the injection and withdrawal are at the same node, of course, then there is no charge. This market design ensures that the “correct” price applies at every point on the network and at every point in time, and makes free riding impossible – those who cause congestion on the grid will have to pay for it. The calculations need not be performed only in real time – PJM runs a day-ahead market which is financially binding on those who participate in it, so that only the differences between participants’ actual positions and their day-ahead trades are cashed out at the real-time prices.

There is a challenge here, however, or perhaps two. Most electricity markets use an apparently simpler system, with a single price for large zones, or even whole countries. NordPool splits Norway into three, and Denmark into two, but has single prices for Sweden and Finland. Where there is congestion inside a country or a zone, the (national) transmission operator has to resolve it through “counter-trading”, buying back power in an area of surplus, and buying more inside the import constraint. The cost of doing this has to be passed back to market participants in some way, and it is important not to create incentives to artificially cause congestion.
A research challenge is thus to discover whether zonal pricing systems are really simpler to run than the nodal spot pricing approach. If they are simpler, does this come at a significant cost in terms of the price signals that are sent, and participants’ reactions to those signals? Rule 35 of the British Highway Code instructs drivers to “give signals … if they would help other road users”, and there may be little point in sending an elaborate signal to someone who is never going to react to that signal. (On the other hand, some participants will in time learn to react to signals, even if they do not respond at first, and so we should be cautious about ruling out the use of price signalling.) So a better way of phrasing the challenge might be as “how far can we move away from nodal spot pricing without a significant loss of economic efficiency?” If it turns out that the answer is “not very far”, then we face a political challenge: “how can we persuade decision makers that the nodal spot pricing approach is in practice the best option?”

If contracting in advance is an important feature of a good market design, what kind of contracts should be promoted? There is a trade-off between liquidity and customisation – there may be many market participants willing to trade a base-load contract with no specific location, while very few would be interested in a contract for an unusual load shape at a particular node on the system. One implication, of course, is that there is likely to be greater liquidity in zonal markets, in which there are few locations to choose from. An alternative approach, which should be functionally equivalent to nodal pricing, is to require participants to buy the right to send power across any congested “flowgate”. The number of rights required depends on the participant’s impact on flows across the congested line, so that those located close to the line are likely to require the greatest number of rights. The approach assumes that there will be relatively few congested lines in any market, and so liquid markets can grow for the right to access each of these. The equivalence with nodal pricing comes from combining the prices for each flowgate (common for all participants) with the varying number of rights required at each node. Will participants be able to keep track of the number of rights required in practice? If a nodal system is used, will most participants be able to get an adequate hedge if liquidity is concentrated on a few hubs? The related issue of the best load shape for contracts is less a matter for economic researchers than for the market participants. There are no fundamental issues of market design over whether a four-hour or a six-hour contract, for example, should be favoured, and futures markets are frequently introducing new contracts in the hope of filling a need. Unless we get evidence to the contrary, we should assume that this is one area in which the market can sort itself out.

Another task that still presents a challenge is to understand the factors driving prices in electricity markets. In a few years’ time, this may become routine, as we gain more experience of markets operating under different conditions, but at present, I think some fundamental questions remain unanswered. Should we be following the approach of the finance literature, which treats the price of electricity as a stochastic variable and concentrates on studying its properties in terms of volatility, jumps, and mean reversion? Or should we concentrate on the fact that the short-term price for every period is set by some intersection of demand and supply, and study the interaction of these factors with the market rules? I would be uneasy if the stochastic approach was taken to imply that we cannot understand the out-turn values for each period’s price. However, it may be that explaining every price in turn is too cumbersome, and randomness should be taken as a shortcut for “things that we could explain, but don’t have time for in this application.”

3. MARKET POWER

The vulnerability of electricity markets to market power is now well known. Bolle (1992) and Green and Newbery (1992) predicted that this would be a problem in England and Wales, and many studies have now shown that generators have raised prices above marginal cost when a market is concentrated, or is short of capacity. We need to be careful, of course, because prices should rise above marginal cost when capacity is tight, or generators would never recover their fixed costs, but there have clearly been some times when prices have exceeded desirable levels.
Ex ante studies of market power in particular markets can prove a useful guide (or perhaps warning) to policy makers, but I am not sure that they automatically present a research challenge. Gathering the data takes time and care, but the procedures for building a Cournot model (or even a supply function model) are well-known, and I think the qualitative outputs are generally predictable. We also know that many of the studies to date have over-predicted the level of prices actually seen in the market. (Bushnell (2003), which models the California market for the summer of 2000 using actual demand, fuel cost, and import data, does produce Cournot outcomes that are close to the prices observed in that market, but the case may be exceptional.)

There is a research challenge associated with ex ante studies of market power, but it lies in improving their methodology. If Cournot models over-predict, and supply functions are hard to compute and suffer from multiple equilibria, are simulation models a possible way forward? These use “intelligent agents” to find profit-maximising strategies for their firms, changing a few variables at a time. These strategies may not be equilibria, since all the firms are continuously changing their strategies, but if they appear to converge to a small range, the average of this range may be a good prediction of what will happen. Interestingly, it appears that the range of strategies may be smaller if the agents are not too intelligent, and can only change a few variables at each iteration. The great advantage of this method is that the calculation of profits can take into account many factors that may be hard to model in an equilibrium framework, such as the firms’ holdings of long-term contracts, or transmission issues. This could allow us to obtain more realistic predictions of future prices.

An alternative, which has the advantage of using real people rather than machines, is to conduct experiments on electricity markets – these can give interesting results, but the situations studied must be kept relatively straightforward if the participants are to learn enough to function effectively in a reasonable time.

A second research challenge lies in finding the best way to deal with market power. Are structural remedies the only effective way to prevent problems, and if so, how unconcentrated must the market become? Does long-term contracting, or its logical extension of vertical integration, provide sufficient incentives to keep prices down, or would generators wish to raise wholesale prices in order to raise retail prices and the price of future contracts? Would vertical integration actually make things worse, in raising barriers to entry by non-integrated firms? Can price caps play a useful role, either on individual stations’ offers or on the overall level of prices? If there is a cap on the stations’ offers, should it be tailored to their individual costs, or set at a blanket level? Can ex post regulation, concentrating on movements in bids that do not appear to be linked to movements in costs, act as a sufficient deterrent, or would it amount to retrospective micro-management?

4. RETAIL MARKETS

All of the largest electricity customers in the European Union can choose their supplier, and the EU has decided to allow all electricity consumers a choice of supplier by 2007 – in some countries, this choice already exists. Large customers have quickly learned to shop around, and in doing so, have protected themselves from exploitation by suppliers – there is no future in offering an uncompetitive contract in this part of the market. There is a paradox, of course – even the minority of customers that are not willing to switch their supplier are getting fairly competitive prices, and so the incentive to shop around might seem limited, despite the large electricity bills that most of these customers have to pay.

Small customers seem to be much less interested in shopping around – while nearly 40% of domestic customers in the UK have switched their supplier, the proportions in those Nordic countries that have opened their markets, and in Germany, have been much lower. This is despite the significant savings that are available to customers who switch. One question for research is thus why so many people are not switching, when it is usually an easy process, and savings are
available for the asking. In an economist’s standard model, people will always switch if the gains from doing so exceed the cost of switching. It is possible to estimate the switching costs that would be consistent with the observed gains from switching and low rate of switching (e.g. Sturluson, 2003). To me, however, the resulting estimates seem implausible, given what is actually involved in making the switch. We probably need better models of consumer inertia and apparent irrationality – perhaps not really a matter for specialists in electricity markets. A matter for policy-makers is whether anything can be done to encourage switching.

If we can’t encourage switching, are we right to move towards unregulated competition in retail markets – will the consumers who are not willing to switch be protected by the minority who are? I assume, by the way, that there is a role for public policy in protecting consumers of this type. Can we design a system of incumbent regulation that allows meaningful competition from other suppliers, without unfairly handicapping either the incumbent or the entrants? Littlechild (2003) has suggested that transitional price-caps in the run-up to competition can provide adequate protection, and will prove superior to long-term regulation.

If switching does take off, what impact will this have on the market for long-term contracts and the prospects for investment? Newbery (2002) suggests that large amounts of switching will discourage long-term contracting, although numerical modelling by Green (2002) suggests that the quantitative impact might be small. If long-term contracting is discouraged, will that make investment less attractive?

5. INVESTMENT AND SECURITY OF SUPPLY

Can we be confident that electricity markets will deliver security of supply? The answer that economic theory says that they should will not win us many friends if it is given in the middle of a blackout. Investment will only happen, however, if investors are sufficiently confident that they will be able to cover all their costs, including a reasonable return on capital. Since we will inevitably have spare capacity for much of the year, and a competitive market would tend to keep prices close to marginal cost at those times, we will need very high prices at the other times if generators are to recover their fixed costs from energy prices alone. The research challenge is to discover whether this is the best way for generators to recover their fixed costs, and to find the best way to set prices when capacity is short.

The textbook model suggests that prices should be set by consumers’ willingness to pay, so can demand side participation be used to clear the market? What are the obstacles in terms of metering and behavioural changes that must be overcome before demand side bidding becomes a significant part of the market? Should there be some other mechanism, involving additional payments for capacity? If so, should they be linked to the actual balance between demand and supply, or should they be based on longer-term forecasts? (We can be fairly clear that a scheme like the Spanish “capacity payment” that simply raised all electricity prices by a flat-rate amount is unlikely to generate accurate investment signals.)

Is there a bias against efficient investment decisions? Bunn and Larsen (1992) and Ford (2002) have both suggested that electricity markets would be prone to cycles of over- and under-investment, as investors base their decisions on current prices and fail to take account of the similar moves being made by others. Can a market-friendly mechanism for co-ordinating investment be found? Quite apart from the level of investment, should we be concerned about its type? Building combined cycle gas turbine stations to replace coal- and oil-fired stations will be good for the environment, but could it lead to a dangerous dependence on a single fuel? Can we measure the benefits of diversity, and can we find policies that encourage it, without relying on governments to lay down the proportions of the market that each fuel can take – always assuming that is not a role that we want the government to assume.
Most governments are worried about emissions of Carbon Dioxide, and are promoting investment in renewable generation. How much should we be willing to pay to reduce CO$_2$ emissions? What is the most effective way to promote renewable generation – should we be using direct subsidies, or taxes on carbon, or quantitative schemes – either to promote renewables or to restrict carbon emissions. Almost the only thing that we know for sure at present is that it should be possible to improve upon the UK’s Climate Change Levy, a carbon tax that was designed not to discriminate against the coal industry.

6. LINKING MARKETS TOGETHER

Despite the list of research challenges I have set out, we actually know quite a lot about the workings of individual electricity markets. Now that deregulation is spreading, we face the challenge of linking neighbouring markets together in an efficient manner.

Where the lines joining the markets have adequate capacity, the problem is merely one of coordination, ensuring that the market operators know how much power is due to flow between the two markets. Arbitrage should ensure that the prices in the markets are equalised, at least in expectation. In most cases, however, transmission systems have been designed to carry power within a country, region or company area. The links between systems are much weaker, and likely to be constrained as more agents attempt to trade power over them. Again, if the markets at each end are competitive, the problem is basically one of coordination. Cadwalader et al (1999) describe a mechanism that could be used to ensure that separate electricity markets (based up on nodal pricing, of course) would successfully coordinate inter-system trades. In such a market, it would be necessary to buy the right to access the congested transmission links, and the value of the right should equal the difference in prices between the two lines. In Europe, some international links are being effectively arbitraged in this way, but others are not (Bower, 2002; Newbery and McDaniel, 2003). An important research challenge is to find out why arbitrage has been ineffective, and devise more efficient ways of allocating transmission capacity.

If the markets are not competitive, then manipulating access to transmission becomes another way in which dominant companies can exploit their market power. Joskow and Tirole (2000) showed how transmission rights and output decisions interact in a two-node network, representing the simplest case of trading electricity between regions. Contracts that increase a generator’s exposure to a price that it can affect, such as those that allow it to import electricity to its home market, are likely to enhance the generator’s market power. Contracts that reduce the generator’s exposure to a price that it can affect, such as export contracts from its home market, are likely to reduce its market power. Gilbert et al (2003) have extended this work to show that if transmission contracts for a two-node network are allocated in a uniform price auction, generators will only obtain contracts that mitigate their market power. If generators inherit the wrong kind of contracts or a pay-as-bid auction is used, however, the contracts will end up enhancing the generators’ market power.

Most of the papers in this area have concentrated on the case of a single dominant firm facing a competitive fringe, perhaps in each region. The next research challenge is to study what happens when we have more than one strategic player in the market – either there is a large company at each end of the line, or there are two large companies in one of the markets. The modelling in these cases would be more complex, but more realistic. How far can we get with equilibrium models, and when will we have to resort to simulation or experiments? The European market is fast moving towards one with a small number of large firms, each active in most of the main countries, and a fringe of smaller firms. Can we model the interactions between the larger oligopolists? Given that this is a repeated game and one involving multi-market contact, we might expect rather higher profits than if we studied each market in isolation. Are there steps that regulators or competition authorities could take in order to improve competition? Is there a critical point at which a company’s activities in one market will have an impact on other,
nearby, markets? If so, could this be used as a guide in European merger policy, which has typically been permissive towards cross-border mergers in the electricity industry?

7. REGULATING THE WIRES

If anyone had forgotten that electricity markets do not take place in some abstract space, but depend upon the transmission and distribution systems to deliver power, the recent blackouts in North America should have reminded them. The regulated sectors of the industry have received much less attention than electricity markets over the last few years – regulation is perhaps seen as traditional and boring, while markets are new and glamorous.

How should we regulate transmission and distribution companies to promote efficient operation and investment? Rate of return regulation might be prone to over-investment, while the second-guessing of ex post prudential review is unlikely to lead to the optimal outcome, either. The UK’s system of price caps was originally seen as something very different from rate of return regulation. The main difference, that prices were de-linked from the company’s actual costs, was only feasible because the price cap was intended as a short-term transition to competition. Once it was realised that many of the price caps were here to stay, a link to the company’s actual costs had to be established, even if only at the times when the price caps were re-set. There are differences in approach, of course, but I would argue that these are less than often suggested, particularly as incentives to keep costs down have been introduced into rate of return regulation.

How strong should the incentives for cost reduction be? How do the costs of a major failure compare with those of holding some slack inside the transmission and distribution operators? How much slack is desirable? Kay (2003) has recently pointed out that companies can almost always find some apparent slack to cut when seeking to reduce their costs, but that this frequently represents people who are important if quality is to be maintained in the long term. How far should the regulator attempt to determine the level of investment required? Can adjusting the regulatory asset base (on which the company is allowed a return) on a rolling basis combine incentives to trim costs with a mechanism that ensures the company will recover its investment over time? Can the threat of having to pay fines or compensate affected customers give companies sufficient incentive to maintain their quality of service?

Is there a way of replacing regulation with competition, at least for new investments in transmission? Financial Transmission Rights (FTRs), originally introduced as Transmission Congestion Contracts by Hogan (1992), pay the holder the difference in prices across a congested line. Bushnell and Stoft (1996) have shown that if the builder of a new line is given all the FTRs associated with that line, the value of their revenues should exactly equal the social value of that line, in a competitive market. The caveat about all the FTRs is important, because in an interconnected network, adding capacity in one place may reduce the amount of power that can safely be sent down another link. The builder of the new line should be made to bear the cost of this reduction in capacity. If the revenues associated with a line equal its social value, and free entry to line-building is permitted, can we rely on the market to build all the lines with social values greater than their costs? This is a potentially appealing prospect, but Joskow and Tirole (2003) give a number of examples in which moving away from the assumptions of the perfectly competitive model implies that the market signal for investment will be at an incorrect level. The challenge for research is to decide whether the imperfections in the merchant transmission model outweigh those of regulation, and to search for improvements to both.
8. CONCLUDING THOUGHTS

There has been a lot of research on electricity markets over the past twenty years, but there are still many interesting challenges for research. As a researcher, that is good news, but we should not forget that a major purpose of our research is to find ways in which the performance of the electricity industry can be improved, in terms of price, cost, reliability, and environmental impact. The research challenges that I have listed may well be associated with opportunities for improved performance that are not being taken up.

So I would like to sum up our task as one of finding ways of making electricity markets work better. If we think we’ve found one, we must check it carefully for flaws – the cost of getting things wrong in this industry can be very high. And if we still think we’ve found an improvement, we may find that the greatest challenge of all is getting it adopted in the real world.
REFERENCES


