POLICY PAPER

Market reform: rationale, options and implementation

Dieter Helm
Professor of Energy Policy,
University of Oxford

8th October 2010

1. Introduction

As a result of Ofgem's Project Discovery and the DECC Market Assessment, and with the Climate Change Committee's Fourth Climate Change Budget due in December 2010, the government is embarked on a market reform process, first through a consultation paper and then through legislation. This short paper summarises the rationale for electricity market reform, considers three related options, and sets out the case for a long-term capacity auction mechanism.

2. The context

Electricity is hard to store and therefore requires a margin of excess capacity and over-build of networks to meet volatility in demand. Customers buy both energy and the insurance that when they turn on the light switch, power will be provided. For most electricity systems for most of the twentieth century, in consequence, electricity has been priced through a two-part tariff: an energy price and a capacity price.

In Britain, this two-part approach was embedded in the CEGB's Bulk Supply Tariff under nationalisation, and carried over in privatisation to an energy price (the system short run marginal cost as the system margin price (SMP)) plus a short-term capacity price (the administered Value of Loss of Load (VOLL) x the short run Loss of Load Probability (LOLP)). The SMP component was determined in a compulsory pool – all power had to be sold into it, and anyone could buy at the pool price. It was therefore transparent and liquid (important features to which we return below).

The capacity element proved not only open to manipulation by the then dominant duopoly, but also relied on the idea that what was in effect a spot price for capacity now would provide a long-term appropriate signal for investors who would be rewarded through this mechanism some years in the future.

For reasons that had much to do with lobbying by the dominant incumbents, the pool was pushed over in favour of a single unified voluntary market. Generators no longer had to set contracts in relation to the pool as a benchmark, and in practice could pluralise contracts as they saw fit. This voluntary market facilitated vertical integration, and each of the major companies sort to increasingly physically match their generation with their customer base. As a result, liquidity and transparency were reduced, and entry by merchant generators became a rare event.

In the process the much-criticised capacity element was abandoned – the single price in the new market (NETA – which became BETTA when England and Wales were integrated with Scotland) was assumed to incorporate the capacity costs. In part this was simply a reflection of the fact that the market remained in the excess supply hanging over from the early 1980s (and reinforced by the 1990s dash-for-gas, underpinned by the supply monopoly which endured until 1998). If the market was in excess supply, security of supply took a back seat – and in effect was ignored.

Ignoring the investment incentives meant that the problem of long term fixed and sunk costs could be parked too. For at least three of the key technologies that might be required for future generation – nuclear, renewables, CCS coal and CCS gas – the gap between marginal and average costs is very considerable, and this gap is partly made up of sunk costs. In a competitive market, the efficient allocation of risk dictates that investors and consumers would share this risk through long-term contracts, and that these long-term contracts would have particular regard to the sunk costs. A key feature of NETA/BETTA is the absence of such contracts, and indeed the dominant incumbents have sort to solve this problem through internalisation through vertical integration.

3. Climate change and security of supply

By 2009, it was increasingly apparent that the long years of excess supply were coming to an end, and that the incumbents lacked sufficient incentives to invest in the capacity margin. Ofgem – wrongly as it has turned out – considered the problem to be one of the choice of technology. In Project Discovery, Ofgem set out a case for market reform based upon a fear that the market would choose to build too many gas CCGTs at a time when North sea gas production was running down, and Britain was becoming reliant on imports from more distant sources, like Russia and the Caspian. Ofgem did not see a problem with the level of investment, and its timing on gas was spectacularly wrong – for just as Ofgem had woken up to the problem, shale gas technologies had created a very different position, with the prospect of plentiful, geographically widely-distributed and cheap gas for decades to come. If anything, the Ofgem analysis could be turned on its head: gas now provides an opportunity – at very low cost and a small fraction of the cost of offshore wind – to rapidly reduce emission by replacing coal generation as a transition strategy to a lower carbon world.

DECC's concerns were rather different, and at the heart of its concerns were both that the market might not invest enough and in a timely fashion, and that it might not build sufficient renewables to meet the EU Directive targets. The security of supply concerns had been going up the political agenda, most notably through John Hutton's Energy Review in 2007. This had been part of the rationale for the U-turn on nuclear which he effected, and also the push for building a number of new coal fired power stations, most of which (75%) would be unconstrained by CCS for some time.

It is hard to argue with DECC's concerns that the market would not deliver as currently constructed. NETA provides no incentive for the generators to build the excess capacity that security of supply requires: indeed no rational capitalist would deliberately create excess supply. Furthermore, since vertical integration effectively killed off new entrants, the profit maximising strategy for the incumbents would be to keep the market tight in the presence of an element of market power. The tighter the demand-supply balance, and in the absence of long-term contracts, then the higher the NETA price, increasing the returns to existing power stations. Provided that politicians and regulators allowed the price to spike, demand would always equal supply. Security of supply therefore showed up in the price rather than the quantity – and the result could be a profitable one.

On climate change, the problem facing the government is that the chosen technology to meet the renewables directive at the margin is extremely expensive. Offshore wind is one of the few technologies that makes even nuclear look cheap. Thus the subsidies have had to be piled up. Currently offshore wind gets two ROCs, preferential treatment on grid investments, the benefits of the CCL and a host of other subsidies and supports. Even with this, and the prospect of a more carbon focussed CCL, and the EUETS weighing down on fossil fuels, it is widely acknowledged that this is not likely to be sufficient to meet the target. (Curiously, the one contrary argument is Chris Huhne's claim that we face a doubling of oil prices, and his related assumption that gas prices will remain hitched to the oil price. If this were true, then renewables would not need much subsidy at all. There is however little reason to believe this claim: the oil market has considerable supply flexibility, gas prices are decoupling from oil, and the world is now awash with both natural and shale gas).

Assuming that the meeting the target is a binding constraint, DECC and the CCC have been considering how the market might be reformed to ensure that the trajectory of new investment is consistent with the overall carbon targets, the carbon budgets and the renewables directive. In effect, the climate change rationale for market reform is to force the market to choose a particular technology profile for investment – in the absence of the first best policy of setting an appropriate carbon price without picking winners.

Thus the rationale for market reform is to achieve two targets: security of supply through a capacity margin (excess supply) and a technology profile. It is about both the level and composition of generation.

4. Three mechanisms

There are at least three broad ways in which market reform might meet these targets:

4.1 A change to the supplier licence obligations

The supplier licence could be changed to mandate suppliers to contract for a specified amount of capacity and for a composition of that capacity. In the simplest model, the government could take powers to instruct the suppliers on their contracting behaviour as and when it sees fit. This is, in effect, a regulated solution.

The advantages of this approach are obvious: the government determines the capacity level and mix, and the suppliers recover the costs from customers. Since all suppliers would have the same obligations (with *ex post* reconciliation rules), the impact on competition is limited – and importantly, the requirements cannot be undercut by switching. In effect, it replicates the CEGB model.

The disadvantages from what is a form of central planning are considerable – both generally in respect of this sort of intervention, and particularly in the electricity market. The information requirements on government are considerable, and the scope for capture by lobbyists is maximised. Government would be picking winners on a continuing basis.

Yet some form of supplier obligation is required for any of the main market reform proposals. The reason is that any requirement on the choice and level of investment requires counterparty. Suppliers will have to be forced to buy the capacity level and profile imposed on a market that would otherwise choose different levels and composition. So some form of Option One is going to be required.

4.2 Long term capacity auctions

One way of getting round the informational and capture problems of direct regulation is to revert to an electricity market which distinguishes between capacity and energy elements, and to introduce a capacity market based upon a capacity auction mechanism.

A capacity auction starts by fixing the quantity of investment required and then invites bids from investors to provide incremental blocks of capacity. As with the EUETS, by fixing quantity, the market then determines the price. The bid is for a (long-term) contract to supply, and the counterparty is the supplier - hence the need for a change to the supplier licences to mandate the purchase of the contracts.

The determination of how much capacity is required and how far ahead to require auctions to cover is outside the market (analogous to the number and time duration of the EUETS permits). Below in section five we consider some of the possibilities.

In addition to auctioning capacity contracts, the climate change targets require that the investment follows the low carbon transition path. The auction can accommodate this by specifying a constraint on the choice of technology. This can be either specific (for example nuclear or offshore wind) or it can be more general (for example a level of carbon reduction). In principle the auction could bring in the demand side, with bids for amounts of demand capacity reductions.

The capacity auction can combine both of these in one mechanism, or there can be separate markets in carbon and in capacity. In principle the UK could translate its domestic carbon targets and carbon budgets into a set of domestic permits for the electricity sector, and run a UK-only permits auction. However, since we already have the EUETS, and this would not take account of the renewables directive target, it is probably best to combine the two elements – the quantity and the technology – within a single mechanism. This is the approach taken below.

The advantages of an auction approach is that: it maximises the information revelation which the government would otherwise have to research; it maximises competition by allowing anyone to bid and thereby gain a contract; it minimises the scope for capture by vested interests; and it provides for secondary markets to develop.

The disadvantages are: that it is likely to be opposed by the incumbents and interest groups and hence may be more difficult to implement; that it would require DECC to design and implement the capacity auctions and designate the institutional architecture; that it would take time; and if different kinds of technology are allowed to bid (at least initially) it would make explicit the costs of offshore wind and other politically favoured technologies.

4.3 A low carbon obligation

A low carbon obligation (LCO) is a generalisation of the Renewables Obligation (RO) already in place. The suppliers would be obligated to buy a proportion of their supplies from low carbon sources, following the required projections from the CCC's carbon budgets. It would be up to suppliers to decide how to meet the obligation (and there would probably be a buy out price).

Given however that the renewables directive requires that the reduction in carbon emissions must be met in part by a fixed renewables target, the LCO would in practice be in two parts – a reserved RO and a residual - at least until

2020 to which the EU target for renewables applies. The non-RO part of the LCO might in practice be largely nuclear (though it could include energy efficiency packages too).

The advantages of the LCO are: that the vertically integrated generators are likely to be in favour and hence implementation would be easier; that it is administratively relatively simple and does not require DECC to design an auction; and that it leaves it to the suppliers to determine the time duration and form of contracting (and places the risk of failure to comply upon them).

The disadvantages are: that it would enhance rather than reduce market power for the vertically integrated incumbents; that entry would be limited; that it would not tend to elicit innovative solutions; and that the form of the LCO would be open to the sorts of lobbying and capture that have been so prevalent in the RO. In this last regard, the banding that has been introduced into the ROCs system could be extended to the LCO – with all the rent seeking that might entail. (It should be remembered in this context that the RO + ROCs regime is one of the most expensive renewables support schemes in the developed world which has to date delivered one of the lowest shares of renewables.)

5. The design of capacity auctions

The auction approach has significant advantages over the other two mechanisms. The security of supply and climate change targets can be met by any of the three – by direct regulation, by auctions, or by the LCO. The first and third are likely to be higher cost because of the lack of competition and extent of capture.

There are lots of different types of capacity markets and it is beyond the scope of this short paper to provide a comprehensive survey. The main distinctions are between: the time duration; the determination of the price of capacity; and the design of the institutional architecture.

There are many short-term capacity markets that are in many ways similar to the capacity mechanism in the original privatisation model described above. The PJM market in the US is an example: a three-year market, with extensive spot trading. The capacity auction recommended here is at the opposite end of the spectrum: it is explicitly for new investments and has a corresponding time horizon (though it may include plant life extension as an option). Rather than the spot price signalling to investors the future returns on capacity, the capacity auctions advanced here determine the future price for the capacity and award contracts accordingly. The secondary market is derivative from these fixed contracts, backed by suppliers who are obliged to buy them. The price is therefore determined in the auction, and not through some mechanism that reflects the LOLP and sets a VOLL – as in the privatisation model.

The institutional architecture differs too between conventional capacity markets and the capacity auction model. In the contract auction model, someone has to determine the quantities, the low carbon transition path, and to an extent even the technologies. The capacity market is a market to meet the exogenously determined requirements on quantity and technologies.

There are lots of ways in which the capacity auctions could be designed. The key issues are:

(i) how to determine the capacity margin – lower bands and the declining quantity schedule

Historically the optimal capacity margin has been assumed to be around 20%, and given the uncertainties about demand, plant performance and the networks, precision is likely to be spurious. There have been lots of attempts to estimate the consumers' willingness to pay, and all suffer from the difficulties in estimating the demand for what is a public good – a system characteristic – the security of supply.

Given the uncertainties it is not sensible to estimate the capacity required for the future at a single point in time. In the short to medium term (say 2010-2020), most of the requirement can be set. For example, the quantity could be set as a minimum level consistent with a low electricity demand growth scenario. The government's ambitious energy efficiency programme might all be delivered and economic growth might be subdued. Then, as information becomes available through the period to 2020 (for example energy efficiency might be much slower to achieve or the economy might grow rapidly), the extra amounts needed can be added to the auctions. These might be met by shorter-term investments – for example CCGTs can be built in around 2 years, whilst plant extensions and outputs might be raised.

In the medium to longer term (say 2020-2030), the uncertainties are magnified and it makes sense to build up only some of the capacity requirements now by committing to capacity contracts. In this period for example, nuclear might be an important technology, and given its planning, licensing and construction timetables, contracts for this capacity could be auctioned now, but much to the total requirement held over to later auctions. This would in effect be a pragmatic and flexible building block approach – what might be described as the declining quantity schedule.

(ii) how to use the auction process to reveal information – a two stage process

In many commercial settings where contracts are bid for, the buyer sets up a process to reveal information from the seller. Indeed, in an important sense, auctions are best seen as an informational revelation process. There are lots of ways this is achieved – from initial expressions of interest; negotiated contracts; identification of preferred bidders; to formal bidding structures.

A key problem in designing capacity markets is that the government is both poorly informed and subject to intense lobbying by vested interests. The experience of major capture in the renewables example provides a warning of what happens if it is left to the government to determine the capacity level and technology without utilising the market to reveal information.

Using the experience of commercial contracting, one approach would be to design a two-stage auction as follows:

> **Stage one:** initial bids are invited for a unit or units of capacity at some future date, unconstrained by technology. Bids are invited across the technology landscape – from nuclear to CCGTs to coal and energy efficiency and renewables. There might also be time bands bid for delivery rather than precise dates. Intermediate aggregation agents might bid composite capacity from a variety of sources.
>
> **Stage two**: given the information revealed about the relative costs of different technologies, and the costs of contracting for more or less total capacity, the second stage is set for a precise quantity to be bid, and which technologies (or jut a low carbon requirement) qualify for round two bidding.

(iii) how to decide who does what – an open institutional process

The process described above does not require an exact capacity forecast, especially as the time horizon expands. It does nevertheless require some modelling of the electricity system and its possible demand and supply paths. Since security is a public good, there can be no assumption that the set of private incentives and interests will add up to the public requirement, and in any event governments know that they will be held responsible by voters for any serious power failures and indeed will be forced to intervene if prices spike to reflect deficient investment.

So the issue under any electricity structure (except in excess supply) is not whether the government forecasts, but how. The two-stage auction helps – particularly revealing the costs of meeting different capacity margins and different carbon targets. But the residual forecasting need to set the quantities remains and the best way to ensure it is robust is to make it an open process. This is best achieved at arms' length form the government, by an independent agency. Again a staged process would be helpful. The agency might publish its model on the web, inviting initial comments. It could allow interested parties to simulate different scenarios. The "expert" elements introduced to the MPC and the OBR would add credibility.

This requirement has implications for the review of Ofgem and its functions currently underway. If Ofgem is split into two – a narrow technical regulator and E-serve – then the modelling and capacity setting exercise would be best carried out separate from the core technical activities. E-serve is primarily a delivery body, and it could be broken up into a series of separate delivery contracts. A new capacity agency would then be created along MPC/OBR lines, taking on some of the modelling work from DECC and having a close interface with the CCC. These functions might all be incorporated in a wider energy agency, and even incorporate the CCC. (In other electricity industries, this function might naturally fall to the system operator of the main transmission grid, but where this is a private company as in the UK, the incentive problems make this undesirable).

(iv) Other issues

There are a host of practical issues that would have to be addressed if the capacity auction approach was to be adopted. These include:

- The consequences for the existing NETA market, which could reduce back into an energy only market. Though not necessary for the development of capacity auctions, there would be an opportunity to recreate an open, compulsory and transparent electricity pool.
- The short term balancing of the market would still be required since the capacity auctions strike *ex ante* contracts in the face of demand and supply side uncertainty. An obvious approach would be to develop short-

term contract auctions to balance the market building on existing practice and the role of the grid operator.

- In stage one of the auction, bidders may be reluctant to come forward if they anticipate their technology will be ruled out in stage two. Since the information revealed is valuable, and the bidding costs are sunk, some mechanism to incentivise bidding and to recover the costs would be appropriate.

6. Conclusions and recommendations

i. the NETA/BETTA electricity market is not fit for purpose (and arguably never was)

ii. the key elements of energy policy – security of supply and climate change – require intervention to set a capacity margin and to influence the choice of technology

iii. there are at least three ways of achieving these objectives – by a supplier obligation to contract according to an externally determined path through regulation; a long term capacity auction market; and a low carbon obligation

iv. the capacity auction approach is to be preferred: it is a market-based mechanism, which maximises competition, reveals information and has resilience to capture

v. the design of the capacity auction should follow a series of stages drawing on lessons from commercial markets. A two-stage process is recommended for serious consideration – stage one open to all technologies and with varying capacity levels; and a second stage which fixes the quantity and makes the required technology judgements

vi. given the need to determine levels and technologies, an independent agency would increase transparency for modelling, and its design should draw on the lessons and experience from the MPC and OBR, and may be part of a wider energy agency, perhaps even incorporating the CCC.