DAE Working Paper WP 0233

UNIVERSITY OF CAMBRIDGE
Department of Applied Economics

Auctions and trading in energy markets : an economic analysis

David Newbery and Tanga McDaniel

The Cambridge-MIT Institute

Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research

CMI Working Paper 15

DAE Working Paper Series

UNIVERSITY OF
CAMBRIDGE
Department of
Applied Economics

not to be quoted without permisson

The
Cambridge-MIT
Institute

Massachusetts Institute of Technology
Center for Energy and
Environmental Policy Research

CMI Working Paper Series

Auctions and trading in energy markets - an economic analysis[1]

-

David Newbery and Tanga McDaniel
Department of Applied Economics, University of Cambridge
22 October 2002

Abstract

Auctions are playing an increasingly prominent role in the planning and operation of energy markets. Comparing the New Electricity Trading Arrangements to the former electricity Pool in England and Wales requires some analysis of the relative merits of uniform versus discriminatory pricing rules, and use of the gas network in Britain and electricity interconnectors around Europe is allocated on the basis of auction results. In this paper we discuss the changes in the trading arrangements in the electricity industry in England and Wales as well as some of the results to date. We also look at the wider issue of using auctions to replace regulation by market solutions for managing the natural monopolies in energy markets.

JEL: D44, L5, L94, L95
Key words: auctions, electricity, gas, interconnectors, networks, regulation

Auctions and energy markets

Auction design is at the heart of the evolving energy markets in Britain. In electricity, the New Electricity Trading Arrangements (NETA) replaced the former Electricity Pool on 27 March 2001. This represents the most important change in the design of the wholesale electricity market since privatisation in 1990. NETA also reflects a major change in the philosophy of market design, a subject that remains controversial. The Pool that NETA replaced operated as a single (or uniform) price daily auction that allowed the System Operator (SO) to match all demand and supply and determine the market-clearing price. In contrast, under NETA, the SO confines his actions to operating a residual balancing market, which is run as a pay-as-bid or discriminatory auction, and which produces two imbalance prices.

In gas, auctions have been introduced for allocating gas transmission capacity and gas storage, and have been proposed for guiding capacity investment decisions. Proposals were made for auctioning access to British electricity transmission capacity, but these have not been implemented. The French-English electricity interconnector is, however, auctioned, as are a number of electricity interconnectors on the Continent, including Netherlands-Germany, Netherlands-Belgium and Denmark-Germany.

---

[1] For CRI *Regulatory Review 2002/2003*. Support from the ESRC under the project R000 238563 *Efficient and sustainable regulation and competition in network industries* and under the CMI Electricity Project is gratefully acknowledged. We are indebted to John Bower for his data on capacity ownership, and to Gert Brunekreeft Richard Green, Alex Henney and Ralph Turvey for helpful comments.

This paper will briefly describe the reasons for changing the trading arrangements in the electricity industry, the nature of those changes, and the results to date. We also look at the wider issue of using auctions to replace regulation by market solutions for managing the natural monopolies in energy markets.

The reform of electricity trading arrangements

The reform of electricity trading arrangements was prompted by growing dissatisfaction with market manipulation in the Electricity Pool. In October 1997 the Minister for Science, Energy and Technology asked the Director General of Electricity Supply to consider how a review of electricity trading arrangements might be undertaken and to report results by July 1998. The timetable of that consultation is set out in Ofgem (2002), the deadline was met, the proposed reforms approved and set in motion with the planned date of "Go-live" set at 21 November, 2000. The Utilities Act 2000 introduced the necessary license changes on 8 August 2000, and on 14 August the Balancing and Settlement Code (BSC) came into effect. There were perhaps inevitable delays in ensuring that the IT systems were robust and reliable, and on 27 October the electricity and gas regulator, Ofgem, announced that "Go-live" date would be delayed until 27 March 2001.

The objectives of Ofgem's predecessor (the electricity regulator, Offer), which were approved by the Government, were to consider whether, and if so what, changes in the electricity arrangements would: best meet the needs of customers with respect to price, choice, quality and security of supply; enable demand to be met efficiently and economically; enable costs and risks to be reduced and shared efficiently, provide transparency; respond flexibly to changing circumstances; promote competition in electricity markets, facilitating entry and exit from such markets; avoid discrimination against particular energy sources; and be compatible with Government policies (Offer, 1998d, pp83-4).

The Pool also set up a Pool Review Steering Group to propose a set of objectives for these trading arrangements. They agreed the overall objective was 'that trading arrangements should deliver the lowest possible sustainable prices to all customers, for a supply that is reliable in both the short and long run' (Electricity Pool, 1998). They also listed a number of subsidiary objectives similar to Offer's, though they placed greater emphasis on the importance of minimising costs and ensuring longer run sustainability.

The workings of the Pool and its shortcomings are well described in the series of Offer papers produced in the course of the inquiry (Offer, 1998a-c). They have been discussed in two earlier issues of the *Regulatory Review*: Shuttleworth (1999) and Littlechild (2001). Shuttleworth, writing after the publication of Offer's *Interim Conclusion* (Offer, 1998d), noted that "it is difficult to find any rigorous analysis to underpin the reform proposals", while Newbery (1998c) concluded that "(T)he present review appears to have relied mainly upon unsubstantiated claims, inappropriate analogies, unquantified criticisms, and a remarkably uncritical assessment by the participants of the debate, without commissioning the kind of detailed analysis one might have expected from a regulatory agency claiming industry expertise."

Briefly, the distinctive features of the English Pool are that it was a compulsory, day-ahead, uniform last-price auction in which generators had firm rights to transmission but no firm obligations to generate. Dispatch was centrally organised by the Grid Company acting as System Operator (SO). Bids were submitted a day-ahead and were to be valid for every half-hour of the following day, though in practice generators could prescribe a large number of technical parameters to restrict the validity of bids at various times or under various conditions. The SO selected the least (financial) cost set of generating sets to run and issued operating instructions to these plants. The Pool set a System Marginal Price (SMP) each half hour as the computed unit cost of electricity from the most expensive unconstrained generation set called on to operate. Unconstrained generators were all paid the same SMP plus a capacity payment, equal to the Value of Lost Load (VOLL) *less* SMP, times the Loss of Load Probability (LoLP), which together made the Pool Purchase Price, or PPP. Generators available but not dispatched also received the capacity payments, while constrained generators received their bid price, if they were required to generate within an import-constrained zone and bid above the SMP, or their lost profit (SMP *less* bid price) if they were in an export constrained zone.

Although all electricity was dispatched through the Pool, more than 90% of electricity was sold under contract, and less than 10% actually received the Pool price. The typical contract was a Contract for Differences (CfD) for a fixed amount (M MWh, say) with an agreed strike price, f, say. If the Pool price was p, buyers paid $M(f\text{-}p)$ to the generator, and the generator collected Mp from the Pool for that electricity, thus receiving in total Mf. Under or overproduction relative to that contract would be paid the Pool price, so the Pool acted also as a balancing market. Actions taken by the SO to achieve overall balance and quality of supply would be charged out to all consumers and not targeted on those causing the imbalances or other problems (though the Grid Code specified quality requirements). Although the Pool only set the price for a small fraction of total output, that price was transparent, published a day ahead, and clearly influenced contract prices, as any contract party retained the option to buy or sell at the spot Pool price.

The enquiry that Offer conducted (the '*Pool Review*') recognised that the main reason for past market manipulation was the market power of the incumbent generators. The restructuring at privatisation created a duopoly of fossil-generators that set the price more than 90% of the time (Newbery, 1995). Subsequent substantial entry by 'Independent' Power Producers (IPPs) who built Combined Cycle Gas Turbine (CCGT) plant gradually eroded the market share of the incumbents, but almost all their output was sold on long-term contracts, usually to their part-owners, the Regional Electricity Companies, (RECs). The RECs retained the franchise to supply initially about 30% of total demand, but after 1994 just the domestic and small industrial and commercial market (half of the total), and were allowed to pass through their contract costs to these customers. A second complaint against the Pool was that these generators could bid zero but receive the Pool price (plus the CfD payments). They therefore appeared to offer no competition to the incumbents.[2]

Regulatory activism encouraged the incumbents to divest 6,000 MW of price-setting coal-fired plant in 1996. The incumbents succeeded in selling the plant for a fixed amount *plus* an 'earn-out' payment of £6/MWh of actual output, thus ensuring that prices remained high.[3]

[2] The criticism seems misplaced. Bidding zero guarantees dispatch and hence reduces the size of the market for the incumbents, who lose that market for the life of the entering plant. The incumbents therefore have a strong incentive to offer contracts at entry-deterring prices. A more telling criticism is that IPPs bid zero to ensure contract performance for passing through the prices into the franchise market. A yardstick purchase formula on the RECs for their franchise sales would have provided better incentives to secure power economically and encouraged the IPPs to bid their marginal cost.

[3] Justified as the implicit price of the sulphur emission permits associated with the plant, and as a risk-sharing device, defences that succeeded in passing regulatory scrutiny.

Subsequent plant sales continued to lower concentration in the important price-setting mid-merit part of the market, as shown in Figure 1 below. By the time NETA was introduced, the earn-outs had ended, and the wholesale market was arguably one of the least concentrated in the world.

The *Pool Review* criticised the rules for price determination in the Pool, which it argued allowed complex bids for small amounts of marginal plant to set prices for the whole market that bore little relation to production costs. In contrast to other commodity markets, the demand side had little role in setting the price, which emerged from a complex computer program designed for the old vertically-integrated Central Electricity Generating Board (and which had no use for half-hourly wholesale prices in any case). The fact that buyers and sellers could always trade in the spot market reduced the incentive to contract. Contracting not only forces buyers and sellers to bargain over the terms (and hence allows competition to work) but also, by pre-committing the price of the covered output, reduces the incentive to manipulate the Pool price, as manipulation only rewards the uncontracted output. Perhaps the most telling criticism was that the governance structure of the Pool made it almost impossible to make changes to rectify any of these perceived faults, as the Pooling and Settlement Agreement was an enforceable contract that could only be changed by mutual consent.

The proposed trading arrangements

The *Pool Review* argued that the complexities of price formation in the Pool allowed generators to exercise more market power than would have been possible had the market been structured more like a classic commodity market. Crucially, central dispatch was abandoned and generators were now free to arrange their own operation and had to find buyers for their output. The SO's task was to ensure system stability by balancing demand and supply. This was achieved by accepting offers and bids from generators or suppliers to increase or reduce output or demand. NETA replaced the Pooling and Settlement Agreement by a Balancing and Settlement Code with a well-defined method of making modifications. The Pool ceased to exist. Electricity was now to be traded in four voluntary, overlapping and interdependent markets operating over different time scales. Bilateral contract markets cover the medium and long run, while forward (and at some stage futures) markets offer standard contracts (base-load, peak hours) for periods up to several years ahead. A short-term bilateral market, operating from at least 24 hours to Gate Closure ($3^1/_2$ hours before a trading period),[4] allowed parties to adjust their portfolio of contracts to match their predicted physical positions. This short-term market would yield information to construct a spot price for each half-hour (e.g. the UKPX Reference Price Data).

At Gate Closure, the official end of the bilateral markets, all parties had to announce their Final Physical Notifications (FPN) to the System Operator (SO). The SO would then accept bids and offers in the last of these markets, the balancing market, to keep the system stable. Actions in this balancing market would be fed into the Balancing Mechanism to produce cash-out prices for clearing imbalances between traders' FPNs and their actual (metered) positions. This structure mirrors that emerging in the British gas market, though the Balancing Mechanism differs sharply from spot commodity markets, to which NETA aspired.

The key differences between NETA and the Pool

The most obvious difference between NETA and the Pool is that under the Pool generation was centrally dispatched while under NETA plant is self-dispatched. The obligation to balance output with demand is now placed on each generator, with the SO confined to ensuring system stability. The Pool, that acted as both a wholesale market for all electricity and allowed NGC as SO to balance the system, is replaced by a Balancing Mechanism (also operated by NGC as SO) for the residual imbalances that arise because of a failure or unwillingness of parties to self-balance. Whereas the Pool operated as a uniform single-price auction for buying and selling all power (including that needed for system balance), the Balancing Mechanism is run as a discriminatory (pay-as-bid) auction. NGC charges for balancing through the Balancing Services Use of System charge.[5] In addition there is an imbalance settlement process, operated by Elexon, "for charging participants whose notified contracted positions do not match their actual metered electricity production or consumption, and for clearing certain other costs of balancing the system." (BSC, 2002).

Parties may submit bids or offers to the Balancing Mechanism, specifying the price they wish to be paid (or pay) to move away from their FPN. A generator may thus *offer* to increase output (relative to his FPN), in which case if accepted he will be paid his offer price. He may *bid* to reduce generation and pay to do so (to avoid the cost generating that amount). A consumer may *offer* to reduce demand (and be paid in compensation) or *bid* (and pay) to increase demand. The SO then selects the bids and offers needed to balance the system and passes on details of bids and offers for Elexon to determine the cash-out prices for imbalances. This process of determining which bids are needed to balance the system and which remaining bids are to be used to determine the cash-out prices is both complex and to a considerable extent arbitrary, and this summary is therefore necessarily abbreviated.

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

One of the consequences of charging all imbalances (and not just those which contribute to net imbalance) is that Elexon may make a profit (or loss, if most parties are long and being paid to spill). Profits and losses are attributed to the "Residual Cash-flow Reallocation" (RCRC, or the so-called "beer fund") and paid (or charged back) to generators and suppliers in proportion to their output or demand.

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

---

[4] Reduced to one hour in July 2002.

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

[6] The prices are equal by about 25% of the time, and SSP>SBP very occasionally (0.1% of the time).

the average cost of securing the services is then charged out.[7] One consequence of this combination is that it is more risky for a generator to offer balancing services. If a generator has an accepted offer to increase output, and then suffers a loss of output, he is almost certain to have to pay more than he is paid. He may therefore prefer to retain the spinning reserve for his own insurance. A single final balancing price would make such an offer never any worse than self-insuring and normally better, and would thus promote a more liquid balancing market.

The theoretical case for change

The original intention for the Balancing Mechanism was to target the costs of balancing the system on those parties who cause the imbalance, so that they could take efficient decisions on whether to self-balance, or leave it to the SO to secure those balancing services more economically (and there are obvious economies of scale in providing these services at the system level). Imbalance is a system-wide phenomenon, in that the SO only needs to manage the net imbalance (either long or short) of the whole system.[8] If a party is long when the system is short, then that party is helping to reduce the imbalance, and should logically be rewarded for reducing balancing costs, not penalised. The cost-responsibility intention thus argues for a single cash-out price.

Ofgem subsequently and strongly defended the market-unfriendly dual cash-out Balancing Mechanism as a method of forcing all parties to contract well ahead of time, to avoid these penal charges. The long period over which deals can be struck, and the wide variety of possible counter-parties, makes contracting more competitive. The imbalance market is deliberately penal, with a low price for selling and a high price for buying, in order to encourage contracting, as shown in Figure 1. The figure smoothes a very volatile series by taking weekly moving averages of the 336 half-hourly prices. To give some idea of this volatility, the graph of the weekly average of the daily average price plus one standard deviation of the daily price is also shown, and the cash-out prices are compared with the weekly average spot price from the UKPX.

The lack of a single transparent price at which trades can always be made was also argued to reduce the incentive to collude in setting that price (by bidding in the Pool). Although the Pool price may only have applied to less than 10% of traded power, the Pool price was clearly salient in influencing the strike price for CfDs and hence the terms of the remaining 90% contracted.

[7] The Dutch balancing market is at the other extreme. It operates a uniform price auction to determine a single price for those 15-minute periods in which the system is either long or short for the whole period, and charges those who are short while rewarding those long. There is the potential (not yet used) to add a penalty of 1Euro/MWh to both imbalances. If the system is both short and long within the 15 minute period it determines two prices, effectively one for each sub-period in which the imbalance is in one direction.

[8] Transmission constraints may mean that imbalances have to be addressed within each constrained region, but the idea remains of (local) aggregate system balance.

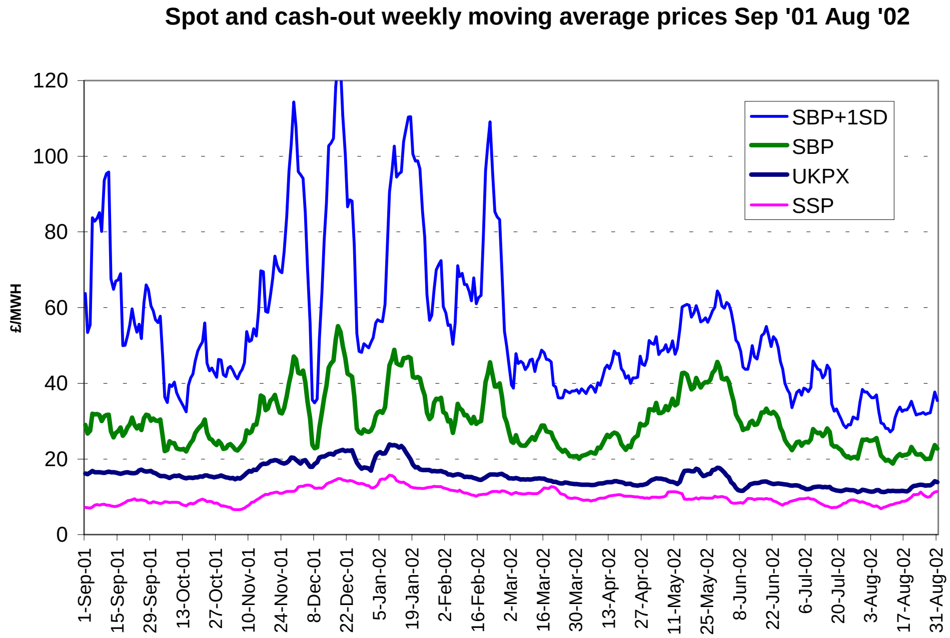

Figure 1 The range of system buy and sell prices

The debate over the market design for NETA was initially very confused. It was suggested (Offer, 1998e, 3.11-3.12, Offer, 1999) that if the single-price auction of the Pool were replaced with a pay-as-bid auction, average prices would be lower.[9] Some pointed to the revenue-equivalence theorem in auction theory, according to which under stringent assumptions (e.g. independent/private values, single-unit demands, risk neutrality and symmetric bidders) the average price would be the same under either design.[10] In practice those assumptions, particularly regarding symmetry and single-unit demands, are unlikely to be satisfied, making it an empirical or experimental issue as to which design is superior under what market conditions.

Green and McDaniel (1999) argued that a single marginal price auction was likely to be more efficient and required less burdensome information assumptions to work well than a pay-as-bid auction. One source of inefficiency is that the merit order may be compromised – lower-cost plant may not be dispatched before higher cost plant, particularly if the owner of the higher cost plant has a larger portfolio of plant and can predict more accurately the highest accepted price. We shall discuss other sources of inefficiency below.

Laboratory evidence has produced strong support in favour of uniform price auctions relative to discriminatory/pay-bid auctions within the framework of electricity markets. Rassenti, Smith and Wilson (2002) find less volatile prices with a discriminatory auction; yet, their discriminatory auction design without market power resulted in prices significantly above prices in the uniform design *with* market power! Abbink, Brandts and McDaniel (2002) report results from a laboratory experiment in which sellers are asymmetrically informed about market demand. They argue that informational asymmetries are characteristic

[9] "prices have been higher .. in part due to trading arrangements, .. a single Pool Price, which inhibits supply side pressure." (Offer, 1999, p3). "Any worthwhile reforms would require … the implementation of 'pay-as-bid' pricing." (ibid, p30).

[10] See Klemperer (1999) for a review of the revenue equivalence literature.

of decentralised electricity markets, such as England and Wales under NETA, and find that discriminatory remuneration leads to significantly less efficiency than uniform remuneration when information is not symmetrically distributed. Both Rassenti et. al (2002) and Abbink et. al (2002) use multi-unit demand auctions in their experiments (a relevant feature of utility industries). Bunn and Bower (2001) provide simulation results suggesting that pay-as-bid auctions would produce higher prices than a single marginal price auction.

The debate over market/auction design arguably paid too little attention to two important features of the electricity market. The first is that the substantial wholesale price volatility strongly encourages extensive contract coverage. Under the Pool over 90% was under contracts struck in bilateral deals that were clearly pay-as-bid. The issue then resolves into handling the residual spot market. Under NETA this is through screen trading the day before that produces a narrow spread, and then in the Balancing Mechanism. One argument against a marginal price auction for a very thin market with a very unpredictable steepness of the supply schedule is that the price may be very volatile. Paying-as-bid and averaging the cost then may reduce the volatility. The other solution is to increase the volume and liquidity of the residual market, and the Pool did that by requiring all generation to be bid in. An unresolved question is whether a single-price auction for balancing services would increase liquidity enough to reduce volatility, and lower the average cost of securing these services.

Under NETA, the amount of electricity handled by the Balancing Mechanism is only about 3% of the total, and the form of the auction is probably less important than the penal dual-price cash-out. If that were replaced by a single price targeted on net contributions to system imbalance (as in The Netherlands), then the total volume of imbalance trading (not the net imbalance) may increase, and the design issue could reassert itself as significant. Traded volume and liquidity would also increase with a single marginal price, as generators would be exposed to less risk bidding in to the balancing market. If they were then to go short, they would receive the same price for their offers as they had to pay in imbalance, while at present they risk making a loss unless they bid high.

The dual cash-out prices and the instability and unpredictability of the SBP have probably attracted the most concentrated criticism of all aspects of NETA. Two proposed modifications (P74 and P78) were considered by the BSC Panel in July 2002. Modification P74 proposed a single price, reflecting the net position of the whole system (a SBP when short, and a SSP when long), with the SSP < the spot price < SBP. The other, P78, nominally retains two cash-out prices, though the effect is rather like a single cash-out price. The SBP would be set at the market price when the system is long, and retain the SSP for those spilling power, while setting the SSP equal to the market price when the system is short, but charge those short the SBP. The BSC Panel advised Ofgem that the "alternative" versions of both modifications were preferable, but Ofgem rejected this advice on 9 September 2002 (and in particular the single cash-out price of P74) and instead accepted the original P78, to come into effect on 25 February 2003). While this reform may reduce the incentive to go long, it does not remove the discriminatory nature of the auction (including the different price paid to an offering generator and charged if he then goes short).[11]

One of the key questions is whether NETA would deliver lower prices than the Pool in the presence of market power, as its proponents argued (Currie, 2000). Rassenti et. al (2002)

[11] This particular discouragement to a more liquid balancing market could be overcome by guaranteeing that any offer that is matched by a call is cancelled financially.

provide some empirical evidence casting some doubt on this argument. We find results on the theoretical side as well, but they are far from decisive. Allaz and Vila (1993) consider a Cournot oligopoly model in which producers take the output decisions of their rivals as given. If they can contract ahead, then they can pre-commit some part of their output decision, and thereby influence the output decisions of their rivals, inducing them to reduce their output and helping raise price while ceding market share to the original producer. If all producers simultaneously sign contracts, on the assumption that they cannot influence other contracts, then collectively they commit to produce more output and final prices will be lower. Under strong assumptions on the game form, the only perfect equilibrium in a static Cournot game (in which producers can repeatedly offer additional contracts for the same future period) is the competitive equilibrium.

The problem with this argument is that these conditions are not satisfied in the electricity market. Producers meet daily in a continuing sequence of encounters, and the theory of repeated games is delicate and inconclusive. In addition, the Cournot assumption may be inappropriate (Green and Newbery, 1992), and there is ample evidence that the outcome can be above the competitive price (as was clearly the case for many years in the Pool). Contracting does reduce the incentive to exercise market power in the spot market (Newbery, 1998a), but market power can still influence the terms on which contracts are struck.

If generators have market power, and if they need not fear entry, then conditions in the electricity market would seem to favour generators. Consumers have a highly inelastic demand and are clearly willing to pay high prices rather than be cut off. The Value of Lost Load in the Pool was set at £2,000/MWh (1990 prices), based on some rather anecdotal evidence and a view of the desirable level of capacity payments. This figure was also used in the Victorian electricity market, but after several embarrassing power cuts, there were suggestions to raise the figure, perhaps by a factor of 10.[12] Generators, on the other hand, only forgo price *less* variable cost for the period they are not supplying, which, in competitive markets, may be very modest. Any generator that knows that without some output, remaining supply cannot meet demand at a reasonable price, can demand a high price for contracting (or supplying spot) in that period. It is difficult to see how increasing pressure on consumers to contract ahead of time (as under NETA) does much to reduce the market power the contract providers possess.

The critical proviso in this argument is that generators ignore the threat of entry. Newbery (1998a) argued that in a contestable market, oligopoly generators would offer contracts at an entry-deterring price (assuming they had enough capacity), to avoid the costly risk of permanently losing a fraction of their market. Under the Pool, and certainly before the domestic franchise was ended in 1999, the market was contestable. IPPs could sign 15-year contracts with RECs for power off-take, then sign 15-year gas contracts, borrow on the back of these contracts and finance CCGTs with performance guarantees, all with negligible risk. The Pool offered a guaranteed market and a clear marker price for contracts, and was

[12] In the Pool, the Loss of Load Probability (LoLP) was arguably over-estimated (Newbery, 1998b), so that the product of VOLL and LoLP may have produced acceptable results. In Australia, VOLL was only paid after load was lost, so generators had to make their own predictions of the LoLP that may have been more accurate.

apparently protected by an impressively strong Pooling and Settlement Agreement.[13] The result was that the market was contestable in the classic sense that the post-entry price could be locked in before entry in the contract, and the incumbents would lose that part of the market for 15+ years.

The real test of NETA is whether the market would be as contestable as under a marginal priced Pool, in the presence of market power. Newbery (1998c) argued that the increased risks of trading under NETA with penal imbalance charges would encourage vertical integration between generation and supply or retailing (as has happened on a wide scale). A vertically integrated industry with risky imbalance penalties would if anything be harder for IPPs to enter. Finally, and not directly connected with NETA, the ending of the domestic franchise has removed the natural counter-parties to the kind of long-term contract that makes entry contestable. One can imagine various combinations of policy that could hamper or facilitate contestability. A supply franchise with vertical integration would likely be worse, while a supply franchise and a prohibition on more than a certain fraction of that market served by own or affiliated generation (as in Britain pre-1999) might preserve contestability even with NETA.

Apart from the conditions of entry, it is far from clear that bilateral physical contracting combined with a discriminatory Balancing Mechanism mitigates market power more than bilateral financial contracting combined with a single price auction market of last resort. The practical question is whether, as Ofgem fears, a shift to a single price in the Balancing Mechanism would significantly discourage contracting ahead of time, and if so whether that would give rise to higher average prices for electricity. In order to address that question we need first to examine wholesale prices under NETA and compare them with costs and pre-NETA (Pool) evidence. Ofgem (2002) provides much of the relevant evidence in graphical form, so that presented here is much abbreviated.

The evolution of wholesale prices pre and post-NETA

NETA went live on 27 March 2001 after a delay of six months. The introduction of NETA was not the only dramatic change in the British generation industry. Between November 1999 and February 2000 the concentration of coal fired generation capacity (measured by the Herfindal Herschmann Index, or HHI, fell from 2644 to 1688. The standard US anti-trust measure of a concentrated industry is one with a HHI higher than 1800, and below 1300 the industry is considered unconcentrated. The concentration of coal plant continued to fall, and fell below 1300 in January 2002. The concentration of fossil plant (i.e. excluding nuclear) fell below 1800 in July 1999 and below 1300 in April 2000 – *before* NETA.

---

[13] In the even, the PSA was destroyed by legislation, and the Pool price in contracts ceased to exist, forcing contract renegotiation. Future IPPs may be more cautious when contemplating entry into an apparently liberalised market.

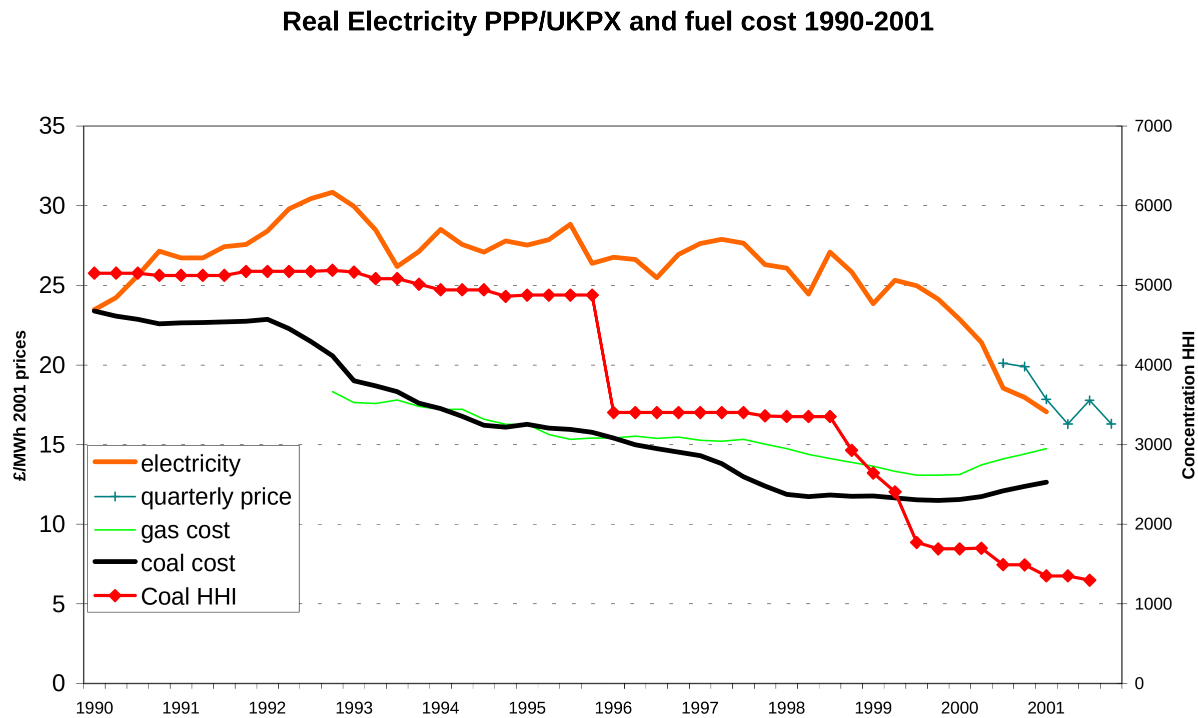

Figure 2 The effect of generation ownership on competition and wholesale prices

Figure 2 shows the evolution of coal capacity concentration and the real (2001 prices) wholesale price of electricity, taken as the Pool Purchase Price before March 2001 (and thus including capacity payments that now must be recovered in simple bids), and the average daily UKPX price thereafter. The graph of prices is a 12-month centred moving average to remove seasonality, extended with recent quarterly figures. They show a striking fall in the price shortly after the critical change in industrial concentration.[14] The crucial point to note is that the decline in prices preceded NETA 'go-live' and followed shortly after the rapid divestiture of plant in the latter part of 1999. John Bower (2002) (who supplied the plant data used in Figure 2) has already demonstrated this more rigorously using formal econometric tests. Further evidence of the effect of competition (and spare capacity, which was also increasing in this period) can be seen by comparing the average wholesale price for each half-hour settlement period in the winter months (Dec-Feb), when demand is at its seasonal peak. Figure 3 shows that prices were more volatile across the day in the winters of 1998/9 and 1999/2000 (rewarding flexible and mid-merit plant owned by the incumbents at the expense of base-load new CCGT entry). In the winter 2000/01, the last winter before NETA go-live, the time profile collapsed to the same real level as in the first winter after NETA (all prices are reflated to 2001 average values).Ofgem claimed that the fall in prices was in anticipation of NETA, but the prices shown are not contract prices, but spot prices for which arguably the winter of 2000/01 was the last chance to enjoy the opportunities to manipulate the Pool before it was replaced by NETA. It seems far more likely that the fall in prices was primarily driven by the fall in concentration. This may have been aided by the ending of the long-term

[14] The impact of the earlier fall in concentration on prices (in 1996 after Offer encouraged divestiture of 6,000 MW of coal-fired plant) was largely offset by an "earn-out" payment of £6/MWh to the incumbents that supported the prices bid by divested plant.

PPA contracts between the IPPs and the RECs, which the ending of the franchise, and the ending of the reference Pool strike price, forced to lapse.

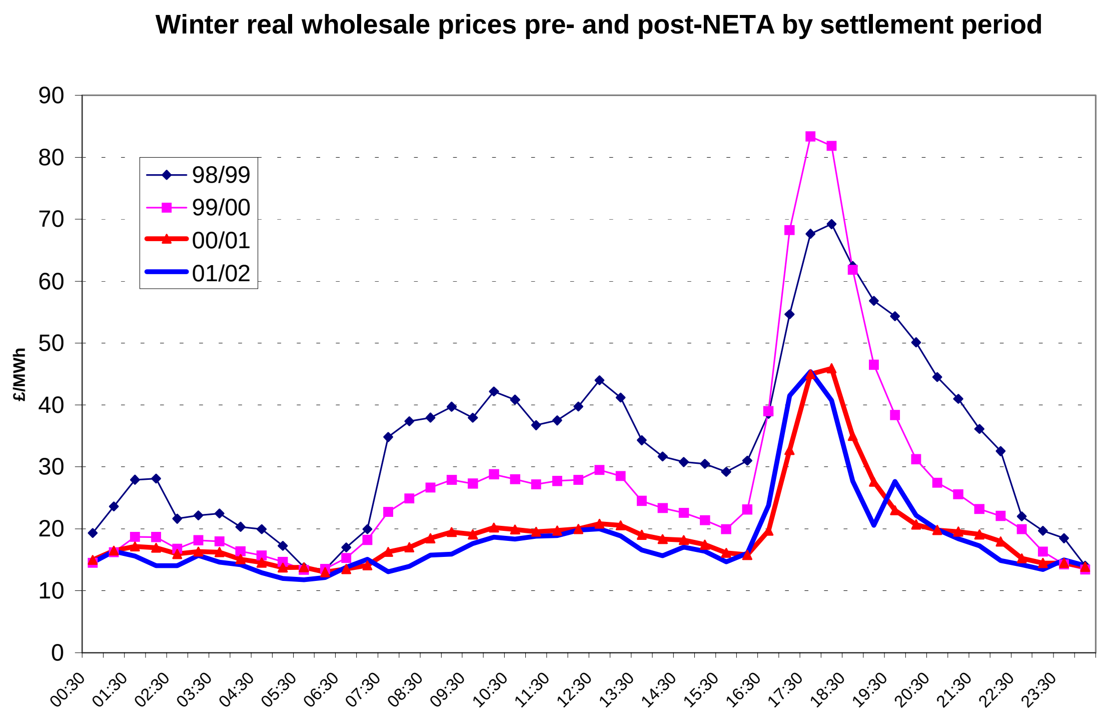

Figure 3 Average daily profile of prices in the Pool and UKPX, 2001 prices

Assessment of NETA

NETA has addressed one obvious shortcoming of the Pool – the difficulty of making socially desirable improvements that may adversely impact the profits of one of the parties. It may be that as a result some of the more obvious present problems can be addressed. One of these is the operation of the Balancing Mechanism, where the imbalance prices penalise each party individually. This encourages over-contracting (as shown in the Appendix). On average suppliers should be over-contracted by 80% of the Standard Deviation (SD) of their forecast demand, so that if their *individual* forecast uncertainty is 5%, they should be on average 4% over-contracted. The average cost of being over-contracted by the optimal amount (which varies by half-hour) from December 2001 to August 2002 was £3.32/MWh of SD, and the average base-load spot price over the same period was £15.13. Thus for a 5% forecast error the extra cost of over-contracting was 5% of £3.32 = £0.17 or 1.1% of the purchase cost. Compared to the fall in *prices* that is small, though not compared to the fall in costs.

Generators are likely to be able to forecast output more accurately, except for unexpected outages, and respond to the penal imbalances somewhat differently. One form of insurance against penal imbalance prices is holding plant part-loaded in case of generator outages. This is costly, as the thermal efficiency of a part-loaded 500 MW coal-fired plant falls from 35% to 32% (net calorific basis, Henney, 2002). This raises fuel costs by nearly 10% or perhaps £1/MWh. The estimate annual extra cost of running an additional 3,600 MW of part-loaded plant is 1.3 million tonnes of extra coal, or £40 million per year. The extra CO_2 emissions are estimated at 0.5-1 million tonnes carbon (MtC), a figure that can be compared with the target for the Renewables Obligations for 2010 of 2.5 MtC.

The Law of Large Numbers means that NGC is able to provide more accurate system-wide forecasts and hence would be able to balance the whole system more cheaply than individuals. Ofgem's *NETA Review* (Ofgem, 2002, p3) states that "NGC's gross balancing actions (buying and selling electricity) have been around 2% of demand," confirming that the incentives to balance are high and effective. One consequence is that NGC can secure reserves more cheaply as more plant is already running, so reserve costs are not properly allocated.

The costs of switching to NETA have been estimated at about £700 million (spread over a five year period) followed by annual costs of £30 million.[15] Some of the switching costs might have been needed to update old software even if the Pool had continued, but the extra costs of operating a trading desk 24 hours per day, 7 days per week, to deal with half-hourly balancing, are certainly substantial, and probably exceed £30 million/year. Whether these large costs are exceeded by *social* (rather than transfer) benefits must be doubtful, and is being investigated by the National Audit Office. Our pessimistic conclusion is that most and perhaps all of the price fall is attributable to increased competition, not NETA, and so the costs and disruption were unnecessary. In the longer run, entry may be more difficult and security of supply will not have been enhanced by removing capacity payments.

Development of auctions for managing natural monopolies

The current use of auctions for allocating access to the National Transmission System (NTS) in Britain replaced previous methods of negotiation and "grandfathering" with regulated access tariffs set at the long-run marginal cost of increasing capacity at that node - both of which were shown to be flawed as market conditions changed. Neither approach was adequate for allocating capacity rights in the face of real scarcity. McDaniel and Neuhoff (2002a) use the history of the British gas industry to support the claim that auctions are an appropriate means of allocating scarce network capacity given that there is competition in production as well as supply. Without competition in production expected prices will equal the reserve price (normally related to the previous regulated price), in which case auctions are unnecessary and costly. Without competition in supply, bidders can pass on uncompetitive auction prices to downstream consumers. In what follows we describe some of the important features of the industry and how they relate to the current short-term auctions.

The market players in the gas industry are producers, suppliers, and traders; these are not mutually exclusive so, for example, a producer may also supply gas to final consumers. Most producers using the NTS have gas fields in the North Sea and land their gas on the beach at one of the main entry terminals to the gas network. The production side of the gas industry has been competitive for some time, but it is only since 1995 that British Gas' control of the supply side of the market has fallen below 50 percent. In 2000 the supply arm of the former British Gas, Centrica, had a market share below 40 percent. The movement in the industry towards auctions appears to be in line with changes in the industry structure and concentration. The former methods of allocating rights by negotiation and grandfathering

[15] "The costs of implementing and operating the new trading arrangements are estimated to be between about £136m to £146m per annum, for a five year period. Thereafter the operating costs are expected to be of the order of £30m per annum." (Offer, 1999, p14).

were discriminatory by design, discouraged entry and were therefore incompatible with full liberalisation.[16]

There are six major gas terminals on the British coast: St. Fergus, Teeside, Theddlethorpe, Easington, Bacton and Barrow, plus a number of (small) on-shore sites and LNG storage facilities. The most utilised of the major terminals is St. Fergus in Scotland (where there are considerable transmission constraints) and Bacton in the south-east (where gas heading for the Zeebrugge interconnector enters). Gas from producers is landed at these beach terminals where it can be sold to traders or entered into the network by the producer. The value of entry rights is the expected difference between the spot price on the beach and the price where gas is traded downstream at the National Balancing Point (NBP). Presently entry capacity is auctioned concurrently by entry terminal and by month twice each year.[17] Auctions in March allocate rights at each terminal for the period April to September; auctions in September allocate rights from October to March. Entry rights are use-it-or-lose-it and unsold/undeclared rights are auctioned in subsequent daily and within-day auctions. Each auction has four rounds (plus a fifth 'residual' round with a slightly different dynamic), and in all cases the auctions are sealed-bid and pricing is pay-as-bid.[18]

The design favoured by the network owner and auctioneer, Transco, during the consultation process was a modified second price (i.e., Vickrey) open outcry auction that accommodates multi-unit demands. This particular auction is due to Ausubel and is explained in Ausubel (1997). The design ultimately implemented was a sealed-bid, multi-round option. It can be seen as a compromise given that it is more familiar and less complex than the Ausubel auction. This was partly to do with time restrictions and having the requisite systems in place before the September 1999 launch, and partly due to the generic unfamiliarity with the auction of choice. The auction design can also be seen as a compromise given the trade-off between discouraging collusion on the one hand (as would be more easily achieved in an open outcry design) and minimising the winner's curse on the other. As shippers' values are positively correlated, multiple rounds allow for learning, thereby minimising the chance of 'paying too much'. Although the general design of the auction was retained through subsequent semi-annual rounds, the details, especially the measurement of capacity available, have changed in every auction (particularly that for Winter 2001).[19]

---

[16] Gas storage has also been allocated using auctions since March 1999. We do not discuss the storage auctions in this chapter but details of background and an analysis of the March-April 1999 auctions can be found in Hawdon and Stevens (2001).

[17] *Concurrently* implies that all terminals and all months are auctioned at the same time.

[18] The rules of the auction can be found in the Network Code which can be downloaded from Transco's WebPages: http\\www.transco.co.uk.

[19] This is not the place for an extended discussion of the problem of defining capacity. Ofgem has argued throughout for a "more robust, independently verifiable measure of capacity …. linked to the maximum physical capacity available at each terminal" (Ofgem, 2001b). Maximum physical capacity is defined as the maximum volume of gas that can be injected at a terminal on the day with the highest demand that can be expected in 20 years, setting the injections at all other terminals to maximise capacity at the terminal in question. It is a theoretical measure of the maximum that could be made available. Until September 2002, the capacity made available was related to the Seasonal Normal Demand, presumably because the system is capable of actually delivering this. The obvious objection to that measure is that the system can frequently take considerably more, so the capacity is being artificially restricted. That may not matter if everyone knows that the remaining capacity actually available will be marketed on the day (ideally in a liquid spot capacity market).

There are a number of ways to measure the performance of the short-term auctions. One might be interested in the amount of revenue generated, the relative cost of alleviating network constraints (as measured by the cost to the network owner of buying back rights when demand for access exceeds supply), or the efficiency of the allocation. Jannsen and Moldovanu (2002) provide examples where various plausible auction objectives do not coincide. For example, if there is a market after the auction in which the bidders are competitors, then the bidders' valuations may be contingent on the identity of the other winners. In this case, bidder A's value may be lower if B is a winner than if C is a winner, and A may bid higher if bidding against B in order to prevent B from winning. The outcome may be socially inefficient even though bidders have maximised their individual valuations. This possibility is particularly relevant for utility auctions since bidders later compete in the downstream supply market.

When objectives cannot be simultaneously met then a choice between them should be made before the auction is designed. It is fairly easy to contend in the case of gas that efficiency should take precedence over revenue maximisation when the two do not coincide. This is arguably the case for any auction where the auctioneer is a regulated monopolist. Also secondary markets may not always resolve primary market misallocations. Jehiel and Moldovanu (1999) illustrate the shortcomings of secondary markets when one must rely on bilateral trading while, in a model of auctions for long-term capacity rights, McDaniel and Neuhoff (2002b) show that if there are costs for entrants then traders will reduce but not eliminate differences between socially optimal outcomes and auction outcomes.

McDaniel and Neuhoff (2002a) measure the performance of the short-term auctions according to two metrics: the closeness between monthly and daily auction prices at St. Fergus, and costs of constraint alleviation. First, if the monthly auctions are performing well then the prices observed should resemble daily auction prices. As a proxy for the daily auction price at St. Fergus, McDaniel and Neuhoff use a 30-day moving average of the difference between spot prices on the beach at St. Fergus and prices at the NBP. This proxy for daily prices is compared to the realised monthly auction prices between September 1999 and March 2002. They conclude that the auction price is a good predictor of the spot prices and take this as evidence that the auctions are working well according to this metric. Deviations between these prices can be traced to features external to (and thus not anticipated by) the auction, such as a post-auction increase in oil prices which increases the demand for interconnector access to the Continent and consequently the demand for domestic gas. When events after the auction cause the price at the NBP to rise there will be a gap between spot prices at St. Fergus and NBP prices when the network is constrained since gas landed at St. Fergus cannot enter the network.

Second, again using data on average spot prices at St. Fergus and the NBP, McDaniel and Neuhoff (2002a) show that the value of access rights increased by a factor of ten during the period September 1999 to October 2000 (from 0.7p/therm to 6.7p/therm) but that the cost of alleviating constraints fell over the same period. The increased transparency of the auction over previous allocation methods means that Transco has an objective measure for what the rights are worth to shippers; however, since capacity buy backs may not occur until the *gas-day* the price of buy-backs can still be above the prices paid in the auction.

McDaniel and Neuhoff confined their analysis to St. Fergus on the grounds that, at least for later auctions, the mark-up of the auction price above the reserve price is low at other terminals,[20] indicating either that there are few bidders and/or that there are no constraints.[21] Because there are only seven auctions to date and the rules have changed substantially between the early ones, it is difficult to make reliable statements about outcomes between auctions. However, trends *within* auctions can be observed and differences between terminals within auctions. It is possible, for example, to look for relationships between scarcity, the number of bidders and competition. Even if there are only a few bidders (who are unable to collude) the existence of network constraints should lead to competition among bidders. Conversely, in the absence of constraints, having many bidders may not ensure prices above the reserve. This latter point is important because, if the price always equals the reserve price, why have an auction? Auctions can be costly and time consuming for both the bidders and the auctioneer, and if there is lack of competition or lack of scarcity, then setting regulated tariffs might be a preferable alternative.

For some of the early auctions prices are grossly exaggerated relative to the reserve price at terminals where constraints are normally not a problem. This could either indicate a need for learning or be a result of privately held information and expectations about macroeconomic conditions. Figure 4 shows the mark-ups of the average auction price over the reserve price in each month at St. Fergus and Bacton for the last six auctions.[22] The two graphs show mark-ups for auctions 2-4 (left panel) and auctions 5-7 (right panel) separately to contrast the convergence of the reserve price at Bacton in later auctions to the sustained high mark-ups at St. Fergus. Expectations about macroeconomic conditions such as changing oil prices should especially affect bidders at Bacton since the price of gas on the Continent is linked to the price of oil, and rising Continental gas prices increases the demand for Bacton-Zeebrugge interconnector capacity. Such a situation may help to explain the extremely high mark-ups at Bacton in the March 2001 auctions. That is, if shippers were expecting the high oil prices of Summer 2000 to continue then they would anticipate continuing higher Continental gas prices and consequent scarce NTS entry capacity at Bacton. One disadvantage of auctioning capacity rights for six month intervals is therefore that uncertainty over future economic conditions means there will often be deviations between the monthly and daily access prices, in which case the monthly price will be a poor indicator of daily prices.

Figure 4 also shows the average number of bidders for each of the auctions.[23] The legend for the graphs show the terminal (F = Fergus and B = Bacton), the month (M = March, S = September) and the year (2000 or 2001). There are more shippers bidding for rights at St. Fergus than elsewhere on the network; the lines for St. Fergus are mostly to the right of Bacton in the left graph and strictly to the right of Bacton in the right graph.

---

[20] The reserve price is computed from the previous regulated price (based on long-run marginal cost of injections at the terminal) and adjusted for the degree of competition, with a lower reserve as competition increases (measured by the HHI).

[21] The extreme case is Barrow where there are at most two bidders, one of which is Centrica.

[22] The lines link the set of months in each auction, graphing mark-ups as a function of number of bidders, not the order in time of the months.

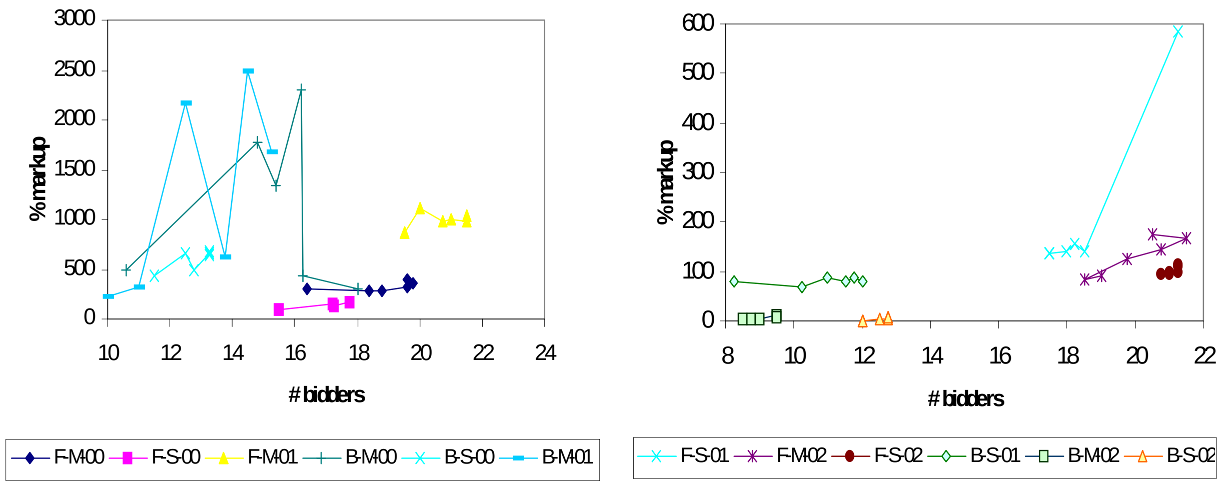

Figure 4: Mark-ups vs. number of bidders, St. Fergus and Bacton. Points give averages over rounds for each of the six months (linked together by lines).

For the early auctions there is much more variability in both the number of bidders and the mark-up at Bacton. The range of bidders is unusually large at Bacton in March 2001, ranging between 4 and 18 in some months over the four auction rounds. In all auctions except September 2001 there is little variation within auctions at St. Fergus in either the number of bidders or mark-ups. Another difference between the two terminals is that, unlike at Bacton, there is an obvious (though subtle) positive trend between the number of bidders and the percentage mark-up within auctions at St. Fergus. This is reassuring since it suggests that an increase in the number of bidders can increase competition when there are constraints (as is the usual case at St. Fergus). At Bacton the trend between bidders and mark-ups is less obvious and much more erratic within auctions. What is interesting is that although the number of bidders does not change substantially between early and later auctions, there is a convergence to the reserve price. Again this is reassuring as it supports the claim that when there is no scarcity the benefits of auctions relative to regulated tariffs is unclear.

Bacton stands in the middle of two extremes for when Continental gas prices are higher than in Britain there will be high demand for entry at Bacton. Thus, rationing may indeed be required at those times. When Continental gas prices are equal to or lower than British prices demand will fall, possibly to zero. The opposite extreme is no competition at no scarcity, which is illustrated by Barrow in the south-west that is dominated by a single bidder and where the auction always clears at the reserve price.

Longer term auction to guide investment

McDaniel and Neuhoff (2002a) argue that the monthly auctions appear to be successful with respect to: anticipating daily auction prices, reducing the cost of alleviating constraints, and capturing producer rents (as indicated by the high revenues obtained in the auctions). Ofgem now wants to take the auctions a step further by allowing future investment in the network to be guided by auctions of long-term entry rights and funded by the auction revenues obtained.

[23] This is the average of total bidders, not total successful bidders.

This is a radical step and it does not follow that this is an obvious extension of the current monthly auction. The idea that auctions might provide for the correct amount of network investment overlooks a number of important market failures and fails to fully consider the strategic objectives of bidders who already use the auction and the potential impact on future entry. This use of auctions is still quite controversial and there is still ambiguity about the auction design (as well as the proper measure of capacity).[24]

One feature of the long-term auctions is that some portion of capacity rights will be reserved for short-term markets (e.g., yearly, monthly, daily, with-in day).[25] This portion might either be a fixed fraction of existing capacity or it might vary depending on the outcome in the long-term auction. Reserving some rights for short-term markets can improve competition by making it easier for traders and entrants to obtain rights. McDaniel and Neuhoff (2002b) build a simple model that captures the salient features of the potential designs cited in Ofgem (2001a). Their paper considers the division of rights between short and long-term auctions and suggests that market power or shippers' private information about production plans can distort the auction away from the socially optimal outcome. The magnitude of this distortion depends on the proportion of rights allocated to short-term markets. The auctioneer must therefore choose between different objectives: (i) removing distortions arising from market power, (ii) extracting private information about production plans, or (iii) removing distortions created by network effects; i.e., the fact that shippers land their gas at different points on the network.

While the use of auctions for allocating short-term scarce capacity on the British network has proven to be an improvement relative to the previous methods of selective negotiation and grandfathering with regulated tariffs this is not sufficient justification for using them as a guide for future investment. Long-term auction outcomes can be highly sensitive to the auction parameters and there will be conflicts between the myriad of objectives the auction is meant to achieve. It is not clear that such auctions will provide for a more efficient network and allocation of access rights relative to regulatory approved capacity expansion.

Auctions for electricity capacity

Another rapidly developing opportunity for using auctions in the electricity market is for access to inter-connectors. Inter-connector auctions into The Netherlands came into effect on 1 January 2000. The England-France inter-connector capacity has been competitively allocated since 1 April 2001. It is interesting to compare and contrast these two different inter-connector auctions and discuss their role in the integration of European electricity markets.

One of the key issues in inter-connector auction design is whether the capacity to be sold is to be defined on the net or gross flows. At present, inter-connectors to The Netherlands and to England are only sold for gross flows – that is, the entire capacity

[24] Ofgem have just recommended that annual auctions should begin in January 2003, Ofgem (2002).

[25] We use "long-term" to refer to auctions for capacity rights more than three years in the future while "short-term" usually refers to capacity rights up to six months in the future (as with the current gas entry capacity auctions). Where long-term auctions exists, short-term might then refer to any period in which the capacity which is being bid for already exists or is already being built; e.g., auctions for rights less than three years in advance.

available in either direction is sold separately.[26] An alternative would be to algebraically sum bids for import and export capacity (where exports would be defined as negative imports) and allocate the total capacity to the net flow. This principle of superposition would mean that if traders wished to export 500MW from The Netherlands to Germany, when the import capacity into The Netherlands were 1000MW, then the auction would allocate capacity of 1500MW for available imports. The auction price for imports would be paid by importers and paid to exporters. For this to work, exporters would have an obligation to export (in practice to net off against imports) in order that their production decision (in The Netherlands) and consumption decision (in Germany) could be balanced against demand in The Netherlands and supply in Germany. That way, the actual flow of electricity on the inter-connector would indeed be the algebraic difference between imports and exports. Failure to export would have to be subject to imbalance charges as import commitments could no longer be honoured and would need to be compensated.

The most active auction into The Netherlands is from Germany, where two grid connections are separately auctioned. Capacity is auctioned for yearly capacity (i.e. for base-load), monthly capacity and, on the day ahead, for each hour of the following day. Capacity obtained is subject to a "use it or lose it" principle so that any capacity that is not nominated to the system operators before a certain time is released and sold on the daily auction. The auctions are uniform price and the lowest accepted price determines the price for all capacity in that auction. The two German-Dutch auctions (from RWE and Eon to TenneT) clear simultaneously but independently, so prices normally differ, though on average are similar.

Any import capacity secured must be bid through the Dutch spot market (APX). Traders can bid for hourly capacity in the day-ahead inter-connector auction, then buy electricity in the German spot market (LPX or EEX), and sell electricity into the APX. Normally, prices in Germany are below those in The Netherlands, and as a result the inter-connectors are fully used and the auction price clears at a positive level. Figure 5 shows the average of the cost of bidding for base-load capacity on the two German-Dutch interconnectors in the month-ahead and day-ahead auctions. This can be compared with the average price difference between the two markets for those hours where the price difference was positive (and hence worth using the interconnector).

The yearly and monthly auctions appear efficient, in that the 2001 yearly auction cleared at close to the price difference of year-ahead base-load OTC contracts in the two different markets, as did the monthly market. The daily market is not so well arbitraged, and the daily inter-connect auction clears on average at a lower price than the spot market price difference. This may reflect the risk that the hourly spot prices are not known at the time of bidding for inter-connector capacity, in contrast to the yearly and monthly auctions where there are suitable financial contracts to guide price discovery.

[26] System security issues may mean that the capacity in each direction that is reliably available may differ.

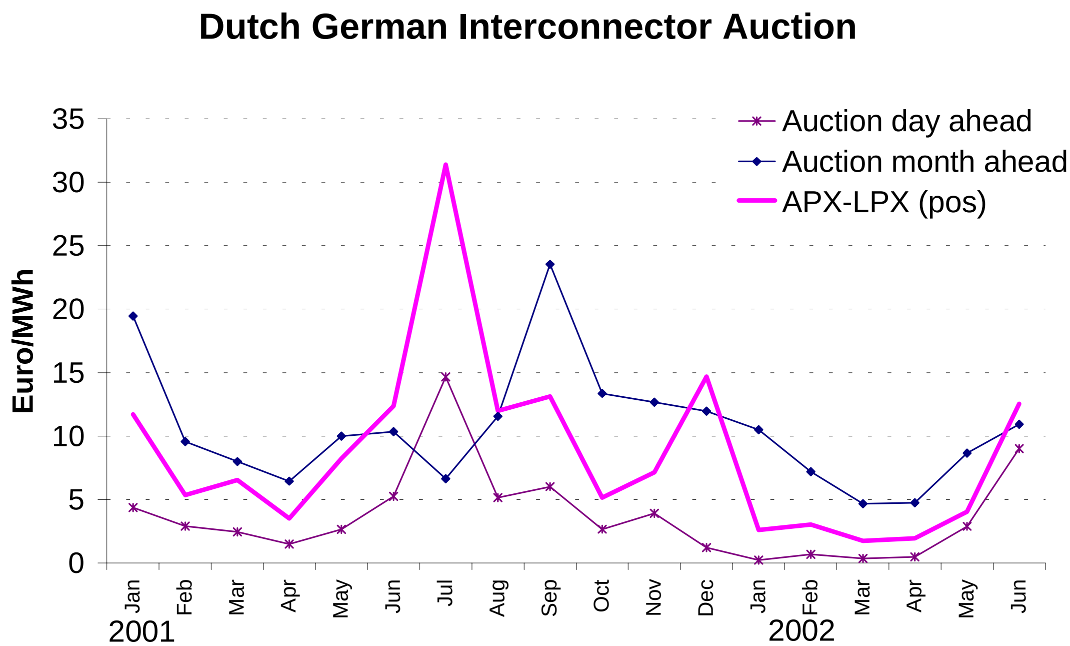

Figure 5 Comparison of costs and returns to buying interconnection

The England-France inter-connector has a rather different design. As in The Netherlands, capacity in each direction is sold separately, so there is no superposition of flows. Since the Powernext market in France provided a spot market price, the prices have been higher in France in 42% of the hours (Dec-01 to Aug-02).[27] At some periods traders may wish to both export and import and in such hours the flows are in fact netted, but the inter-connector is then under-utilised.

The capacity is sold under three-year contracts through sealed bid tenders, one-year and quarterly contracts through pay-as-bid auctions, and daily capacity also in pay-as-bid auctions. Again, the capacity is subject to the use-it-or-lose-it principle, and is then resold on the daily auction. The lack of liquidity in the spot market Powernext in France appears to restrict the number of potential bidders (normally to 2-3), and may justify a pay-as-bid auction. The spread of prices is normally quite small, with the marginal bid varying between 101-125% of the weighted average bid. Figure 6 graphs the most recent annual and quarterly weighted average bids of capacity from France to England (expressed per MWh, (dividing the effective daily cost of base-load by 24), and the excess of the UK spot price over the hourly Powernext price (again for hours when this is positive). The average (positive) price difference (Dec-01 to Aug-02) is 5.07 Euro/MWh, while the cost of annual auction for calendar year 2002 was 4.14 Euro/MWh. The average of the quarterly auctions was 2.46 Euro/MWh, suggesting either risk aversion, high transaction costs or illiquidity as reasons for the imperfect arbitrage.

[27] Interestingly, French prices are higher than English prices more than half the time in hours 7-10 and 15-17 (French time).

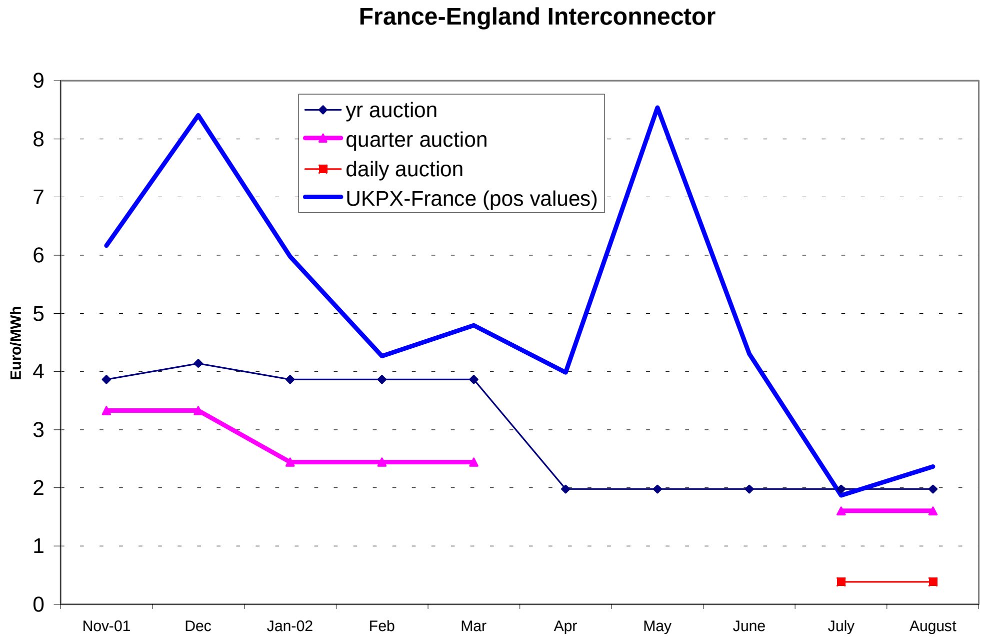

Figure 6 Cost and profit of trading from France to England

The most obvious reform to improve the efficiency of the interconnector auctions would be integrate spot and auction markets so that they clear simultaneously on net flows. This would in effect require or create a single EU-wide power exchange, and would be resisted by individual power exchanges. It would have the additional and possibly decisive advantage of reducing market power in local markets. At present, interconnect capacity is bought ahead of spot electricity, and creates either inelastic demand or supply, which enhances market power. If demand for interconnection were jointly determined with local supply and demand, the effective demand and supply schedules would be more elastic.

Conclusions

Auctions are increasingly used to determine prices in gas and electricity markets. In electricity, they are used for pricing scarce inter-connector capacity, and for determining the spot price, while in gas they are used to price entry into the high pressure gas system, and for allocating storage. In all these cases, auctions determine the price and allocation of currently available resources, whether it is existing transport capacity or the output of existing production facilities. Auctions have been used for commodity spot markets probably for as long as producers and traders have met to exchange, but the idea of using auctions to price natural monopoly facilities is relatively new and attractive to those who consider regulation a costly and imperfect way of setting prices and allocating resources.

The temptation is to believe that the price discovery role of auctions can be used to guide investment decisions in long-lived natural monopoly elements. Clearly the price signals emerging from spot auctions are informative, but there are risks involved in relying solely on auctions to determine future capacity, at least without clearly specifying how and when the regulator may step in to ensure adequate capacity if the auctions deliver apparently inadequate investment. There are then dangers that investors will fear future excess capacity (which has social and external benefits in increasing reliability and reducing market power) and will thus underbid for new investment, validating the need for additional investment and low prices.

As for electricity trading arrangements, Britain has abandoned the Pool model that has been so influential in other electricity reforms, and has adopted a market structure that has more similarity to Continental electricity markets, though these still prefer non-discriminatory or last-price auctions, even for balancing. Wholesale prices have fallen dramatically. It has not yet clear whether this is because of a massive increase in the competitiveness of the wholesale market just predating NETA, or the end of the domestic franchise, or the removal of the contractual basis for long-term power purchase agreements by ending the Pool, or as a result of NETA itself. Sceptics argue that the price fall is not due to NETA and would have happened anyway in a competitive Pool, at much lower cost, with a more sustainably contestable market providing better supply security. No doubt future issues of the *Review* will be better placed to allocate praise and blame for the final convergence of British prices to Continental levels.

References

Abbink, K., J. Brandts and T. McDaniel (2002), "Asymmetric Demand Information in Uniform and Discriminatory Call Auctions: an experimental analysis motivated by electricity markets," forthcoming in *Journal of Regulatory Economics*.

Allaz, B. and J-L Vila (1993) `Cournot competition, futures markets and efficiency', *Journal of Economic Theory*, 59(1), Feb. 1-16

Ausubel, L. (1997) "An Efficient Ascending-Bid Auction for Multiple Objects," University of Maryland (revision, 7 August 2002). On-line at: http://www.ausubel.com/auction-papers.htm

Bunn, D. and J. Bower (2001) "Experimental Analysis of the Efficiency of uniform-price versus Discriminatory Auctions in the England and Wales Electricity Market," *Journal of Economic Dynamics and Control*, vol. 25:3-4, 561-592.

Bower, J. (2002) 'Why Did Electricity Prices Fall in England and Wales?: Market Mechanism of Market Structure?', Oxford Institute for Energy Studies Working Paper, EL O2, Version 1.1. (September 2002). ISBN 1 901 795 22 5. Published by OIES, Oxford. http://www.oxfordenergy.org.

BSC Panel (2002) *Annual BSC Report* available at http://www.elexon.co.uk/elexon/about_elexon/bus_plan.html

Currie, D. (2000) "The New Electricity Trading Arrangements in England & Wales: A Review", The Beesley Lectures on Regulation, The Institute of Economic Affairs and the London Business School, 10 October 2000.

Electricity Pool (1998) *Pool Review Steering Group Response to Offer's Interim Conclusions of June 1998*, London: Electricity Pool

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 T. McDaniel (1999) 'Modeling the RETA: a model of forward trading and the balancing mechanism', DAE Working Paper, 0002

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

Hawdon, D. and N. Stevens (2001), "Regulatory Reform of the UK Gas Market: the case of the storage auction," *Fiscal Studies*, vol. 22, no 2, p. 217-232.Jannsen, Maarten and Benny Moldovanu (2002), "Interaction between allocation mechanisms and future market outcomes," in *Auctions and Beauty Contests: a policy perspective* SEOR-Erasmus Competition and Regulation Institute, Rotterdam.

Jehiel, Philippe and Benny Moldovanu (1999), "Resale Markets and the Assignment of Property Rights," *Review of Economic Studies*, 66, 971-991.

Klemperer, P. (1999), "Auction Theory: a guide to the literature," *Journal of Economic Surveys*, Vol. 13(3), 227-286.

Littlechild, S. (2001) `Electricity Regulation', ch. 2, pp 25-62 in *Regulatory Review 2000/2001*, ed. P. Vass, CRI, University of Bath

McDaniel, Tanga (2002), "Auctioning Access to Networks: evidence and expectations," mimeo, Department of Applied Economics, University of Cambridge.

McDaniel, Tanga and Karsten Neuhoff (2002a), "Auctions to Gas Transmission Access: the British Experience," in *Auctions and Beauty Contests: a policy perspective* SEOR-Erasmus Competition and Regulation Institute, Rotterdam.

McDaniel, Tanga and Karsten Neuhoff (2002b), "Entry Capacity Auctions and Investment Signals," *Cambridge-MIT EP-04* and *Department of Applied Economics Working Paper 0213*, 2002.

Newbery, D.M. (1998a) `Competition, Contracts and Entry in the Electricity Spot Market', *RAND Journal of Economics*, 29(4), Winter, 726-49. (DAE Working Paper 9707)

Newbery, D.M., (1998b) 'Pool Reform and Competition in Electricity', in M. Beesley (ed.) *Regulating Utilities: Understanding the Issues*, London: Institute of Economic Affairs.

Newbery, D. M. (1998c) `The Regulator's Review of the English Electricity Pool', *Utilities Policy*, **7**(3), 129-41

Newbery, D.M. and M.G. Pollitt (1997), 'The Restructuring and Privatisation of the CEGB- Was it worth it', *Journal of Industrial Economics*, XLV (3), 269-303

Offer (1998a) *Review of Electricity Trading Arrangements: Background paper 1 - Electricity Trading Arrangements in England and Wales*, February, Birmingham: Office of Electricity Regulation

Offer (1998b) *Review of Electricity Trading Arrangements: Background Paper 2 - Electricity Trading Arrangements in other Countries*, February, Birmingham: Office of Electricity Regulation

Offer (1998c) *Review of Electricity Trading Arrangements: Working Paper on Trading Inside and Outside the Pool*, March, Birmingham: Office of Electricity Regulation

Offer (1998d) *Review of Electricity Trading Arrangements - Interim Conclusions*, June, Birmingham: Office of Electricity Regulation

Offer (1998e) *Review of Electricity Trading Arrangements: Proposals*, July, Birmingham: Office of Electricity Regulation

Offer (1998f) *Report on Pool Price Increases in Winter 1997/98*, June, Birmingham: Office of Electricity Regulation

Offer (1999) *The new electricity trading arrangements* (Volume 1), July, Birmingham: Office of Electricity Regulation

Ofgem (2001a) *Long term signals and incentives for investment in transmission capacity on Transco's National Transmission System - The New Regime*, March.

Ofgem (2001b) *Ofgem decision – Winter 2001 Capacity Auction modifications*, letter to shippers, etc of 9 Aug 2001

Ofgem (2002) "Ofgem's view on long-term entry capacity auctions: modification proposals 0500 and 0508," 1 October 2002.

Rassenti, S., V. Smith, and B. Wilson (2002), "Discriminatory Price Auctions in Electricity Markets: low volatility at the expense of high price levels," forthcoming in *Journal of Regulatory Economics*.

Shuttleworth, G. (1999) 'The Electricity Industry 1997-98', ch. 2, pp 19-34 in *Regulatory Review 1998/9*, ed. P. Vass, CRI, University of Bath

Stoft, S. (1997), 'What should a power market want', *Electricity Journal*, 10(5), 34-45.

Appendix: The optimal degree of over-contracting by a supplier

Suppose that forecast demand is a random variable $Q = Q(1+\theta)$, where θ is distributed as $F(\theta)$. The amount contracted will be $K = (1+x\sigma)Q$, where σ is the coefficient of variation (and SD) of θ. The selling price of Q is P, the contract price for purchasing wholesale is p, and the system buy price (SBP) for imbalances is b, the SSP is s. The expected profit is then Π, where:

$$\Pi = PQ - Kp + s\int_{-\infty}^{x\sigma}(K - Q)dF(\theta) - b\int_{x\sigma}^{\infty}(Q - K)dF(\theta),$$

or

$$\frac{\Pi}{Q} = P - (1 + x\sigma)p - s\int_{-\infty}^{x\sigma}(\theta - x\sigma)dF(\theta) - b\int_{x\sigma}^{\infty}(\theta - x\sigma)dF(\theta).$$

The optimal degree of over-contracting is then x, which maximises Π:

$$\frac{1}{\sigma Q}\frac{\partial \Pi}{\partial x} = -p + s\int_{-\infty}^{x\sigma}dF(\theta) + b\int_{x\sigma}^{\infty}dF(\theta) = -p + sF(x\sigma) + b(1 - F(x\sigma)),$$

so the value of x solves

$$F(x\sigma) = \frac{b/p - 1}{b/p - s/p}.$$

Given values for b/p and s/p and the probability distribution of θ (which to a first approximation can be taken as normal), the optimal degree of over-contracting x can be solved. Thus taking the data from 3 Dec 2001 to 15 June 2002, the average for the same half-hour each day of $F(x\sigma)$ is 0.79 with a range from 0.54 to 0.94 and a SD of 0.09. If F is normal, then the supplier should on average be 0.8σ over-contracted, ranging from 0.1σ in HH25 (1230-1300) to 1.56σ in HH32 (1600-1630).

A slightly more sophisticated calculation takes the average value for each half-hour period of x=Min(0, F^{-1}(Max[0.99, Min{(b/p-1)/(b/p-s/p)}])).[28] This produces almost exactly the same answer of x ranging from 0.3 to 1.3 with an average value of 0.83.

[28] The constraints are to rule out infeasible or implausible numbers (i.e. under-contracting).