Missing Money and Missing Markets: Reliability, Capacity Auctions and Interconnectors

EPRG Working Paper 1508
Cambridge Working Paper in Economics 1513

David Newbery

Abstract
In the energy trilemma of reliability, sustainability and affordability, politicians treat reliability as over-riding. The EU assumes the energy-only Target Electricity Model will deliver reliability but the UK argues that a capacity remuneration mechanism is needed. This paper argues that capacity auctions tend to over-procure capacity, exacerbating the missing money problem they were designed to address. The bias is further exacerbated by failing to address some of the missing market problems also neglected in the debate. It examines the case for, criticisms of, and outcome of the first GB capacity auction and problems of trading between different capacity markets.

Keywords
capacity markets, renewables, procurement volume, interconnectors

JEL Classification L94, D44

Contact dmgn@cam.ac.uk
Publication February 2015, updated July 2015
Financial Support no specific financial support

www.eprg.group.cam.ac.uk
Missing Money and Missing Markets: Reliability, Capacity Auctions and Interconnectors

Abstract

In the energy trilemma of reliability, sustainability and affordability, politicians treat reliability as over-riding. The EU assumes the energy-only Target Electricity Model will deliver reliability but the UK argues that a capacity remuneration mechanism is needed. This paper argues that capacity auctions tend to over-procure capacity, exacerbating the missing money problem they were designed to address. The bias is further exacerbated by failing to address some of the missing market problems also neglected in the debate. It examines the case for, criticisms of, and outcome of the first GB capacity auction and problems of trading between different capacity markets.

1. Introduction

Britain was the first country to introduce a capacity auction to deliver capacity adequacy after EU Third Package2 (to deliver the Target Electricity Model, TEM) was announced and it coincided with the date by which the TEM was to come into effect. The TEM is designed as an energy-only market that leaves the delivery of capacity adequacy to profit-motivated investment decisions by liberalized and unbundled generation companies. The UK’s Energy Act 2013 that set out the Electricity Market Reform (EMR) rejected relying on an energy-only market and legislated for auctions to deliver capacity adequacy.

This paper examines the design and justification of that capacity auction, its relation to the wider issue of reliability, and criticizes the under-studied issue of how the amount of capacity to procure was determined. It argues that typical capacity auction designs have a bias towards excess procurement, in contrast to fears that the energy-only market would lead to under-procurement. While capacity remuneration mechanisms, of which auctions are potentially the best, are intended to address the missing money problem, by ignoring the missing market problem they perversely exacerbate the missing money problem. Capacity auction design also raises important questions for cross-border trading and the role of interconnectors, which this paper addresses.

2. Reliability, security of supply and capacity adequacy

Energy policy aims to deliver security, sustainability and affordability, but of these three objectives politicians treat security,3 or more broadly, reliability, as over-riding. In electricity markets, reliability is the more inclusive term, measured by long-term

---

1 Paper arising from the Symposium on Energy Markets and Sustainability, Barcelona, 3 Feb. 2015
2 See e.g. http://www2.nationalgrid.com/UK/Industry-information/Europe/Third-energy-package/
3 Bompard et al (2013) provides a useful taxonomy of terms used to describe security.
satisfactory operation “so that instability, uncontrolled separation, or cascading failures of such system will not occur as a result of a sudden disturbance, including a cybersecurity incident, or unanticipated failure of system elements.” (NERC, 2015). It requires both security and adequacy. Security is “(T)he ability to withstand sudden disturbances, such as electric short circuits or unanticipated losses of system components …” (ENTSO-E, 2015). Adequacy is the ability “to supply the aggregate electrical demand and energy requirements of the end-use customers at all times, taking into account scheduled and reasonably expected unscheduled outages of system elements.” (NERC, 2015). Security is a public good supplied by the System Operator (SO) through his acquisition of a range of ancillary and balancing services, while adequacy could, in principle, be delivered by competitive energy-only markets, as the TEM envisages (Oren, 2000).

The TEM was designed as the next step in delivering the EU Integrated Electricity Market (IEM) due to come into effect by December 2014. Its core is an energy-only market with a single auction platform, EUPHEMIA (Pan-European Hybrid Electricity Market Integration Algorithm), for day-ahead, intra-day and balancing trades, which simultaneously clears bids and offers and the use of all interconnectors across the EU, fragmenting the market into different price zones only after interconnectors are fully used. Its working hypothesis, that energy-only markets will deliver capacity adequacy, was based on the experience of Nord Pool, which served as the model for the TEM.

Nord Pool has operated a successful energy-only trading system for many years, as have the major power exchanges such as EEX and APX, without any apparent problems of capacity adequacy, but not all EU countries have (or once) followed this model. Many markets have made or continue to make capacity payments, and DG COMP has been very critical of this practice, arguing that they often have more to do with compensating generators for stranded assets than delivering reliability at least cost. The GB capacity market is, as of early 2015, the only capacity market to be explicitly designed and operating since the announcement of the Third Energy Package.

As a number of countries are now considering whether, and if so how, to introduce a (or reform their) Capacity Revenue Mechanisms it is timely to examine the British experience. Eurelectric is the latest organization to recognize that not all EU countries will be happy with the reference energy-only markets of the TEM, and “recognizes that properly designed capacity markets, developed in line with the objective of the IEM, are an integral part of a future market design.” (Eurelectric, 2015, p4.) While that document discusses what might be required to deliver a reliability standard, it is somewhat sceptical on how this might be achieved, instead arguing that “whatever reliability standard is chosen, Regulators and TSOs should compute it with methodologies and tools that are publicly available.” A second objective of this paper is

---

4 Northern Ireland is part of the SEM discussed below, leaving Great Britain, GB, as the rest of the UK.
to assess how this might best be done, guided by the principle of addressing the missing market problem.

3. Missing money and missing markets
Given the need to instantly balance supply and demand in the electricity system, ensuring short-term security of supply is normally an obligation placed on the SO, while longer term capacity adequacy is often the subject of regulatory and political concern. EU electricity markets are now liberalized, and generation is, for the most part, not subject to traditional utility regulation, but to normal competition policy both domestically and under the scrutiny of DG COMP. If investment decisions could be solely guided by strictly commercial decisions and if markets were not subject to policy interventions or price caps, it is plausible that capacity adequacy could be delivered by profit-motivated generation investment without explicit policy guidance. For this to be the case, investors need confidence that the revenue they earn from the energy markets (including those supplying the ancillary services that the SO needs to ensure short-term stability) will be adequate to cover investment and operating costs.

If this revenue is not adequate, there is a “missing money” problem (Joskow, 2013), but if it is adequate but not perceived to be so by generation companies or their financiers, then there is a “missing market” problem (Newbery, 1989). Missing money problems arise if price caps are set too low (below the Value of Lost Load, VoLL), or ancillary services, such as flexibility, ramp-rates, frequency response, black start capability, etc. and/or balancing services are inadequately remunerated, or transmission access charges are inefficiently high (important in distorting exit decisions), and/or, energy prices are inefficiently low. Inefficiently low wholesale prices seem less likely as the normal problem is one of market power raising prices above their competitive level, and prices are not necessarily inefficiently low just because there is excess capacity.

Missing markets create problems if risks cannot be efficiently allocated with minimal transaction costs through futures and contract markets, or if important externalities such as CO2 and other pollutants are not properly priced. The concept of missing markets can be usefully extended to cases in which politicians and/or regulators are not willing to offer hedges against future market interventions that could adversely affect generator profits. These arguments have been extensively covered in the literature, recently in the Symposium on ‘Capacity Markets’, (Joskow, 2013; Cramton, Ockenfels and Stoft, 2013). Almost all the discussion about capacity mechanisms concentrates on the missing money problem and whether the various market and regulatory/political
failures are sufficient to justify a capacity mechanism, and if so, what form it should best take.5

Both the missing money and missing market failures have risen in salience as renewable electricity targets have become more ambitious at the same time as the EU Emissions Trading System has failed to deliver an adequate, durable and credible carbon price, and as such is under constant threat of reform. Absent a futures market with a credible counter-party it is hard to be confident that future electricity prices will be remunerative for unsubsidized generation, and harder to convince bankers or shareholders of the credibility of investment plans based on forecast revenues. If renewables continue their planned increase in market share mandated by the EU

Renewables Directive (2009/28/EC) they will depress average energy prices. This does not of itself give rise to an adequacy problem, although utilities may justifiably complain that their past investment decisions have been partially expropriated by unanticipated political actions. However, it increases the demand for existing services such as primary reserves, fast frequency response and inertia and may also increase the need for additional ancillary services. If these services are not yet adequately defined and/or their future prices are hard to predict there is a missing market problem. If these services are underpriced by SOs whose powers of balancing supply and demand may be met by administrative or regulatory means (e.g. by requiring those connecting to the grid to make some of these services available as part of the grid code), there is a missing money problem. In either case these may precipitate a capacity adequacy problem.

Newbery (2013) documents the analysis that led to the UK’s EMR and its embodiment in the Energy Act 2013, which was primarily designed to rescue the UK’s failing attempt to meet its renewable energy targets at least cost. Renewable electricity suffered from both the missing money and missing market problems, as support was provided by Premium Feed-in Tariffs via Renewable Obligation Certificates (ROCs). Their value depended on future electricity prices and the supply-demand balance for the ROCs, both potentially volatile. While fossil generation enjoyed a natural hedge in that electricity prices mirrored fossil fuel prices (see Newbery, 2013, fig 2), renewables and nuclear power, whose fuel costs are zero or very small, are exposed to the full volatility of electricity prices (Roques et al., 2006; 2008), which increases risk unnecessarily and hence raises the cost of capital, the major part of the total cost. The EMR instead proposed long-term Contracts for Differences (CfDs) with fixed indexed strike prices to solve both the missing money and the missing futures market problems.

3.1 Market failures in delivering reliability

Before the electricity industry was liberalized and unbundled, the SO had ownership control of generation and transmission and was responsible for both system security and adequacy. Planned investment ensured that both capabilities would be assessed, which was also much easier when essentially all plant had (at least in aggregate) a predictable and controllable output. The main security aspects were handling very short-run increases in demand (notably during intermissions in major sporting events when consumers all simultaneously switch on their electric kettles) or those caused by the loss of a large station or transmission link. The standard approach was to specify a reserve margin and ensure adequate short-run flexibility by the choice of plant type. Thus the Central Electricity Generating Board, CEGB, that pre-dated the British restructuring of 1989, computed the required gross reserve margin at 19% based on a Loss of Load Probability (LoLP) calculation and a reliability measure (disconnecting some consumers in three winters over a 100 years, decided in 1962 (Bates and Fraser, 1974, p122). It built pumped storage systems to provide fast response, peaking capacity and to use surplus night-time nuclear power (Williams, 1991), as well as jet-derivative gas turbines for fast ramping.

With liberalization and unbundling all these security services had to be separately procured by the SO. Some had to be remunerated, others, such as inertia and the additional security offered by interconnectors, came at no cost to the SO. The SO now had to manage transmission constraints using price signals (distorted through a failure to adopt efficient locational marginal pricing) dealing bilaterally with generators out to maximize profits and typically possessing short-run market power. In addition the SO had to secure various ancillary and balancing services through spot and contract markets, just when the challenges of handling increasing volumes of intermittent and less predictable wind were increasing (Newbery, 2010, introducing “Large-scale wind power in electricity markets” in that issue of Energy Policy; Newbery, 2012).

Although the delivery of security services is a public good provided by the SO, these services can be procured through some combination of forward, prompt and real-time markets and contracts. As problems of intermittency increase, so does the challenge of ensuring that these services are efficiently priced and procured (Pöyry, 2014).

The Single Electricity market (SEM) of the island of Ireland is probably at the forefront of addressing these problems, as it is a small moderately isolated system in which individual power stations are large and lumpy relative to peak demand (up to 10%) and the system is being adapted to handle up to 70% non-synchronous wind penetration.\(^6\) One (implausible) solution would have price signals varying over very short periods of

---

\(^6\) Conventional rotating generation turbines are synchronised to the grid frequency and have substantial inertia, so that if there is a momentary loss of supply, that inertia prevents the frequency falling too fast. Wind power has effectively no inertia, which has to be provided in some form to maintain frequency within acceptable limits.
time. A sudden fall in frequency caused by a sudden increase in demand relative to supply means that the value of power in the next cycle (1/50th of a second) has increased, and the speed of response is key to minimizing the disruption. Figure 1 lists the various products and their time domain that the SEM Committee defined in their consultation document (SEM, 2014), grouped into the three categories of Synchronous Inertial Response (SIR, usually delivered by the inertia of synchronized generators), operating reserves (primary, POR; secondary, SOR; and tertiary, TOR), and Ramping or Replacement Reserves (RR), with some overlap between them.

If prices were to move in response to instantaneous system conditions, then it would be potentially profitable to have the capability to respond on the appropriate time scale. In practice, market designs vary in their granularity, with the most flexible having 5 minute settlement periods (Australia). Typically Continental balancing markets have a 15 minute settlement period, while GB and the SEM have a half-hourly settlement period in the Day Ahead Market (DAM), and most Continental power exchanges and the EUPHEMIA auction platform have hourly resolution in the DAM. Increasing granularity improves the accuracy of the temporal pattern of price signals to guide the choice of flexibility, but runs up against the practical constraint that the system state requires a finite amount of time to re-estimate, probably of the order of minutes, while the transaction costs of dealing at high frequency make very short-term markets illiquid.

Figure 1 Various flexible systems services defined for the SEM
Source: SEM (2014)

Given the inability and absence of energy markets at this level of time granularity, new products are needed to supplement existing products. The SEM Committee consulted on how to procure them in early 2015, given that the supply side for some

---

7 Also suggested by Mott MacDonald (2013). California is interested in 5 minute granularity.
services is highly concentrated, raising market power issues. In this case many markets are missing as their procurement is still undecided, making it hard to estimate their future revenues. When choosing what type of generating plant to build, investors have a choice of characteristics with inevitable trade-offs: flexible plant with high ramp rates is either more expensive or less efficient than less flexible plant with lower ramp rates, but unless investors can forecast the revenue from selling these security services, it is hard to make efficient plant choices.

The classic public good problem facing the regulators and the SO is how to value these various services, given that they are provided in bundles of varying proportions by different technologies. Some of the services reduce LoLP, and should be informed by the VoLL, but others reduce the need to curtail renewable generation, which is a political objective of uncertain value.

3.2 Defining and measuring reliability

Eurelectric (2015) provides a useful summary of the various ways used to measure reliability. Most EU electricity systems specify the “Loss of Load Expectation” (LoLE), which for most and for GB is three hours per year. This is a forward looking measure, that taking a representative and large number of possible outcomes (of weather, plant reliability, demand, etc.) for some future period, the electricity system should perform better than averaging “Losses of Load events” of three hours per year. Clearly there is a difference between losing load for the entire population and controlled disconnections for a modest number of consumers, and Eurelectric (2015) argues that a better measure is the Loss of Expected Energy (LoEE) measured in MWh/year. National Grid (2014a) uses the cost of LoEE to determine the procurement amount for the capacity auction.

The former CEGB had a standard of disconnections in three winters in 100 years or 3%, while Belgium had a 1% standard and Spain a 5% standard (Webb, 1977). This was often translated into a gross reserve margin (Transmission Entry Capacity less peak load) required to deliver the reliability standard. Thus the CEGB set it at 19% , made up of a de-rating factor of 11% and remaining outage risks of 6%, so that the de-rated reliability margin was 7% (6%/0.89). Such quantitative measures sit well within the planning framework of a centrally controlled electricity system pre-liberalization.

The other more market-oriented approach to reliability is to specify the Value of Lost Load, VoLL, the amount that consumers should be willing to pay to avoid disconnection. In a future with sufficiently smart meters, consumers would be able to sign a contract with the electricity supplier stating the maximum amount they would be willing to pay for each tranche of firm power, with the smart meter disconnecting appliances at each specified price point, leaving presumably some lights and electronic equipment until last. As such that would represent the complete private good market solution to the problem and would avoid the missing money if not the missing (future)
market problem, provided the short-run prices were efficiently set at their efficient scarcity value. This is the sum of the System Marginal Cost (SMC) plus a Capacity Payment, CP, where

$$CP = \text{LoLP} \times (\text{VoLL} - \text{SMC}), \quad (1)$$

The relationship between the security standard and the VoLL is symmetric, in that if capacity investment decisions are based on revenues determined by (1) and the VoLL is pre-determined, then the resulting capacity will give rise to a LoLE. If the standard is as a predetermined LoLE, the cost of new capacity implies a cost of delivering the LoLP and hence the VoLL.

Britain has followed both models. The English Pool set the VoLL at £(2012)5,000 (€6,250/MWh at £1 = €1.25), letting the market determine capacity. After the Pool was replaced with an energy-only market in 2001, the Department of Energy and Climate Change (DECC) specified the LoLE. National Grid (2014a) deduced the 2018 VoLL as £(2012)17,000/MWh (€(2012)21,250/MWh), higher than direct estimates of the willingness to pay to avoid disconnections (London Economics, 2013).

3.3 Can energy-only markets deliver adequate reliability?

One completely legitimate case for a capacity payment is that if generators are required to bid their Short Run Marginal Cost (SRMC, mostly fuel costs), as under the Bidding Code of Practice of the SEM (SEM, 2007), they will fail to recover their fixed costs without such an addition. The Electricity Pool of England and Wales also added the CP of (1), but allowed generators to offer an unrestricted supply function (which, given their market power, was often above SRMC, Green and Newbery, 1992; Newbery, 1995; Sweeting, 2007). In this period of benign liberalization, high prices led to considerable entry and an excessive reserve margin.

In the energy-only market envisaged by the TEM, generators will offer supply functions that should reflect the scarcity value of electricity (and their degree of market power). Figure 2 shows the day-ahead price duration curves for several European power exchanges in 2012. What is striking is that most exchange prices do not exceed €200/MWh, and even the most peaky, France, only does so 0.25 of 1% of the time (about 22 hours per year). Given that the VoLL in the English Pool until 2001 was €(2012)6,250/MWh and the current implied VoLL in GB is €(2012)21,250/MWh (both at £1 = €1.25), these prices indicate a low LoLP or high reliability. Given existing capacity levels that is a reasonable inference, but the problem again is one of missing (futures) markets. Investment lags in delivering capacity adequacy are 2-4 years for peaking plant (longer for most base-load plant), beyond the time horizon of adequately liquid futures markets (and they only offer one year’s hedge).
Figure 2 Price Duration curves of day-ahead hourly prices, 2012

Sources: MIP (Market Index Price) and NL prices from APX, Germany and France from EEX

On the other hand, Figure 3 shows that the 2008 balancing buy prices\(^8\) in the energy-only market that replaced the Pool were considerably peakier than the old Pool prices (which included an explicit CP and also probably reflected more market power). Thus energy-only markets can reflect scarcity, and properly calculated capacity payments may be very low if the reserve margin is adequate as LoLP is roughly exponential in demand less derated capacity (Newbery, 2005). However, by 2013-14, the GB Balancing Mechanism had a price duration curve quite similar to those shown in Figure 2, with prices above €200/MWh for less than 0.25 of 1% of the time, and well below the French day-ahead price duration curve.

Thus one might conclude that energy-only markets (which include balancing markets) can deliver sufficiently sharp scarcity prices that should signal the profitability of adequate new investment, provided all the other security services are adequately remunerated (i.e. resolving any of those missing market problems). This might be plausible if all investment decisions were taken on commercial grounds as in the 1990s, that prices were not capped, that the policy environment were predictable and stable, and that either liquid forward market existed for a reasonably fraction of the proposed plant life (i.e. 20+ years ahead of the final investment decision) or credible long-term power purchase agreements could be signed with credit-worthy counterparties. Unfortunately, hardly any of these conditions hold in the TEM.

\(^8\) For a description of the British Balancing Mechanism see Newbery (2005).
3.4 Market, institutional and political/regulatory failures

While price caps are set at rather low levels in the US, exacerbating the “missing money” problem, there are also, if much higher, price caps in EUPHEMIA (for day-ahead at €3,000/MWh, a price that France has hit on numerous occasions). The lack of forward markets and long-term contracts might not be so critical if the future were reasonably predictable and stable, but this is far from the case at present. EU Climate Change policy is failing, in conflict with Renewables Directive (2009/28/EC), and surely ripe for as yet uncertain reform. Large volumes of unreliable renewables increase the need for flexible reserves, which in the past came from obsolescing plant, mostly oil or coal. This plant is now being decommissioned because of the Large Combustion Plant Directive and the Integrated Emissions Directive.

Increasing renewables (mainly wind and solar PV) add little to reliable capacity, as it is unavailable on still cold dark winter nights, but reduces average wholesale prices. If the average capacity factor of on-shore wind is 25%, then the GB target of 30% of electricity from wind requires capacity of 30/25 times or 120% of average demand. In windy conditions that would often displace all conventional plant and could lead, under present subsidy structures, to negative prices.

Intermittent generation increases the need for additional flexible plant that can be called up at short notice if the wind falls or the sun fades. In addition, new plant will be needed to replace retiring plant (not just coal, but in the UK, France and Germany, substantial volumes of nuclear plant as well). That plant will need considerably higher
prices when renewable output is low than has been recently experienced. Even if the carbon price is currently low, the EU is committed to an 80% reduction in Greenhouse Gas emissions by 2050. Coal is twice the carbon intensity of gas, so utilities are unlikely to build durable (40-60 year) coal-fired plant that would face tight future emissions limits, leaving gas-fired plant as the only alternative. Unfortunately, crashed electricity prices and high gas prices precipitated by the closure of Japan’s nuclear fleet has made their economics very unattractive.

The UK introduced a carbon price floor in the 2011 Budget (HMT, 2011) that would support the price of CO₂ at £16/tonne in 2013, rising to £30/tonne (€35/tonne) in 2020, and projected to rise to £70/tonne by 2030 (all at 2009 prices). This threatened the operation of existing coal-fired plant. As an example of policy instability, the 2014 Budget froze the carbon price floor – clearly an instrument subject to the whim of chancellors creates additional investment uncertainty. It would be a brave politician who trusted these markets to deliver reliability.

4. The GB capacity auction

In response to the looming capacity crunch and other market failures set out in Ofgem (2009) and subsequent DECC consultations (Newbery, 2012a), the UK passed the Energy Act 2013 setting out the EMR, which includes a Capacity Mechanism to ensure adequate capacity. GB now has the Secretary of State for Energy & Climate Change, advised by DECC, deciding how much capacity is required. The capacity auction is a single-price descending clock auction with a demand schedule as shown in figure 4. National Grid as SO was charged to recommend the target volume of capacity to secure four years after the auction (which was termed the T-4 auction).

National Grid (2014a) chose the amount to procure balancing the cost of additional capacity against the cost of the Loss of Expected Energy, as shown in Figure 5. National Grid (2014a) projected that the auction clearing price would likely be set at the Cost of New Entry (CONE), estimated at £49/kWyr. This was the missing money a Combined Cycle Gas Turbine (CCGT) might need given its revenues from all other markets and after paying the Transmission Network Use of System (TNUoS) charges. These range from £30/kWyr (in NW Scotland), to negative (-£5/kWyr in Cornwall) (National Grid, 2013). Entrants are given 15-year indexed contracts, while existing plant receive one-year contracts to defer exit decisions until the next auction.

The missing money can be estimated from the VoLL (£17/kWh) less the maximum the SO pays for balancing actions (£6/kWh) to give £11/kWh, times 3 hrs LoLE, or £33/kWyr. The effective cap in the Balancing Mechanism is £9,999/MWh would reduce the missing money to £21/kWyr.
The auction design was best-practice (Newbery and Grubb, 2015) but flawed in requiring the SO to advise the minister on the procurement amount. The SO stands accountable if “the lights go out” but does not pay for the capacity. The minister wishes to avoid newspaper headlines predicting blackouts resulting from his decision. Both
argue for excess procurement. DECC appointed an independent Panel of Technical Experts (PTE) to comment on the analysis, and they made a number of strong (but, for the 2014 auction, ineffective) criticisms.

### 4.1 Criticisms of the capacity to procure

The PTE first criticized the terminology of “Loss of Load” as emotive and misleading. The GB regulator, Ofgem, defines a “Loss of Load event” as one in which market demand exceeds market supply and as such the SO has to intervene to balance the system. For that purpose the SO can call on a range of increasingly expensive options: asking generators to temporarily exceed rated capacity; invoking ‘new balancing services’; cutting interconnector exports to zero, requesting imports; reducing voltage (“brown outs”), before finally resorting to selective disconnections. The crucial point is that these actions cost less, often much less, than VoLL and hence bias the unserved load cost and the target capacity in figure 5 upward.

Successful Capacity Market Units in the auction receive a Capacity Agreement which requires them to be available in “stress events” signaled four hours ahead. DECC (2014, §391) defines these events as “any settlement periods in which either voltage control or controlled load shedding are experienced at any point on the system for 15 minutes or longer. … Periods of voltage control or load shedding resulting from failures or deficiencies in the transmission or distribution systems are not considered as stress events.” However, these “notices of inadequate system margin” are issued “based on the available capacity (declared ‘maximum export limit’ (MEL) minus transmission system demand and reserve for response capacity.” DECC (2014, box, p107.)

National Grid (2014a) chose the amount to procure using a Least Worst Regrets approach as it was unwilling to attach explicit probabilities to the various scenarios considered. The result of overvaluing the cost of “Loss of Load” is to increase the capacity at which the Least Worst Regret cost schedule is minimized (figure 5). “Slow progression” reaches a cost minimum at 53.3 GW for 2018-19 delivery. The net procurement target is 53.3-w-x-y-z-0.4 GW, where w, x, y and z refer to various distributed energy resources and opt-out plant. The 0.4 GW is existing short-term operating reserve.

The PTE (DECC, 2014a) strongly criticized National Grid for assuming no net imports in stress periods, despite 3.75 GW interconnection capacity and potential new capacity of 2.25 GW that might be available by 2018-19. This seemed perverse, as all

---

9 “The new balancing services are Demand Side Balancing Reserve (DSBR) and Supplemental Balancing Reserve (SBR).” National Grid announced its tender for these new services on 10 June 2014 (http://www.nationalgrid.com/uk/electricity/additionalmeasures).

10 The CEBG estimated that voltage reductions reduce load by 7½% in the 1970s (Bates and Fraser, 1974) but National Grid now estimates only 1½% in the absence of firm evidence.
parties (Ofgem, DECC and National Grid) agreed interconnectors increased security. Three reports commissioned by these parties (Pöyry, 2012, 2013; Redpoint, 2013) argued that interconnector capacity could displace domestic capacity by 50-80% of its value. Even DECC’s *Final Impact Assessment*, published just before the procurement decision (DECC, 2014b), estimated the amount of interconnector capacity to include in the total procurement amount at 2.9 GW. Ignoring interconnectors could move the auction clearing price from the cost of new entry of £49/kW yr to that set by existing plant (maximum of £25/kW yr), increasing the auction cost by £1.3 billion per year.

Ignoring interconnectors seemed particularly perverse as the TEM aims to integrate markets across borders. Market coupling already dispatches GB Continental interconnectors in the EUPHEMIA day-ahead market (DAM). Interconnector flows already reflect willingness to pay in the DAM, and will soon do so in the intra-day and real time markets when network codes are agreed.

### 4.2 Possible consequences of excessive capacity procurement

Excess procured capacity will lower future wholesale prices with a number of effects, not all immediately obvious. First, lower prices reduce the revenue new entrants can expect from the energy markets, increase the CONE and raise the auction price. Second, it undermines the old market design in which investment in conventional generation was at the discretion of private companies making commercial decisions. No company would invest in conventional generation without a capacity agreement given its large disadvantage compared to those with agreements. The amount of new plant will therefore be entirely determined by the minister, ending a key element of the liberalized market. All non-fossil generation will also be granted long-term Contracts for Difference (CfDs) under the EMR, moving GB to the Single Buyer model ruled out in earlier EU Electricity Directives.

Third, lower prices increase payments to low-carbon CfDs, which receive the difference between the contracted strike and wholesale price. As the Government limits total renewables payments through the Levy Control Framework, the perverse effect is to support less renewable electricity, although the EMR was designed to remove obstacles to meeting the renewables target.

Fourth, the commercial case for interconnectors depends on price differences, with GB typically importing from cheaper Continental markets. Lower GB prices reduce arbitrage profits, undermining the investment case for the additional interconnectors when they are increasingly needed to balance growing intermittent generation across wider market areas. Ignoring interconnectors risks a self-fulfilling but expensive policy of autarky.
Fifth, although the future wholesale price will be lower, offsetting possibly a large part of the consumer cost, it will be hard to convince consumers of this. They will see the gross cost estimated at 53.3 GW x £49m/GWyr = £2.6 billion per year.

Finally, on 2nd December 2014, after the PTE had published its critical report and the Secretary of State had decided on the procurement volume, but before the auction on 18th December, the Treasury’s National Infrastructure Plan announced that interconnectors would be eligible for the 2015 capacity market.\footnote{See \url{https://www.gov.uk/government/collections/national-infrastructure-plan}} It would have been easy to have left room for interconnectors (e.g. adding another element to the \( w-x-y-z-0.4 \) GW deduction from the target volume) and lower the net amount to procure.

4.3 The outcome of the 2014 capacity auction

The auction cleared at £(2012) 19.40/kWyr (National Grid, 2014b). The auction produced several surprises. First, the auction cleared at less than 40% of the predicted CONE value of £49/kWyr (although close to the missing money estimated above assuming a balancing cap of £9,999/MWh). The estimated CONE was based on new entry of CCGT, and two CCGTs entered, supplying about 60% of the total 2,795 MW new entry. Second, the next largest (28%) entry category was OCGT/ reciprocating engines, average size 11 MW. The third largest contribution (6%) was from unproven Demand Side Response (DSR, all with a one-year contract, other new entrants have 15-year contracts).

One might expect that DSR and OCGTs would require a lower strike price, particularly as they can contribute to significantly reducing TNUoS charges if they are embedded with major loads, but the low price that CCGTs were willing to accept is surprising, and may be based on optimistic views of gas prices (which were expected to decline by the time of the auction) or high balancing prices. National Grid announced its tender for new balancing services on 10 June 2014,\footnote{At \url{http://www.nationalgrid.com/uk/electricity/additionalmeasures}} reducing the extent of the missing money for these services after DECC had published its estimate of the net CONE.

The final point is that the auction demonstrates the value of market-based methods of revealing entry costs, and the danger of leaving such decisions to SOs or regulators (as in the SEM,\footnote{See \url{http://www.allislandproject.org/en/cp_current-consultations.aspx?article=75c548a7-34ee-497e-afdd2-62f8aa0062df}} where the regulators calculate the cost of Best New Entry and set it at a high price).

5. Biases in capacity auctions and energy-only markets

The arguments above strongly suggest that if procurement decisions are left to politicians advised by the SO, they will err on the high side, and tend to ignore supplies from outside their control area (over interconnectors). Their caution is exacerbated by the emotive and

\footnote{See \url{https://www.gov.uk/government/collections/national-infrastructure-plan}} \footnote{At \url{http://www.nationalgrid.com/uk/electricity/additionalmeasures}} \footnote{See \url{http://www.allislandproject.org/en/cp_current-consultations.aspx?article=75c548a7-34ee-497e-afdd2-62f8aa0062df}}
misleading terminology of “Loss of Load”. Some of these shortcomings can be addressed by requiring the SO to cost and quantify the actions that are taken in stress events that fall short of controlled disconnections. Delegating the decision to independent agencies, perhaps to an Independent Planning and System Operator could depoliticize the decision (but might not remove the fear of disconnections through inadequacy, nor the bias of not paying for capacity).

There is a more fundamental problem in that if future energy prices are competitively delivered and if all security services (ancillary and balancing) are properly priced, the missing market and missing money problems can both be addressed by offering suitable hedging contracts, of which the auctioned capacity agreement is an excellent example. Price caps could be replaced by reliability options or one-way CfDs that have a high strike price, and which allow consumers or their suppliers to hedge against high prices while allowing the spot and balancing market prices to reach scarcity levels needed for efficient actions (in demand reduction and interconnector trade) (see e.g. Vásquez et al, 2002; Bidwell, 2005).

Now consider the costs of under or over-specifying the amount to procure. Over-procurement, as noted above, risks depressing future prices and hence reducing future energy and ancillary service revenues, requiring a higher auction price in compensation. While addressing the missing markets problem it risks amplifying the missing money problem. In contrast, under-procurement leads to expectations of higher future prices, requiring a lower capacity auction bid as the capacity agreement does not preclude earning revenues in all the energy markets. If the price is very low, investors may conclude that investing without a capacity agreement has relatively low risk, particularly as the design of the GB auction offers a T-4 contract of 15 years for new plant, but successive T-1 contracts of one-year for existing plant at the same clearing price, for which speculative plant would be eligible. A signal to err on the side of under-procurement would be underwritten by the ability to true up closer to delivery, reducing risks, as any over-procurement would merely delay the moment at which more capacity was needed in the auction, and should limit the period of inadequate revenue to a year or so.

6. The need for regional coordination
Eureletric (2015) argues strongly for a design of regional capacity markets which places the obligation on generation (or demand) regardless of national location. This requires a common regional capacity adequacy assessment and no double payment (i.e. if capacity has an agreement from country A it would be denied one from B and would be excluded from B’s capacity assessment, subject to adequate interconnector capacity from B to A).

While this may be a desirable long-run objective, it is hardly compatible with the existing Third Package. Meanwhile countries will have to decide how to treat
interconnectors connecting possibly very different (or no) capacity markets. Provided the auction platform can accommodate efficient scarcity prices (i.e. provided at least the intra-day and balancing markets are not capped at too low a level), then trading over any interconnector will only benefit a country that ensures that the relevant prices are efficient, as in (1). Suppose A and B trade but A has a higher VoLL than B, and hence a larger reliability margin. If A and B both have stress events, A can outbid B to secure imports, with B accepting a higher LoLP reflecting its lower willingness to pay to avoid disconnections, while ensuring that domestic consumers were insulated from these high trading prices through the contract coverage provided by the Reliability Options.

The logic of making adequacy as close as possible to a private market good (through allowing efficient pricing) is that there can be gains from trade for the efficiently priced market even when market designs are different. If prices are inefficient in B, then it is they who lose, not A. That provides incentives to reform and avoids the need for politically fraught agreements on harmonization.

7. Conclusions and policy implications
Missing money and missing markets provide compelling reasons for a capacity payment in competitive electricity markets dominated by politically determined and subsidized unreliable generation and where investors lack confidence in future revenues. Capacity auctions (GB provides a good example) address the missing money problem and part of the missing market problem (the missing futures markets), which still needs efficient solutions - markets, auctions and procurement contracts - for location, flexibility, etc. needed to deliver security. More complex category auctions may also be the best way of procuring these services. The part of the adequacy debate that has been neglected is how to, and who should, determine the amount and type of capacity to procure (generation, DSR, interconnection), a problem that is exacerbated by misunderstandings over what a “Loss of Load” event means and what it might cost.

This paper argues that this neglect biases towards over-procurement, which leads to a self-fulfilling prophecy that merchant generation investment can no longer be relied upon. Perversely, this exacerbates the missing money problem that capacity auctions were designed to address. The bias is further exacerbated by failing to address some of the missing market problems that have also been neglected in the debate.

Whether or not interconnectors should be included in auctions is less important than that their contribution should be recognized in determining the procurement amount. All British interconnectors are HVDC controllable links whose flows can be rapidly reversed and as such could provide extra imports at short notice, but they can also impose sudden large loads on the GB system if they switch to exporting. The UK Government, possibly under pressure from DG COMP over State Aid concerns after the PTE had
published their criticisms (DECC, 2014a), decided to include interconnectors in the next auction for 2019-20 delivery, and consult on how to determine their reliable capacity.

There remain a number of policy issues to resolve, not least how the European auction platform EUPHEMIA will determine the direction of flows close to real time, when stress events that the capacity auction was designed to address are likely to emerge. EUPHEMIA has a €3,000/MWh price cap on the DAM, well below the VoLL. It has not yet (early 2015) fixed price caps for intra-day and balancing actions. If prices in the real time European markets could properly reflect scarcity, and if the GB market could deliver the true scarcity prices to EUPHEMIA (including the CP of equation (1)) then good market design and pricing would deliver efficient solutions, and other countries with less good pricing would lose out, motivating them to improve their market design. Price caps hinder this aim, and instead good rules will be needed for out-of-market actions when price caps are reached, and/or markets no longer determine flows, and SOs have to intervene. These rules or bilateral agreements between the SOs at each end of interconnectors are currently lacking or incomplete. If these problems can be addressed, then the more demanding task of harmonizing regional capacity markets may be unnecessary.
References


Missing Money and Missing Markets: Reliability, Capacity Auctions and Interconnectors

EPRG Working Paper 1508
Cambridge Working Paper in Economics 1513

David Newbery

Abstract In the energy trilemma of reliability, sustainability and affordability, politicians treat reliability as over-riding. The EU assumes the energy-only Target Electricity Model will deliver reliability but the UK argues that a capacity remuneration mechanism is needed. This paper argues that capacity auctions tend to over-procure capacity, exacerbating the missing money problem they were designed to address. The bias is further exacerbated by failing to address some of the missing market problems also neglected in the debate. It examines the case for, criticisms of, and outcome of the first GB capacity auction and problems of trading between different capacity markets.

Keywords capacity markets, renewables, procurement volume, interconnectors

JEL Classification L94, D44

Contact dmgn@cam.ac.uk
Publication February 2015, updated July 2015
Financial Support no specific financial support

www.eprg.group.cam.ac.uk
Abstract
In the energy trilemma of reliability, sustainability and affordability, politicians treat reliability as over-riding. The EU assumes the energy-only Target Electricity Model will deliver reliability but the UK argues that a capacity remuneration mechanism is needed. This paper argues that capacity auctions tend to over-procure capacity, exacerbating the missing money problem they were designed to address. The bias is further exacerbated by failing to address some of the missing market problems also neglected in the debate. It examines the case for, criticisms of, and outcome of the first GB capacity auction and problems of trading between different capacity markets.

1. Introduction
Britain was the first country to introduce a capacity auction to deliver capacity adequacy after EU Third Package (to deliver the Target Electricity Model, TEM) was announced and it coincided with the date by which the TEM was to come into effect. The TEM is designed as an energy-only market that leaves the delivery of capacity adequacy to profit-motivated investment decisions by liberalized and unbundled generation companies. The UK’s Energy Act 2013 that set out the Electricity Market Reform (EMR) rejected relying on an energy-only market and legislated for auctions to deliver capacity adequacy.

This paper examines the design and justification of that capacity auction, its relation to the wider issue of reliability, and criticizes the under-studied issue of how the amount of capacity to procure was determined. It argues that typical capacity auction designs have a bias towards excess procurement, in contrast to fears that the energy-only market would lead to under-procurement. While capacity remuneration mechanisms, of which auctions are potentially the best, are intended to address the missing money problem, by ignoring the missing market problem they perversely exacerbate the missing money problem. Capacity auction design also raises important questions for cross-border trading and the role of interconnectors, which this paper addresses.

2. Reliability, security of supply and capacity adequacy
Energy policy aims to deliver security, sustainability and affordability, but of these three objectives politicians treat security, or more broadly, reliability, as over-riding. In electricity markets, reliability is the more inclusive term, measured by long-term

---

1 Paper arising from the Symposium on Energy Markets and Sustainability, Barcelona, 3 Feb. 2015
2 See e.g. http://www2.nationalgrid.com/UK/Industry-information/Europe/Third-energy-package/
3 Bompard et al (2013) provides a useful taxonomy of terms used to describe security.
satisfactory operation “so that instability, uncontrolled separation, or cascading failures of such system will not occur as a result of a sudden disturbance, including a cybersecurity incident, or unanticipated failure of system elements.” (NERC, 2015). It requires both security and adequacy. Security is “(T)he ability to withstand sudden disturbances, such as electric short circuits or unanticipated losses of system components … ” (ENTSO-E, 2015). Adequacy is the ability “to supply the aggregate electrical demand and energy requirements of the end-use customers at all times, taking into account scheduled and reasonably expected unscheduled outages of system elements.” (NERC, 2015). Security is a public good supplied by the System Operator (SO) through his acquisition of a range of ancillary and balancing services, while adequacy could, in principle, be delivered by competitive energy-only markets, as the TEM envisions (Oren, 2000).

The TEM was designed as the next step in delivering the EU Integrated Electricity Market (IEM) due to come into effect by December 2014. Its core is an energy-only market with a single auction platform, EUPHEMIA (Pan-European Hybrid Electricity Market Integration Algorithm), for day-ahead, intra-day and balancing trades, which simultaneously clears bids and offers and the use of all interconnectors across the EU, fragmenting the market into different price zones only after interconnectors are fully used. Its working hypothesis, that energy-only markets will deliver capacity adequacy, was based on the experience of Nord Pool, which served as the model for the TEM.

Nord Pool has operated a successful energy-only trading system for many years, as have the major power exchanges such as EEX and APX, without any apparent problems of capacity adequacy, but not all EU countries have (or once) followed this model. Many markets have made or continue to make capacity payments, and DG COMP has been very critical of this practice, arguing that they often have more to do with compensating generators for stranded assets than delivering reliability at least cost. The GB capacity market is, as of early 2015, the only capacity market to be explicitly designed and operating since the announcement of the Third Energy Package.

As a number of countries are now considering whether, and if so how, to introduce a (or reform their) Capacity Revenue Mechanisms it is timely to examine the British experience. Eurelectric is the latest organization to recognize that not all EU countries will be happy with the reference energy-only markets of the TEM, and “recognizes that properly designed capacity markets, developed in line with the objective of the IEM, are an integral part of a future market design.” (Eurelectric, 2015, p4.) While that document discusses what might be required to deliver a reliability standard, it is somewhat sceptical on how this might be achieved, instead arguing that “whatever reliability standard is chosen, Regulators and TSOs should compute it with methodologies and tools that are publicly available.” A second objective of this paper is

4 Northern Ireland is part of the SEM discussed below, leaving Great Britain, GB, as the rest of the UK.
to assess how this might best be done, guided by the principle of addressing the missing market problem.

3. Missing money and missing markets
Given the need to instantly balance supply and demand in the electricity system, ensuring short-term security of supply is normally an obligation placed on the SO, while longer term capacity adequacy is often the subject of regulatory and political concern. EU electricity markets are now liberalized, and generation is, for the most part, not subject to traditional utility regulation, but to normal competition policy both domestically and under the scrutiny of DG COMP. If investment decisions could be solely guided by strictly commercial decisions and if markets were not subject to policy interventions or price caps, it is plausible that capacity adequacy could be delivered by profit-motivated generation investment without explicit policy guidance. For this to be the case, investors need confidence that the revenue they earn from the energy markets (including those supplying the ancillary services that the SO needs to ensure short-term stability) will be adequate to cover investment and operating costs.

If this revenue is not adequate, there is a “missing money” problem (Joskow, 2013), but if it is adequate but not perceived to be so by generation companies or their financiers, then there is a “missing market” problem (Newbery, 1989). Missing money problems arise if price caps are set too low (below the Value of Lost Load, VoLL), or ancillary services, such as flexibility, ramp-rates, frequency response, black start capability, etc. and/or balancing services are inadequately remunerated, or transmission access charges are inefficiently high (important in distorting exit decisions), and/or, energy prices are inefficiently low. Inefficiently low wholesale prices seem less likely as the normal problem is one of market power raising prices above their competitive level, and prices are not necessarily inefficiently low just because there is excess capacity.

Missing markets create problems if risks cannot be efficiently allocated with minimal transaction costs through futures and contract markets, or if important externalities such as CO₂ and other pollutants are not properly priced. The concept of missing markets can be usefully extended to cases in which politicians and/or regulators are not willing to offer hedges against future market interventions that could adversely affect generator profits. These arguments have been extensively covered in the literature, recently in the Symposium on ‘Capacity Markets’, (Joskow, 2013; Cramton, Ockenfels and Stoft, 2013). Almost all the discussion about capacity mechanisms concentrates on the missing money problem and whether the various market and regulatory/political
failures are sufficient to justify a capacity mechanism, and if so, what form it should best take.\textsuperscript{5}

Both the missing money and missing market failures have risen in salience as renewable electricity targets have become more ambitious at the same time as the EU Emissions Trading System has failed to deliver an adequate, durable and credible carbon price, and as such is under constant threat of reform. Absent a futures market with a credible counter-party it is hard to be confident that future electricity prices will be remunerative for unsubsidized generation, and harder to convince bankers or shareholders of the credibility of investment plans based on forecast revenues. If renewables continue their planned increase in market share mandated by the EU \textit{Renewables Directive (2009/28/EC)} they will depress average energy prices. This does not of itself give rise to an adequacy problem, although utilities may justifiably complain that their past investment decisions have been partially expropriated by unanticipated political actions. However, it increases the demand for existing services such as primary reserves, fast frequency response and inertia and may also increase the need for additional ancillary services. If these services are not yet adequately defined and/or their future prices are hard to predict there is a missing market problem. If these services are underpriced by SOs whose powers of balancing supply and demand may be met by administrative or regulatory means (e.g. by requiring those connecting to the grid to make some of these services available as part of the grid code), there is a missing money problem. In either case these may precipitate a capacity adequacy problem.

Newbery (2013) documents the analysis that led to the UK’s EMR and its embodiment in the \textit{Energy Act 2013}, which was primarily designed to rescue the UK’s failing attempt to meet its renewable energy targets at least cost. Renewable electricity suffered from both the missing money and missing market problems, as support was provided by Premium Feed-in Tariffs via Renewable Obligation Certificates (ROCs). Their value depended on future electricity prices and the supply-demand balance for the ROCs, both potentially volatile. While fossil generation enjoyed a natural hedge in that electricity prices mirrored fossil fuel prices (see Newbery, 2013, fig 2), renewables and nuclear power, whose fuel costs are zero or very small, are exposed to the full volatility of electricity prices (Roques et al., 2006; 2008), which increases risk unnecessarily and hence raises the cost of capital, the major part of the total cost. The EMR instead proposed long-term Contracts for Differences (CfDs) with fixed indexed strike prices to solve both the missing money and the missing futures market problems.

3.1 Market failures in delivering reliability
Before the electricity industry was liberalized and unbundled, the SO had ownership control of generation and transmission and was responsible for both system security and adequacy. Planned investment ensured that both capabilities would be assessed, which was also much easier when essentially all plant had (at least in aggregate) a predictable and controllable output. The main security aspects were handling very short-run increases in demand (notably during intermissions in major sporting events when consumers all simultaneously switch on their electric kettles) or those caused by the loss of a large station or transmission link. The standard approach was to specify a reserve margin and ensure adequate short-run flexibility by the choice of plant type. Thus the Central Electricity Generating Board, CEGB, that pre-dated the British restructuring of 1989, computed the required gross reserve margin at 19% based on a Loss of Load Probability (LoLP) calculation and a reliability measure (disconnecting some consumers in three winters over a 100 years, decided in 1962 (Bates and Fraser, 1974, p122). It built pumped storage systems to provide fast response, peaking capacity and to use surplus night-time nuclear power (Williams, 1991), as well as jet-derivative gas turbines for fast ramping.

With liberalization and unbundling all these security services had to be separately procured by the SO. Some had to be remunerated, others, such as inertia and the additional security offered by interconnectors, came at no cost to the SO. The SO now had to manage transmission constraints using price signals (distorted through a failure to adopt efficient locational marginal pricing) dealing bilaterally with generators out to maximize profits and typically possessing short-run market power. In addition the SO had to secure various ancillary and balancing services through spot and contract markets, just when the challenges of handling increasing volumes of intermittent and less predictable wind were increasing (Newbery, 2010, introducing “Large-scale wind power in electricity markets” in that issue of *Energy Policy*; Newbery, 2012).

Although the delivery of security services is a public good provided by the SO, these services can be procured through some combination of forward, prompt and real-time markets and contracts. As problems of intermittency increase, so does the challenge of ensuring that these services are efficiently priced and procured (Pöyry, 2014).

The Single Electricity market (SEM) of the island of Ireland is probably at the forefront of addressing these problems, as it is a small moderately isolated system in which individual power stations are large and lumpy relative to peak demand (up to 10%) and the system is being adapted to handle up to 70% non-synchronous wind penetration.\(^6\)

One (implausible) solution would have price signals varying over very short periods of 

\(^6\) Conventional rotating generation turbines are synchronised to the grid frequency and have substantial inertia, so that if there is a momentary loss of supply, that inertia prevents the frequency falling too fast. Wind power has effectively no inertia, which has to be provided in some form to maintain frequency within acceptable limits.
time. A sudden fall in frequency caused by a sudden increase in demand relative to supply means that the value of power in the next cycle (1/50th of a second) has increased, and the speed of response is key to minimizing the disruption. Figure 1 lists the various products and their time domain that the SEM Committee defined in their consultation document (SEM, 2014), grouped into the three categories of Synchronous Inertial Response (SIR, usually delivered by the inertia of synchronized generators), operating reserves (primary, POR; secondary, SOR; and tertiary, TOR), and Ramping or Replacement Reserves (RR), with some overlap between them.

If prices were to move in response to instantaneous system conditions, then it would be potentially profitable to have the capability to respond on the appropriate time scale. In practice, market designs vary in their granularity, with the most flexible having 5 minute settlement periods (Australia). Typically Continental balancing markets have a 15 minute settlement period, while GB and the SEM have a half-hourly settlement period in the Day Ahead Market (DAM), and most Continental power exchanges and the EUPHEMIA auction platform have hourly resolution in the DAM. Increasing granularity improves the accuracy of the temporal pattern of price signals to guide the choice of flexibility, but runs up against the practical constraint that the system state requires a finite amount of time to re-estimate, probably of the order of minutes, while the transaction costs of dealing at high frequency make very short-term markets illiquid.

Given the inability and absence of energy markets at this level of time granularity, new products are needed to supplement existing products. The SEM Committee consulted on how to procure them in early 2015, given that the supply side for some

---

7 Also suggested by Mott MacDonald (2013). California is interested in 5 minute granularity.
services is highly concentrated, raising market power issues. In this case many markets are missing as their procurement is still undecided, making it hard to estimate their future revenues. When choosing what type of generating plant to build, investors have a choice of characteristics with inevitable trade-offs: flexible plant with high ramp rates is either more expensive or less efficient than less flexible plant with lower ramp rates, but unless investors can forecast the revenue from selling these security services, it is hard to make efficient plant choices.

The classic public good problem facing the regulators and the SO is how to value these various services, given that they are provided in bundles of varying proportions by different technologies. Some of the services reduce LoLP, and should be informed by the VoLL, but others reduce the need to curtail renewable generation, which is a political objective of uncertain value.

3.2 Defining and measuring reliability
Eurelectric (2015) provides a useful summary of the various ways used to measure reliability. Most EU electricity systems specify the “Loss of Load Expectation” (LoLE), which for most and for GB is three hours per year. This is a forward looking measure, that taking a representative and large number of possible outcomes (of weather, plant reliability, demand, etc.) for some future period, the electricity system should perform better than averaging “Losses of Load events” of three hours per year. Clearly there is a difference between losing load for the entire population and controlled disconnections for a modest number of consumers, and Eurelectric (2015) argues that a better measure is the Loss of Expected Energy (LoEE) measured in MWh/year. National Grid (2014a) uses the cost of LoEE to determine the procurement amount for the capacity auction.

The former CEGB had a standard of disconnections in three winters in 100 years or 3%, while Belgium had a 1% standard and Spain a 5% standard (Webb, 1977). This was often translated into a gross reserve margin (Transmission Entry Capacity less peak load) required to deliver the reliability standard. Thus the CEGB set it at 19% , made up of a de-rating factor of 11% and remaining outage risks of 6%, so that the de-rated reliability margin was 7% (6%/0.89). Such quantitative measures sit well within the planning framework of a centrally controlled electricity system pre-liberalization.

The other more market-oriented approach to reliability is to specify the Value of Lost Load, VoLL, the amount that consumers should be willing to pay to avoid disconnection. In a future with sufficiently smart meters, consumers would be able to sign a contract with the electricity supplier stating the maximum amount they would be willing to pay for each tranche of firm power, with the smart meter disconnecting appliances at each specified price point, leaving presumably some lights and electronic equipment until last. As such that would represent the complete private good market solution to the problem and would avoid the missing money if not the missing (future)
market problem, provided the short-run prices were efficiently set at their efficient scarcity value. This is the sum of the System Marginal Cost (SMC) plus a Capacity Payment, CP, where

\[ CP = \text{LoLP} \times (\text{VoLL} - \text{SMC}), \tag{1} \]

The relationship between the security standard and the VoLL is symmetric, in that if capacity investment decisions are based on revenues determined by (1) and the VoLL is pre-determined, then the resulting capacity will give rise to a LoLE. If the standard is as a predetermined LoLE, the cost of new capacity implies a cost of delivering the LoLP and hence the VoLL.

Britain has followed both models. The English Pool set the VoLL at £(2012)5,000 (€6,250/MWh at £1 = €1.25), letting the market determine capacity. After the Pool was replaced with an energy-only market in 2001, the Department of Energy and Climate Change (DECC) specified the LoLE. National Grid (2014a) deduced the 2018 VoLL as £(2012)17,000/MWh (€(2012)21,250/MWh), higher than direct estimates of the willingness to pay to avoid disconnections (London Economics, 2013).

3.3 Can energy-only markets deliver adequate reliability?

One completely legitimate case for a capacity payment is that if generators are required to bid their Short Run Marginal Cost (SRMC, mostly fuel costs), as under the Bidding Code of Practice of the SEM (SEM, 2007), they will fail to recover their fixed costs without such an addition. The Electricity Pool of England and Wales also added the CP of (1), but allowed generators to offer an unrestricted supply function (which, given their market power, was often above SRMC, Green and Newbery, 1992; Newbery, 1995; Sweeting, 2007). In this period of benign liberalization, high prices led to considerable entry and an excessive reserve margin.

In the energy-only market envisaged by the TEM, generators will offer supply functions that should reflect the scarcity value of electricity (and their degree of market power). Figure 2 shows the day-ahead price duration curves for several European power exchanges in 2012. What is striking is that most exchange prices do not exceed €200/MWh, and even the most peaky, France, only does so 0.25 of 1% of the time (about 22 hours per year). Given that the VoLL in the English Pool until 2001 was €(2012)6,250/MWh and the current implied VoLL in GB is €(2012)21,250/MWh (both at £1 = €1.25), these prices indicate a low LoLP or high reliability. Given existing capacity levels that is a reasonable inference, but the problem again is one of missing (futures) markets. Investment lags in delivering capacity adequacy are 2-4 years for peaking plant (longer for most base-load plant), beyond the time horizon of adequately liquid futures markets (and they only offer one year’s hedge).
Figure 2 Price Duration curves of day-ahead hourly prices, 2012
Sources: MIP (Market Index Price) and NL prices from APX, Germany and France from EEX

On the other hand, Figure 3 shows that the 2008 balancing buy prices in the energy-only market that replaced the Pool were considerably peakier than the old Pool prices (which included an explicit CP and also probably reflected more market power). Thus energy-only markets can reflect scarcity, and properly calculated capacity payments may be very low if the reserve margin is adequate as LoLP is roughly exponential in demand less derated capacity (Newbery, 2005). However, by 2013-14, the GB Balancing Mechanism had a price duration curve quite similar to those shown in Figure 2, with prices above €200/MWh for less than 0.25 of 1% of the time, and well below the French day-ahead price duration curve.

Thus one might conclude that energy-only markets (which include balancing markets) can deliver sufficiently sharp scarcity prices that should signal the profitability of adequate new investment, provided all the other security services are adequately remunerated (i.e. resolving any of those missing market problems). This might be plausible if all investment decisions were taken on commercial grounds as in the 1990s, that prices were not capped, that the policy environment were predictable and stable, and that either liquid forward market existed for a reasonably fraction of the proposed plant life (i.e. 20+ years ahead of the final investment decision) or credible long-term power purchase agreements could be signed with credit-worthy counterparties. Unfortunately, hardly any of these conditions hold in the TEM.

---

8 For a description of the British Balancing Mechanism see Newbery (2005).
3.4 Market, institutional and political/regulatory failures

While price caps are set at rather low levels in the US, exacerbating the “missing money” problem, there are also, if much higher, price caps in EUPHEMIA (for day-ahead at €3,000/MWh, a price that France has hit on numerous occasions). The lack of forward markets and long-term contracts might not be so critical if the future were reasonably predictable and stable, but this is far from the case at present. EU Climate Change policy is failing, in conflict with Renewables Directive (2009/28/EC), and surely ripe for as yet uncertain reform. Large volumes of unreliable renewables increase the need for flexible reserves, which in the past came from obsolescing plant, mostly oil or coal. This plant is now being decommissioned because of the Large Combustion Plant Directive and the Integrated Emissions Directive.

Increasing renewables (mainly wind and solar PV) add little to reliable capacity, as it is unavailable on still cold dark winter nights, but reduces average wholesale prices. If the average capacity factor of on-shore wind is 25%, then the GB target of 30% of electricity from wind requires capacity of 30/25 times or 120% of average demand. In windy conditions that would often displace all conventional plant and could lead, under present subsidy structures, to negative prices.

Intermittent generation increases the need for additional flexible plant that can be called up at short notice if the wind falls or the sun fades. In addition, new plant will be needed to replace retiring plant (not just coal, but in the UK, France and Germany, substantial volumes of nuclear plant as well). That plant will need considerably higher...
prices when renewable output is low than has been recently experienced. Even if the carbon price is currently low, the EU is committed to an 80% reduction in Greenhouse Gas emissions by 2050. Coal is twice the carbon intensity of gas, so utilities are unlikely to build durable (40-60 year) coal-fired plant that would face tight future emissions limits, leaving gas-fired plant as the only alternative. Unfortunately, crashed electricity prices and high gas prices precipitated by the closure of Japan’s nuclear fleet has made their economics very unattractive.

The UK introduced a carbon price floor in the 2011 Budget (HMT, 2011) that would support the price of CO₂ at £16/tonne in 2013, rising to £30/tonne (£35/tonne) in 2020, and projected to rise to £70/tonne by 2030 (all at 2009 prices). This threatened the operation of existing coal-fired plant. As an example of policy instability, the 2014 Budget froze the carbon price floor – clearly an instrument subject to the whim of chancellors creates additional investment uncertainty. It would be a brave politician who trusted these markets to deliver reliability.

4. The GB capacity auction
In response to the looming capacity crunch and other market failures set out in Ofgem (2009) and subsequent DECC consultations (Newbery, 2012a), the UK passed the Energy Act 2013 setting out the EMR, which includes a Capacity Mechanism to ensure adequate capacity. GB now has the Secretary of State for Energy & Climate Change, advised by DECC, deciding how much capacity is required. The capacity auction is a single-price descending clock auction with a demand schedule as shown in figure 4. National Grid as SO was charged to recommend the target volume of capacity to secure four years after the auction (which was termed the T-4 auction).

National Grid (2014a) chose the amount to procure balancing the cost of additional capacity against the cost of the Loss of Expected Energy, as shown in Figure 5. National Grid (2014a) projected that the auction clearing price would likely be set at the Cost of New Entry (CONE), estimated at £49/kWyr. This was the missing money a Combined Cycle Gas Turbine (CCGT) might need given its revenues from all other markets and after paying the Transmission Network Use of System (TNUoS) charges. These range from £30/kWyr (in NW Scotland), to negative (-£5/kWyr in Cornwall) (National Grid, 2013). Entrants are given 15-year indexed contracts, while existing plant receive one-year contracts to defer exit decisions until the next auction.

The missing money can be estimated from the VoLL (£17/kWh) less the maximum the SO pays for balancing actions (£6/kWh) to give £11/kWh, times 3 hrs LoLE, or £33/kWyr. The effective cap in the Balancing Mechanism is £9,999/MWh would reduce the missing money to £21/kWyr.
The auction design was best-practice (Newbery and Grubb, 2015) but flawed in requiring the SO to advise the minister on the procurement amount. The SO stands accountable if “the lights go out” but does not pay for the capacity. The minister wishes to avoid newspaper headlines predicting blackouts resulting from his decision. Both
argue for excess procurement. DECC appointed an independent Panel of Technical Experts (PTE) to comment on the analysis, and they made a number of strong (but, for the 2014 auction, ineffective) criticisms.

4.1 Criticisms of the capacity to procure
The PTE first criticized the terminology of “Loss of Load” as emotive and misleading. The GB regulator, Ofgem, defines a “Loss of Load event” as one in which market demand exceeds market supply and as such the SO has to intervene to balance the system. For that purpose the SO can call on a range of increasingly expensive options: asking generators to temporarily exceed rated capacity; invoking ‘new balancing services’; cutting interconnector exports to zero, requesting imports; reducing voltage (“brown outs”), before finally resorting to selective disconnections. The crucial point is that these actions cost less, often much less, than VoLL and hence bias the unserved load cost and the target capacity in figure 5 upward.

Successful Capacity Market Units in the auction receive a Capacity Agreement which requires them to be available in “stress events” signaled four hours ahead. DECC (2014, §391) defines these events as “any settlement periods in which either voltage control or controlled load shedding are experienced at any point on the system for 15 minutes or longer. … Periods of voltage control or load shedding resulting from failures or deficiencies in the transmission or distribution systems are not considered as stress events.” However, these “notices of inadequate system margin” are issued “based on the available capacity (declared ‘maximum export limit’ (MEL) minus transmission system demand and reserve for response capacity.” DECC (2014, box, p107.)

National Grid (2014a) chose the amount to procure using a Least Worst Regrets approach as it was unwilling to attach explicit probabilities to the various scenarios considered. The result of overvaluing the cost of “Loss of Load” is to increase the capacity at which the Least Worst Regret cost schedule is minimized (figure 5). “Slow progression” reaches a cost minimum at 53.3 GW for 2018-19 delivery. The net procurement target is 53.3-\(w-x-y-z\)-0.4 GW, where \(w, x, y\) and \(z\) refer to various distributed energy resources and opt-out plant. The 0.4 GW is existing short-term operating reserve.

The PTE (DECC, 2014a) strongly criticized National Grid for assuming no net imports in stress periods, despite 3.75 GW interconnection capacity and potential new capacity of 2.25 GW that might be available by 2018-19. This seemed perverse, as all

---

9 “The new balancing services are Demand Side Balancing Reserve (DSBR) and Supplemental Balancing Reserve (SBR).” National Grid announced its tender for these new services on 10 June 2014 (http://www.nationalgrid.com/uk/electricity/additionalmeasures ).

10 The CEBG estimated that voltage reductions reduce load by 7½% in the 1970s (Bates and Fraser, 1974) but National Grid now estimates only 1½% in the absence of firm evidence.
parties (Ofgem, DECC and National Grid) agreed interconnectors increased security. Three reports commissioned by these parties (Pöyry, 2012, 2013; Redpoint, 2013) argued that interconnector capacity could displace domestic capacity by 50-80% of its value. Even DECC’s Final Impact Assessment, published just before the procurement decision (DECC, 2014b), estimated the amount of interconnector capacity to include in the total procurement amount at 2.9 GW. Ignoring interconnectors could move the auction clearing price from the cost of new entry of £49/kWyr to that set by existing plant (maximum of £25/kWyr), increasing the auction cost by £1.3 billion per year.

Ignoring interconnectors seemed particularly perverse as the TEM aims to integrate markets across borders. Market coupling already dispatches GB Continental interconnectors in the EUPHEMIA day-ahead market (DAM). Interconnector flows already reflect willingness to pay in the DAM, and will soon do so in the intra-day and real time markets when network codes are agreed.

4.2 Possible consequences of excessive capacity procurement
Excess procured capacity will lower future wholesale prices with a number of effects, not all immediately obvious. First, lower prices reduce the revenue new entrants can expect from the energy markets, increase the CONE and raise the auction price. Second, it undermines the old market design in which investment in conventional generation was at the discretion of private companies making commercial decisions. No company would invest in conventional generation without a capacity agreement given its large disadvantage compared to those with agreements. The amount of new plant will therefore be entirely determined by the minister, ending a key element of the liberalized market. All non-fossil generation will also be granted long-term Contracts for Difference (CfDs) under the EMR, moving GB to the Single Buyer model ruled out in earlier EU Electricity Directives.

Third, lower prices increase payments to low-carbon CfDs, which receive the difference between the contracted strike and wholesale price. As the Government limits total renewables payments through the Levy Control Framework, the perverse effect is to support less renewable electricity, although the EMR was designed to remove obstacles to meeting the renewables target.

Fourth, the commercial case for interconnectors depends on price differences, with GB typically importing from cheaper Continental markets. Lower GB prices reduce arbitrage profits, undermining the investment case for the additional interconnectors when they are increasingly needed to balance growing intermittent generation across wider market areas. Ignoring interconnectors risks a self-fulfilling but expensive policy of autarky.
Fifth, although the future wholesale price will be lower, offsetting possibly a large part of the consumer cost, it will be hard to convince consumers of this. They will see the gross cost estimated at 53.3 GW x £49m/GWyr = £2.6 billion per year.

Finally, on 2nd December 2014, after the PTE had published its critical report and the Secretary of State had decided on the procurement volume, but before the auction on 18th December, the Treasury’s National Infrastructure Plan announced that interconnectors would be eligible for the 2015 capacity market.\footnote{See \url{https://www.gov.uk/government/collections/national-infrastructure-plan}} It would have been easy to have left room for interconnectors (e.g. adding another element to the -w-x-y-z-0.4 GW deduction from the target volume) and lower the net amount to procure.

4.3 The outcome of the 2014 capacity auction

The auction cleared at £(2012) 19.40/kWyr (National Grid, 2014b). The auction produced several surprises. First, the auction cleared at less than 40% of the predicted CONE value of £49/kWyr (although close to the missing money estimated above assuming a balancing cap of £9,999/MWh). The estimated CONE was based on new entry of CCGT, and two CCGTs entered, supplying about 60% of the total 2,795 MW new entry. Second, the next largest (28%) entry category was OCGT/ reciprocating engines, average size 11 MW. The third largest contribution (6%) was from unproven Demand Side Response (DSR, all with a one-year contract, other new entrants have 15-year contracts).

One might expect that DSR and OCGTs would require a lower strike price, particularly as they can contribute to significantly reducing TNUoS charges if they are embedded with major loads, but the low price that CCGTs were willing to accept is surprising, and may be based on optimistic views of gas prices (which were expected to decline by the time of the auction) or high balancing prices. National Grid announced its tender for new balancing services on 10 June 2014,\footnote{At \url{http://www.nationalgrid.com/uk/electricity/additionalmeasures}} reducing the extent of the missing money for these services after DECC had published its estimate of the net CONE.

The final point is that the auction demonstrates the value of market-based methods of revealing entry costs, and the danger of leaving such decisions to SOs or regulators (as in the SEM,\footnote{See \url{http://www.allislandproject.org/en/cp_current-consultations.aspx?article=75c548a7-34ee-497c-af6d-262f8aa0062df}} where the regulators calculate the cost of Best New Entry and set it at a high price).

5. Biases in capacity auctions and energy-only markets

The arguments above strongly suggest that if procurement decisions are left to politicians advised by the SO, they will err on the high side, and tend to ignore supplies from outside their control area (over interconnectors). Their caution is exacerbated by the emotive and
misleading terminology of “Loss of Load”. Some of these shortcomings can be addressed by requiring the SO to cost and quantify the actions that are taken in stress events that fall short of controlled disconnections. Delegating the decision to independent agencies, perhaps to an Independent Planning and System Operator could depoliticize the decision (but might not remove the fear of disconnections through inadequacy, nor the bias of not paying for capacity).

There is a more fundamental problem in that if future energy prices are competitively delivered and if all security services (ancillary and balancing) are properly priced, the missing market and missing money problems can both be addressed by offering suitable hedging contracts, of which the auctioned capacity agreement is an excellent example. Price caps could be replaced by reliability options or one-way CfDs that have a high strike price, and which allow consumers or their suppliers to hedge against high prices while allowing the spot and balancing market prices to reach scarcity levels needed for efficient actions (in demand reduction and interconnector trade) (see e.g. Vásquez et al, 2002; Bidwell, 2005).

Now consider the costs of under or over-specifying the amount to procure. Over-procurement, as noted above, risks depressing future prices and hence reducing future energy and ancillary service revenues, requiring a higher auction price in compensation. While addressing the missing markets problem it risks amplifying the missing money problem. In contrast, under-procurement leads to expectations of higher future prices, requiring a lower capacity auction bid as the capacity agreement does not preclude earning revenues in all the energy markets. If the price is very low, investors may conclude that investing without a capacity agreement has relatively low risk, particularly as the design of the GB auction offers a T-4 contract of 15 years for new plant, but successive T-1 contracts of one-year for existing plant at the same clearing price, for which speculative plant would be eligible. A signal to err on the side of under-procurement would be underwritten by the ability to true up closer to delivery, reducing risks, as any over-procurement would merely delay the moment at which more capacity was needed in the auction, and should limit the period of inadequate revenue to a year or so.

6. The need for regional coordination
Eureletric (2015) argues strongly for a design of regional capacity markets which places the obligation on generation (or demand) regardless of national location. This requires a common regional capacity adequacy assessment and no double payment (i.e. if capacity has an agreement from country A it would be denied one from B and would be excluded from B’s capacity assessment, subject to adequate interconnector capacity from B to A).

While this may be a desirable long-run objective, it is hardly compatible with the existing Third Package. Meanwhile countries will have to decide how to treat
interconnectors connecting possibly very different (or no) capacity markets. Provided the auction platform can accommodate efficient scarcity prices (i.e. provided at least the intra-day and balancing markets are not capped at too low a level), then trading over any interconnector will only benefit a country that ensures that the relevant prices are efficient, as in (1). Suppose A and B trade but A has a higher VoLL than B, and hence a larger reliability margin. If A and B both have stress events, A can outbid B to secure imports, with B accepting a higher LoLP reflecting its lower willingness to pay to avoid disconnections, while ensuring that domestic consumers were insulated from these high trading prices through the contract coverage provided by the Reliability Options.

The logic of making adequacy as close as possible to a private market good (through allowing efficient pricing) is that there can be gains from trade for the efficiently priced market even when market designs are different. If prices are inefficient in B, then it is they who lose, not A. That provides incentives to reform and avoids the need for politically fraught agreements on harmonization.

7. Conclusions and policy implications

Missing money and missing markets provide compelling reasons for a capacity payment in competitive electricity markets dominated by politically determined and subsidized unreliable generation and where investors lack confidence in future revenues. Capacity auctions (GB provides a good example) address the missing money problem and part of the missing market problem (the missing futures markets), which still needs efficient solutions - markets, auctions and procurement contracts - for location, flexibility, etc. needed to deliver security. More complex category auctions may also be the best way of procuring these services. The part of the adequacy debate that has been neglected is how to, and who should, determine the amount and type of capacity to procure (generation, DSR, interconnection), a problem that is exacerbated by misunderstandings over what a “Loss of Load” event means and what it might cost.

This paper argues that this neglect biases towards over-procurement, which leads to a self-fulfilling prophecy that merchant generation investment can no longer be relied upon. Perversely, this exacerbates the missing money problem that capacity auctions were designed to address. The bias is further exacerbated by failing to address some of the missing market problems that have also been neglected in the debate.

Whether or not interconnectors should be included in auctions is less important than that their contribution should be recognized in determining the procurement amount. All British interconnectors are HVDC controllable links whose flows can be rapidly reversed and as such could provide extra imports at short notice, but they can also impose sudden large loads on the GB system if they switch to exporting. The UK Government, possibly under pressure from DG COMP over State Aid concerns after the PTE had
published their criticisms (DECC, 2014a), decided to include interconnectors in the next auction for 2019-20 delivery, and consult on how to determine their reliable capacity.

There remain a number of policy issues to resolve, not least how the European auction platform EUPHEMIA will determine the direction of flows close to real time, when stress events that the capacity auction was designed to address are likely to emerge. EUPHEMIA has a €3,000/MWh price cap on the DAM, well below the VoLL. It has not yet (early 2015) fixed price caps for intra-day and balancing actions. If prices in the real time European markets could properly reflect scarcity, and if the GB market could deliver the true scarcity prices to EUPHEMIA (including the CP of equation (1)) then good market design and pricing would deliver efficient solutions, and other countries with less good pricing would lose out, motivating them to improve their market design. Price caps hinder this aim, and instead good rules will be needed for out-of-market actions when price caps are reached, and/or markets no longer determine flows, and SOs have to intervene. These rules or bilateral agreements between the SOs at each end of interconnectors are currently lacking or incomplete. If these problems can be addressed, then the more demanding task of harmonizing regional capacity markets may be unnecessary.
References


Ofgem (2009). *Project Discovery Status Report* at [http://www.ofgem.gov.uk/markets/whlmkts/discovery/Pages/ProjectDiscovery.aspx](http://www.ofgem.gov.uk/markets/whlmkts/discovery/Pages/ProjectDiscovery.aspx)


