Questioning the EU Target Electricity Model – how should it be adapted to deliver the Trilemma?

EPRG Working Paper 1617
Cambridge Working Paper in Economics 1634

David Newbery

Abstract
Britain considers the energy-only EU Target Electricity Model (TEM) wanting in delivering the trilemma of reliability, sustainability and affordability and argues that a capacity auction with long-term contracts for new entrants is the least-cost solution compared to relying on expectations of future prices to deliver adequate generation and demand side response. The Energy Union argues against feed-in tariffs (FiTs) for renewables, pressing for premium FiTs (pFiTs), just as GB has abandoned PFiTs in favour of FiTs. This paper draws on the GB experience of Electricity Market Reform before and after the 2015 change of government, to highlight promising resolutions of the energy trilemma, and the problems that have arisen between the diagnosis of the problem and the delivery of solutions. It sets out the theory and practice of delivering capacity, energy and quality of supply, gives a brief history of GB electricity from the CEGB to its current unbundled, liberalized and privatized structure. That sheds light on the trilemma problem and discusses possible solutions. The island of Ireland Single Electricity Market reforms illustrate the problem and possible answer of how best to deliver quality of service with high intermittency.

Keywords Reliability, sustainability and affordability, capacity auctions, contract design, renewables

JEL Classification D47, H23, L94, Q48, Q54

Contact dmgn @ cam.ac.uk
Publication 2016
1. Introduction

Britain has taken a careful look at the energy-only market model that underpins the EU Target Electricity Model and has found it wanting in delivering the Trilemma of reliability, sustainability and affordability. On reliability or security of supply, Britain argues that a capacity auction with long-term contracts for new entrants is the least-cost solution compared to relying on expectations of future market prices to deliver adequate generation and demand side response in a timely fashion and at acceptable financing costs. On sustainability, or decarbonization, the Energy Union (EC, 2015) is now arguing against supporting renewables with the classic feed-in tariff (FiT) and instead is pressing for premium FiTs (pFiTs), just as GB has abandoned PFIs in favour of something much closer to FiTs. While the EU is beginning to accept that the EU ETS is an inadequate instrument to guide low-carbon investment in the electricity supply industry (ESI), GB has enacted a carbon price floor, which, although not by itself a credible signal, it underwrites long-term contracts for low-carbon investment.

This paper draws on the GB experience of Electricity Market Reform before and after the 2015 change of government, to highlight promising resolutions of the energy trilemma in the ESI, and the problems that have arisen between the diagnosis of the problem and the delivery of solutions. Section 2 sets out the theory and practice of delivering capacity, energy and quality of supply to the wholesale market and final consumers, followed by a brief history of the evolution of the GB ESI from a vertically

1 Paper building on the presentation to “A New Model for Electricity Markets? Towards a Sustainable Division of Labour between Regulation and Market Coordination” held in Paris–Dauphine on 8-9 July 2015 and subsequently published as “Tales of Two Islands - Lessons for EU Energy Policy from Electricity Market Reforms in Britain and Ireland”, Energy Policy, available at http://dx.doi.org/10.1016/j.enpol.2016.10.015. The author is a member of the Panel of Technical Experts advising DECC on capacity auctions and a member of the SEM Committee of the island of Ireland but is writing solely in his academic capacity, using only published material, and the views expressed here do not reflect the views of either organisation. I am indebted to Anette Boom, Thomas Greve and an anonymous referee for helpful comments, but I am responsible for remaining errors.

2 The Target Electricity Model was much influenced by the Nordic markets, which with large volumes of storage hydro and ample capacity and interconnection is arguably one where the energy-only market model might be quite suited.

3 There is a list of acronyms at the end.
integrated centrally planned state-owned company to its current unbundled, liberalized and privatized structure and the problems this presented in delivering the trilemma. Section 4 describes the diagnosis and proposed solution to that problem, which were not peculiar to GB. Section 5 therefore studies the Single Electricity Market (SEM) of the island of Ireland, which faces higher intermittency with a lumpier and more isolated system than almost any other country, and which therefore raises the question of how best to deliver quality of service with high intermittency. Section 6 draws lessons from the experience so far and implications for future market and subsidy reforms.

2. Pricing electricity: from central planning to liberalized markets

Electricity appears the archetypical homogenous commodity that underlies microeconomics – all electrons look the same – but that is deceptive. Capacity (MW) limits peak demand, energy (MWh) and power (MVA) vary over time and space, and quality of service (QoS) includes stability of frequency and voltage, while the phase angle affects the ability to extract power. QoS requires a variety of ancillary services supplied by generation or demand connected to the grid (reserves, reactive power, frequency response, black start capability, etc.) and in turn requires grid codes/standards on those connected (fault ride-through, ability to remain connected up to a specified rate of change of frequency, etc.). Generation plant may have fixed start-up costs, limits on the rate at which it can ramp up to full power, varying efficiencies at different plant loads, minimum stable generation output, minimum down-time between operations, etc. The transmission system has limited capacity to move power between nodes and the system has to be able to withstand the loss of at least one of the largest components (the largest single infeed - generation unit or interconnector - or the largest transmission link, the N-1 constraint).

Determining the least-cost dispatch to meet time and space varying demands is therefore difficult as it is a non-convex problem with strong intertemporal dependencies. In centrally dispatched systems, the System Operator (SO) typically solves this with a Mixed Integer Program optimizing over a future period (a week for thermal systems, longer for hydro systems), to determine the optimal security-constrained dispatch (including necessary reserves and other ancillary services). The dual of this optimal quantity program is the scarcity value of electricity at each node (the Locational Marginal Prices, LMPs). The theory, set out by Schweppe at al. (1988), has been implemented in large areas of the U.S. as the Standard Market Design. In the pioneering region of PJM,4 LMPs are recomputed every five minutes.

In a vertically integrated system in which transmission and generation are under central control within a single company (the standard model for much of the developed world until the 1980s) investment decisions in transmission and generation could be

---

4 Named after the original Pennsylvania, New Jersey and Maryland interconnection, but now much wider.
coordinated to deliver least-cost delivery of power to the grid supply points at which the regional distribution network operators (DNOs) connect. These DNOs were usually under different management (although were often under the same state ownership) and were often charged on their specified peak power in each month, and then a variable energy charge, with higher prices for exceeding the specified peak. The DNOs then translated this Bulk Supply Tariff into charges for different types of consumers (differentiated by voltage level and whether half-hourly metered and with what maximum demand allowed or taken).

Efficient investment planning requires the right type, size, location and delivery date of typically large thermal generation units, with options constrained by access to fuel, cooling water, and a grid connection. Transmission planning had time horizons of 60+ years, and given the constraints on location (local opposition) had long lead times and limited choices, while locating generation assets was in principle easier (unless tightly constrained as with hydro-electric dams). Nevertheless, tight coordination of the location and timing of generation and transmission offered the prospects of considerable saving – important when nuclear power stations need to come off-line to refuel periodically and the grid needs adequate capacity to wheel replacement power in from other sources.

State-ownership provided access to (often very) low-priced capital but limited incentives for efficient investment (operation was usually much better as run by engineers and easy for the SO to monitor), particularly as the unions had enormous threat power and extracted high rents. Privatization without liberalization risked monopoly without improved efficiency, liberalization required unbundling to prevent entry deterrence, and unbundling required markets to replace central decision making. Creating suitable markets and ensuring efficient investment and dispatch is difficult, given non-convexities in operation and synergies in investment. Competitive markets can only guarantee efficient outcomes if there are no market failures, and sufficiently dense risk and futures markets for all products supplied and demanded (capacity, energy and QoS). Some market failures can be addressed by charges and/or subsidies, at the risk of political or regulatory failures. Missing futures and risk markets can be replaced by long-term contracts. Natural monopolies (the wires of grid and DNOs) need incentive regulation, which can improve on the poor governance of state ownership, but setting efficient and cost-recovering tariffs is a non-trivial undertaking. The next section illustrates this for Britain.

3. Brief history of the GB electricity sector

Figure 1 shows generation output by fuel and total capacity since shortly before privatization and restructuring in 1990. Around 90% of the conventional thermal generation is from coal, and the share of oil falls rapidly from 7% to 1% in 2002 (the remainder is largely from by-product gases from iron, coke and chemicals). Declared net
capacity now called Transmission Entry Capacity (TEC) is the maximum that the generator requests to be allowed to inject into the grid. An increasing volume of generation (most but not all renewables) connects to DNs, but data for that is only available for the past few years. The story is quickly told: at privatization the UK was supplied by coal and nuclear power with some imports. Shortly after privatization the coal share rapidly declined as nuclear power improved its performance, and with the “dash for gas”, which was all new entry despite the considerable capacity margin (roughly indicated by the gap between capacity and output). At the end of the century consumption fell with deindustrialization and increased demand efficiency, while renewables displaced gas and/or coal, whose shares depended on the very volatile clean (gas) and dark green (coal) spark spreads (the margin between the wholesale price and the fuel and CO₂ cost).

### Electricity supplied by, and capacity of, UK generators, 1987-2014

Fig 1 Electricity supplied and declared net capacity connected to the grid
Source: Digest of UK Energy Statistics, various years

#### 3.1 Privatization and electricity pricing in the Pool

The state-owned companies were replaced by, in England and Wales (E&W), two fossil and one nuclear (initially state-owned) generation companies, with an unbundled grid (initially collectively owned by the regional privatized Regional Electricity Companies, RECs). In Scotland the two vertically integrated companies were sold bundled, while in Northern Island three generation companies were sold with long-term power purchase agreements (PPAs). National Grid and the RECs were regulated and large customers were free to buy directly from the wholesale market, the mandatory gross Electricity Pool, which was centrally dispatched with a System Marginal Price (SMP) set by the
marginal price offered by the most expensive unconstrained generator required, to which was added a capacity payment, CP.

\[ CP = \text{LoLP} \times (\text{VoLL} - \text{SMP}), \]  

where LoLP is the Loss of Load Probability in that half-hour and VoLL is the value of Lost Load (£2016 5,000/MWh). This would give the efficient scarcity price of electricity if the SMP were the system marginal cost, but generators were free to offer any price, only constrained by the threat of anti-competitive remedies.

The sum of SMP and CP gives the Pool Purchase Price, which, with additional ancillary service and constraint costs made up the Pool Selling Price. National Grid as Transmission System Operator (TSO) received offers from all individual generating sets the day before, which were complex multi-part offers with a raft of additional constraints and characteristics, and used the old scheduling algorithm to determine a feasible dispatch. Adjustments during the day were called off the previous day’s offers and charged out to consumers in the Pool Selling Price (Green and Newbery, 1993).

The two fossil generators dominated the E&W Pool and clearly had considerable market power (Newbery, 1995) which they exercised with caution, given the close regulatory scrutiny, until the regulator intervened in 1994 to “encourage” them to divest 6 GW of coal plant to a third generator, which they did in 1996. The resulting triopoly was less constrained in exercising market power, and the price-cost margin continued to widen, in the successful effort to convince outside companies to buy coal plant with apparently attractive profit margins before the dash for gas eroded this market power and lowered prices (Sweeting, 2007). The dash for gas in turn was aided by high Pool prices, low and falling gas prices, and rapidly improving and low capital cost Combined Cycle Gas Turbines (CCGTs). With energy policy under the privatizing Conservative government absent, or rather, left to the market to guide choices, political risk was considered low and substantial entry by “Independent” Power Producers occurred. These Power Producers entered on the back of long-term fixed price contracts with (and often shared ownership by) the RECs, who could pass on their costs to the captive franchise domestic market, and with long-term gas contracts and performance guarantees on the CCGTs, further reducing risk.

As the wholesale market fragmented towards the current Big Six\(^5\) generators, they were finally allowed to buy the supply businesses (originally integrated with the RECs), and hence acquire an internal built-in hedge rather than repeatedly contracting with final customers. If wholesale prices fall but retail prices remain high, generators lose but retailers gain, and the opposite movements of profits up and downstream makes such hedging critical for reducing the risk that was previously internal to the industry when it

---

\(^5\) Centrica, EDF Energy, npower, E.ON UK, Scottish Power and SSE, with domestic market shares of between 12% (Scottish Power) and 25% (Centrica, the original incumbent gas supplier.)
was all state-owned and vertically integrated. After the initial divestment in 1996, the growing gap between cost and price encouraged the Government to replace the Pool with the New Electricity Trading Arrangements, just at the date (2001) when the price-cost margin collapsed under the weight of competition and excess capacity (Newbery, 1998; 2005).

The New Electricity Trading Arrangements were a retrogressive step, replacing central dispatch and the Pool with a self-dispatched energy-only market (abolishing capacity payments) and imposing a two-priced Balancing Mechanism so flawed it has required many hundreds of painfully negotiated modifications to approximate an efficient balancing market. The claimed logic for the reform was that self-dispatch required generators to submit a balanced offer (i.e. output matched by contracts to purchase from consumers or their agents, the retailers) and that required them to contract all output ahead of time, thus removing the incentive to manipulate the spot market (under-contracting encourages sellers to increase the spot price above the marginal cost, over-contracting to reduce the price below marginal cost, Allaz and Vila, 1992; Newbery, 1995).

The initial form of support for renewable electricity supply (RES) was the non-fossil fuel obligation auctions for FiTs, that were very successful in driving down the clearing price in successive auctions, but increasingly under-delivered on investment as there was no penalty for not building after receiving a contract (Newbery, 2012). It was replaced by the Renewable Obligation (RO) Scheme, which issued one RO Certificate (a ROC, essentially a pFiT) for each MWh produced from wind (and varying fractions for each MWh generated by other renewables). Their price depended on supply from RES and demand from suppliers required to purchase an annually specified and increasing share of RES or pay a penalty price that was recycled to those selling ROCs, uplifting the price and value.

3.2 Transmission charging for efficient location
As there was (and still is) a single wholesale price across the whole of GB and as the number of ROCs per MWh delivered had no spatial variation, the only factor influencing location decisions was the transmission network use of system charge (TNUoS), which vary by zone and again is set annually. Generation connected to the grid pays an annual fee for TEC (the G charge) and Industrial and Commercial Load (L) pays on demand in the three half-hours of highest system demand separated by 10 days (the Triad).\textsuperscript{6} Table 1 shows that the sum of the two charges is roughly constant across the country, but the components vary. In 2016/17 the G charges varied from £20.20/kWyr to -£6.09/kWyr,\textsuperscript{7} a

\textsuperscript{6} See http://www.nationalgridconnecting.com/triads-why-three-is-the-magic-number/
\textsuperscript{7} Negative G charges are only paid to G on actual export in Triad periods, and are designed to relieve shortages in import pockets.
range of £26.29/kWyr, while the sum of G + L was £49+/− £4/kWyr. Intermittent
generation pays on average 72% of the G charge in 2016/17.

The TNUoS charges would give efficient location guidance for new investment if
these various charges reflected the appropriate average of the nodal charges in each zone
over the year for the type of generation connected (baseload, mid-merit, peaking or
intermittent). Actually it is only the differences across space that matter as it is the sum of
G+L that is passed through to final consumers (or it would be if there were no link to
external prices through interconnectors). The EU now limits the average G charge to
ensure better harmonization across borders, but this still allows the full range of
locational G (and complementary L) charges. As TNUoS is a fixed charge, it fails to give
efficient dispatch signals, as the underlying LMPs will vary each half-hour (or 5 minutes
in some designs). Perhaps equally important, the current TNUoS charges change annually
and can only be avoided by exit, but there is no guarantee that they give efficient exit
guidance (the connection point may have value to other potential entrants, but may not).
This became important for the capacity auction of 2014.

Table 1 TNUoS charges for an 80% Generator and for Load

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GB G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>max</td>
<td>£20.20</td>
<td>£29.64</td>
<td>£31.06</td>
<td>£31.36</td>
<td>£28.70</td>
</tr>
<tr>
<td>min</td>
<td>-£6.09</td>
<td>-£8.79</td>
<td>-£14.30</td>
<td>-£16.45</td>
<td>-£21.30</td>
</tr>
<tr>
<td>E&amp;W G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>max</td>
<td>£7.48</td>
<td>£8.13</td>
<td>£6.05</td>
<td>£4.63</td>
<td>£2.75</td>
</tr>
<tr>
<td>min</td>
<td>-£6.09</td>
<td>-£8.79</td>
<td>-£14.30</td>
<td>-£16.45</td>
<td>-£21.30</td>
</tr>
<tr>
<td>GB G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>£6.01</td>
<td>£8.88</td>
<td>£7.63</td>
<td>£6.91</td>
<td>£3.51</td>
</tr>
<tr>
<td>E&amp;W G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average G + L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB</td>
<td>£49.09</td>
<td>£47.95</td>
<td>£52.75</td>
<td>£54.94</td>
<td>£65.84</td>
</tr>
<tr>
<td>E&amp;W</td>
<td>£47.50</td>
<td>£47.44</td>
<td>£51.44</td>
<td>£53.82</td>
<td>£63.61</td>
</tr>
<tr>
<td>SD of G+L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB</td>
<td>£4.07</td>
<td>£3.89</td>
<td>£5.11</td>
<td>£5.03</td>
<td>£6.63</td>
</tr>
<tr>
<td>E&amp;W</td>
<td>£2.58</td>
<td>£2.38</td>
<td>£2.63</td>
<td>£2.77</td>
<td>£2.79</td>
</tr>
</tbody>
</table>

Source: National Grid

Efficient exit can be encouraged by offering a deep connection charge contract on
first connection, reflecting the costs the new entrant would impose on the system over its
life, and recovered in a fixed number of annual payments (as in a mortgage), at the end of
which the owner of the right would be free to sell it to a comparable feed-in. That would
discourage premature exit from locations where the G charge is high but the value of the
connection low, an issue that becomes important in capacity auctions, where the aim is to
find the price needed to keep useful plant on the system as well as to reward new entry
when required.
The problem is greatly exacerbated because efficient nodal prices only recover about one-third of the cost of an expanding transmission system (Pérez-Arriaga et al., 1995) and considerably less for a legacy system where expansions are limited to a few lines. As a regulated entity, National Grid must recover its regulated revenue by an additional and large share that once the average G tariff is set, is imposed on consumers (or should be, but in fact is imposed on the net demand from larger consumers with embedded generation).

To the extent that RES can connect to the distribution network, it not only avoids TNUoS charges but will likely be credited with reducing load (Elexon, 2016), in which case it receives the same locational signal (the differentiation between zones remains roughly the same) but can benefit by the sum of the G+L tariff by reducing net load on the DNO. This is rising from £49/kWyr in 2016/17 to a predicted £66/kWyr in 2020. If these charges really represented the forward looking marginal cost of expanding transmission to accept new generation (G charges) or the saving from not having to invest as much in transmission for L, that would be efficient, but as just noted, only a modest part reflects the transmission needed or avoided, massively distorting the incentive for generation to locate at distribution rather than transmission level.

To restore efficiency in locational signals, ideally one would move to nodal charging (LMP) as in the US Standard Market Design, right down to some minimal G infeed (1 MW?). Nodal pricing is not obviously ruled out by the Target Electricity Model, but is strongly resisted by traders who prefer the liquidity of large price zone, and seems unlikely to be introduced in GB any time soon. The (considerable) shortfall in revenue would then be recovered from the gross (not net) consumption of consumers, ideally concentrated on hours of lowest demand elasticity (Ramsey pricing). Domestic consumers are levied such charges between 4pm and 7pm every day throughout the year, and these are plausibly hours of least elastic demand.

By 2008 it was becoming clear that the Renewables Directive (EC, 2009) had set a challenging target for UK RES that was unlikely to be delivered, while the threat of intervention to secure more RES was seriously undermining investments in the conventional generation needed to replace life-expired nuclear stations and the forced closure of unabated coal plant (facing tough environmental regulations, a new carbon price floor, and, in 2016, a political declaration that all unabated coal should close by 2025). The diagnosis and remedy are discussed in the next section.

---

8 Let the G charge on the grid in zone \( z \) be \( T_z \), then the L charge is \( K - T_z \), where \( K \) is the sum of G + L tariffs, roughly constant across zones. DN-connected generation is credited with \( K - T_z \), or pays \( T_z - K \), so the zonal differentiation in G charges remains the same but DN-connected G gains \( K \).
4. Electricity Market Reform: aims and outcomes
The volatility and unpredictability of future RES revenues under the Renewables Obligation (RO) Scheme made them hard to finance except through rather unattractive PPAs with the incumbent utilities, whose balance sheets were under severe stress. RES generation is highly capital-intensive so the larger part of their cost is the weighted average cost of capital (WACC), which is considerably raised by both the market risk of unpredictable future revenues and the political/regulatory risk that the RO scheme would be reformed or ended. The same policy uncertainty made conventional generation unbankable, in contrast to the 1990s’ dash for gas (although that was predicated on long-term contracts whose counterparties could now no longer pass on such costs after domestic retail was liberalized in 1998-9).

Meanwhile wide cross-party support had delivered the Climate Change Act 2008 (HC, 2008) that provides a legal framework for delivering greenhouse gas (GHG) commitments. In response, the Committee on Climate Change has set a 57% emissions reduction by 2030 (relative to 1990) in its Fifth Carbon Budget (CCC, 2015), much of which will have to be delivered by the ESI. The UK originally thought that a combination of renewables (mainly on and off-shore wind), new nuclear power, and, towards the end of this period, plant with Carbon Capture and Storage would deliver that, but none of these options was commercially viable in the liberalized UK electricity market. In response the Treasury enacted a Carbon Price Support that would bring the EU Emissions Allowance price for carbon dioxide up to an escalating price that was planned to reach £30/tonne by 2020 and £70/tonne by 2030 (HMT, 2011). This was hoped to make new nuclear commercially viable and enable a withdrawal of subsidies from mature RES such as on-shore wind and solar PV, and eventually, even off-shore wind. The credibility of the Carbon Price Support was severely undermined when it was frozen at £16/tonne shortly after in a subsequent budget.

The proposed solution to these various problems was to gradually phase out the RO Scheme and replace it with FiTs with Contracts for Difference (CfDs) that essentially guaranteed the sales price of RES in real terms for 15 years. To ensure adequate capacity for reliability (the primary political requirement within the trilemma objectives) the TSO would hold annual capacity auctions for delivery four years ahead for amounts specified by the minister (advised by his Department of Energy and Climate Change, DECC, and scrutinized for factual accuracy by the Panel of Technical Experts, PTE). In addition the Government has painfully negotiated and, as at April 2016 not yet concluded, a long term fixed real price contract for 30 years with EdF to build a new 3,200 MW nuclear power station at Hinkley Point (for an excellent history of UK nuclear power policy, see Taylor, 2016). After extensive consultation this Electricity Market Reform (EMR) was delivered in the Energy Act 2013 (HC, 2013).
4.1 The capacity auction

How well did it deliver on its objectives? The first objective of reliability was to be delivered by an annual capacity auction, with the amount to procure chosen by the minister in June 2014 on the basis of National Grid’s (2014) Electricity Capacity Report, and the PTE’s rather critical commentary (DECC, 2014). Newbery and Grubb (2015) set out these criticisms in greater detail, and also commented on the outcome of the auction. At the predicted net Cost of New Entry (CoNE, that is, the missing money created by price caps below the Value of Lost Load - see Newbery, 2015) of £49/kWyr the cost of procuring the 53 GW of derated capacity would have been £2.6 bn/yr, but the auction cleared in December 2014 at £19.40/kWyr, or 40% of this value, apparently demonstrating the power of auctions to reveal costs.

The net CoNE was based on a new CCGT, and about 2.6 GW of new entry was successful in securing capacity agreements. Surprisingly, given a clearing price 40% of the predicted net CoNE, a large (1.6 GW) new CCGT won, as did a large number of smaller units (790 MW of combustion turbines or reciprocating engines with an average capacity of 11 MW). By 2016 it was becoming increasingly doubtful that the CCGT would secure financing and might forfeit the modest penalty for non-delivery (€38 million on an €800 m project), and complaints were reasonably directed at the large number of diesel engines. The complaints mainly took the form that they were not the preferred CCGTs and were more carbon-intensive and polluting, neither of which is a good objection. Meeting peak residual demand (demand less RES) for the few hours a year of tightness would normally require cheap capital high variable cost plant which can reach full power rapidly, for which these units are ideal. The correct objection is that they secure an embedded generation benefit equal to the G + L charge which in 2018/19 will be on average £52.75/kWyr in GB (slightly less in England and Wales). With the auction price of £19.40 this gives them a credit of £72/kWyear. If such a unit costs £400/kW the payback is only 5 years even if there is no additional credit for running in Triad periods and the announced stress events for which the capacity was procured.

If transmission and distribution pricing had been efficient, the embedded benefit might have only been 10-30% of the current G + L wedge shown in Table 1, or perhaps

---

9 Capacity to be procured is derated to reflect its availability, and subsidized capacity (RES) is not allowed to receive the capacity payment but its derated value is subtracted from forecast demand. Old plant is given a 1-year contract renewable at annual auctions, new plant 15 years, and plant can choose not to be bid at T-4 (i.e. 4 years before delivery) but wait for the T-1 auction. National Grid has to predict how much plant will still be around to bid at T-1 and deducts their derated capacity from net demand to determine the amount to procure. The auction was run as a last price descending clock auction based on best US practice (see Newbery and Grubb, 2015).

10 The reliability standard is 3 hours Loss of Load Expectation (LoLE), where LoLE = $\sum \text{LoLP}$ over the hours of the year from equation (1), and Loss of Load Event is one in which market supply ahead of gate closure is less than demand, so that the SO has to intervene to maintain reliability, by calling on extra capacity, reducing voltage and other measures discussed in Newbery and Grubb (2015).
£5-15/kW, and the auction clearing price might have indeed been closer to the correct net CoNE and perhaps closer to £49/kWyr. The effect of the distortion is to secure a small amount of needed plant of probably the right type and size and logically connected to the distribution network to relieve constraints, and well below efficient market price, with a beneficial impact on consumer costs (a small increase to cover the lost TNUoS revenue and a large saving through a lower auction price paid to all 53 GW).

The other problem that was rapidly revealed was that with the fall in gas prices and the carbon price support, coal plant started losing money and their owners announced closures, including some that held a capacity agreement. Clearly the losses until this payment started in 2018 outweighed the future capacity revenue. As the capacity payment at £19.40/kWyr was below some of the higher TNUoS charges shown in table 1, the question whether those TNUoS charges were given correct exit signals suddenly became highly relevant, but at the date of writing, remains unresolved.

4.2 Supporting renewables
The new CfDs for RES were initially priced by bureaucrats at DECC, on the advice of consultants and after discussions with investors, using a high WACC (for on-shore wind DECC considered the WACC might fall from 8.3% under the RO scheme to 7.9% with a CfD, DECC 2013) that encouraged a huge flood of applicants, mostly wanting advance contracts at these prices before the full scheme was properly implemented. The National Audit Office argued that large amounts of money were poorly spent on the transitional (FIDeR) contracts paying these strike prices (NAO, 2014) as did the PTE in their first report (DEC, 2014a). Under pressure from DG COMP on state aid grounds, DECC decided to adopt auctions for allocating specified volumes of RES (mature, less mature – off-shore wind – and immature technologies – tidal stream, etc.). Newbery (2016) shows that the resulting clearing prices for on-shore wind lowered the WACC by 3% real. If the implied WACC is reduced by 3% through auctions then the saving on generation investment of £75 billion up to 2020 (DECC, 2011) would be £2.25 billion per year by 2020, continuing for 15 years.

This was a very promising rediscovery of the benefits of auctions rather than bureaucrats in price discovery, first trialled in the non-fossil fuel obligation auctions of the 1990s and then replaced by the costly and ineffective RO Scheme. Unfortunately, the final element of EMR was Treasury’s requirement for a Levy Control Framework that limits the support provided to renewables to sums rising to £7.5 bn/yr in 2020. At the time EMR was under consultation, gas prices were high and so were electricity prices, leaving the gap between the strike prices for the CfDs and the wholesale price modest. As gas prices fell, so the gap and support cost rose, and breached the Levy Control Framework. This is clearly counter-productive, for just as the cost to consumers falls and their ability to support the long-term investment for future decarbonisation increases, so that investment is put at risk.
Germany provides several useful examples of good practice here. First, during the period of liberalization, wholesale prices fell, but the Government gradually increased eco-taxes so that the retail price remained roughly constant and attracted little consumer resistance. Second, Germany provides nominal, not index-linked FiTs (more front-end loaded and more readily financed by nominal bonds) for a specified number of MWh per MW capacity, thus limiting the rents earned by those located in windy areas. This makes sense as increasing the price per MWh for the same contract length over-encourages wind farms to locate in windy areas that are typically far from demand centres (NW Scotland). TNUoS charges fail to properly reflect this, and GB is building extremely expensive off-shore DC links from Scotland down to England to handle the excess wind and to avoid building contentious additional pylons through the Cheviot border. More to the point, the learning benefits that are the argument for supporting RES derive mainly from the manufacture and construction of the plant, not so much in their operation (which is in any case rewarded for its energy revenue). A FiT paid for a fixed number of MWh per MW capacity is effectively a capital subsidy.

5. The All-Island example of adapting to high wind penetration

The Single Electricity Market (SEM) of the island of Ireland is an excellent subject for observing the impact of high wind penetration on a system in transition from a Pool-type centrally dispatched market design to one compatible with the EU Target Electricity Model (the Integrated SEM or I-SEM). It illustrates the challenges in procuring quality of supply through various existing and new flexibility services. The problems are amplified by the fact that the SEM is a moderately small, moderately isolated system with a range of fossil generation, with individual units that are large compared to demand. It has already high wind penetration, and currently experiences occasions with more than 50% system non-synchronous penetration (SNSP), at which point curtailment is required to maintain system stability. The ambition is to develop new ancillary and flexibility services under the DS3 project described in section 5.3 to cope with 75% SNSP, while at the same time transforming the centrally dispatched pool with a bidding code of practice that requires offers at short-run marginal cost augmented by capacity payments, into a model compatible with the EU Target Electricity Model, in which bids and offers are submitted to the European auction platform, EUPHEMIA, to determine market clearing prices and interconnector use, in the Day Ahead Market, the Intra-day market and the Balancing Market (which is effectively a Real Time Market).

The fact that the SEM has such ambitions for SNSP puts it at the forefront of managing intermittent renewable power, whose penetration is expected to increase in all EU markets in the coming years. If wind is not to be excessively curtailed, a suite of new services will need to be defined and procured, as discussed below. The current and
emerging systems are set out in documents published by SEMO,\textsuperscript{11} the market operator, Eirgrid’s annual capacity statements, and the All-Island Project.\textsuperscript{12}

The SEM has a considerable surplus of plant compared to generation with 9,774 MW of conventional plant in 2013 (SONI, 2014), about half of which was gas plant and 15% wind. Total existing wind farms connected to the grid and distribution networks, as of October 2014, was 2,646 MW and planned (contracted) is 5,071 MW, to give a subtotal of 7,717 MW. Average Cold Spell peak demand falls from just under 5,100 MW in 2007/8 (the start of the SEM) to just under 4,800 MW in 2013/14. The median Total Electricity Requirement for the peak is shown as growing from 4,850 MW in 2013/14 to just over 5,000 MW in 2019/20 and to 5,200 MW in 2024/5. The target renewables share for 2020 is 40%, which requires between 3,200 and 3,800 MW of wind capacity. This would require an average of about 240 MW of extra wind capacity installed per year. In the six years from 2008-13 2,000 MW was added, at a rate of 333 MW/yr, and over the past 10 years the increment has averaged nearly 400 MW/yr.

The larger generator units are about 400 MW, which is large compared to even peak demand and certainly compared with minimum demand. Thus in 2009/10 the maximum Market Scheduled Quantity (MSQ) was 6,565 MW and the minimum MSQ was 1,922 MW, a swing of 3.4:1. Under ideal conditions peak demands would include imports of up to 950 MW, so domestic generation need only meet about 6,000 MW in 2018, while the minimum demand can be supplemented by exports to make about 3,000 MEW, a swing of only 2:1.

In the period Nov 2001 to April 2010 the average absolute hourly change in MSQ was 213 MW (5% of MSQ) and the maximum was 1,165 MW or 25% of prevailing MSQ. Clearly these potentially large swings in demand compared to unit sizes presents considerable scheduling challenges, which are currently adequately managed in the centrally dispatched system given its forecasting ability.

\section*{5.1 Price formation in the SEM}

The SEM market design is similar to the old English Pool, except that under the bidding code of practice, generators are required to submit their true costs, which include start-up and no-load costs and their unit variable costs at each level of output. The capacity payment is a highly time-averaged and scaled version of equation (1), based on the net CoNE, akin to the GB capacity auction concept. (Note that the VoLL and LoLE are tightly inter-related in equilibrium with VoLL*LoLE = CoNE.)

\begin{itemize}
\item \textsuperscript{11} www.sem-o.com
\item \textsuperscript{12} http://www.allislandproject.org/en/homepage.aspx
\end{itemize}
5.2 Paying for System Services

In addition, the SO has to procure System Services, which are “those services, aside from energy, that are necessary for the secure operation of the power system.” (Eirgrid/SONI, 2011c).

Table 2 lists the existing services with their abbreviations on the left-hand side, and also a list of proposed new services that are considered necessary to manage future SNSP. The percentages show the extent to which current plant can provide the services needed by 2020, as discussed below. A fuller description of these services is given in the source SEM-13-060.

<table>
<thead>
<tr>
<th>New Services</th>
<th>Now</th>
<th>Existing Services</th>
<th>Now</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIR</td>
<td>65%</td>
<td>SRP</td>
<td>69%</td>
</tr>
<tr>
<td>FFR</td>
<td>54%</td>
<td>POR</td>
<td>87%</td>
</tr>
<tr>
<td>DRR</td>
<td>82%</td>
<td>POR</td>
<td>90%</td>
</tr>
<tr>
<td>RM1</td>
<td>88%</td>
<td>TOR1</td>
<td>91%</td>
</tr>
<tr>
<td>RM3</td>
<td>88%</td>
<td>TOR2</td>
<td>89%</td>
</tr>
<tr>
<td>RM8</td>
<td>66%</td>
<td>RRD</td>
<td>83%</td>
</tr>
<tr>
<td>FPFAPR</td>
<td>88%</td>
<td>RRS (Synchronised)</td>
<td>93%</td>
</tr>
</tbody>
</table>

Operating reserves (OR, Primary, Secondary and Tertiary) account for half the total annual cost of €61 million in 2013/14, with replacement reserves (RR) accounting for an additional 14% and reactive power for 18%.13

To interpret these services, fig. 2 shows the full range of existing and proposed frequency control services listed in Table 2. In addition to these Frequency Control services there are voltage response products shown in Fig. 3.

13 [http://www.eirgrid.com/operations/ancillaryservicesothersystemcharges/]
Figure 2 The existing and proposed Frequency Control Services
Source: SEM-13-060

Figure 3 Voltage Control Services
Source: SEM-13-060

Figure 4 gives a useful indication of the challenges of balancing a system with high and variable levels of wind production.
The dotted blue line should be read on the right hand axis and shows the wind penetration varying from 50% of MSQ (the red line on the left hand axis) in day 1 to 1% in day 2. The System Marginal Price, SMP (in green) peaks at 10 times its normal level as wind falls to its lowest level. These sharp price increases provide an incentive to invest in plant that can be called on at short notice and can ramp up rapidly to meet any shortfalls.

To gain a sense of the future challenge presented by the aspiration of a 40% renewables share, assumed to be a share of 37% of wind, the past levels of wind from 2007-2010 have been scaled up to give this level of penetration, assuming a constant 31% capacity factor. The maximum decrease in wind output over any half-hour period over this period is 826 MW (which is close to the maximum half-hourly increase required in dispatchable generation, including trade over interconnectors of 812 MW). The average absolute change in wind in any half hour is 125 MW. The average non-wind supply required over this simulated period is 3,000 MW and the minimum is -1,000 MW, which requires full exports in that period (and then probably some curtailment).

5.3 Delivering a Secure and Sustainable electricity System (DS3)
In response to the consultation (Eirgrid/SONI, 2011a), the DS3 Programme was set out in Eirgrid/SONI (2011b). The SEM Committee proposed seven new System Services to complement the existing seven illustrated in figs. 2 and 3 above and in more detail in SEM-13-060. There are a range of potential portfolio solutions which would allow the system to be operated at 75% SNSP, which may be delivered by refurbishing existing plant or may better be provided by new plant. There are short-run constraints on the most
practical way to procure existing and new system services in the radically different market structure still being developed for I-SEM, but the hope is that as far as possible the services will be procured through auctions, provided there are sufficiently many providers to provide enough competition (SEMC, 2014; 2015). As there is a strong complementarity between procuring system services and capacity the next section sets out the problem and possible solution.

5.4 Procuring flexible capacity in a high-RES system
With increasing RES, the capacity factors of conventional generation will decline, perhaps drastically, increasing the share of capital costs in annual total costs. This might not matter even in an energy-only market if investors had confidence that they could confidently predict their future revenue from energy and ancillary services and were not subject to price caps or policy interventions. Even in such benign conditions, the absence of adequate liquid futures markets required long-term contracts (PPAs) to encourage entry and investment in the “dash for gas”, and such long-term contracts are increasingly needed even for conventional generation as the electricity price detaches itself for longer periods from the marginal cost of fossil plant (Roques et al., 2005; 2006).

In a stable ESI, the value of various ancillary services would be reasonably predictable, as would be the future electricity prices and capacity factors. National Grid publishes annual *Future Energy Scenarios* which map out possible futures for varying fuel and carbon prices and degrees of decarbonisation, with their implied electricity prices and capacity factors by fuel type, which allow a forecast of energy revenues. In such cases it would be relatively simple to compute the missing money required to justify investment and then bid accordingly in the capacity auction. However, in the I-SEM, seven of the future system services are new with no record of likely value or price. More generally, as the ESI adopts more RES, smart meters, new platform providers to define, offer and aggregate services (Weiller and Pollitt, 2013) and as the markets for demand side response and distributed energy resources develop, the potential supply costs of each of these may change dramatically, making it hard to predict their future prices.

One possible solution is to run a package auction in which participants submit a range of offers for different plant or refurbishment options, specifying the volume of each service offered, including firm de-rated\(^\text{14}\) capacity for reliability, and the total annual required revenue from these services, in €/kWyr. New build would be eligible for a long-term contract, existing plant for a one-year contract, and major refurbishments for possibly an intermediate length. Each participant would be able submit any number of mutually exclusive offers with different bundles of attributes, and additional packages of mutually exclusive offers). The auctioneer might wish to publish demand schedules for

\(^{14}\) In the GB auction National Grid publishes the de-rating factors that will be applied to the full export capacity of the plant, reflecting the probability of forced outages or non-availability in stress periods.
some attributes (e.g. capacity) and impose price caps for others (e.g. when the cost of the service is less than the value of not curtailing RES). Given sufficiently many independent bidders offering a sufficient range of different bundles, the auction algorithm would search for the least-cost set of offers that meet the required demand and compute a set of prices for each attribute, ideally with the property that no bidder has a bundle rejected with higher value than any accepted, although this may not be possible, and least worst regrets approaches may be needed (dot.econ, 2015). These prices can then inform those considering new ways of providing the required system services. Once the required services have been efficiently priced, the missing money needed for capacity may be very small, given the security of long-term contracts for the system services.

5.5 The Energy Union’s approach to flexibility services
ACER (2015) stated that “We do not see a case for creating separate flexibility markets. In our view, fully implementing the Guideline on Capacity Allocation and Congestion Management and the Network Code on Electricity Balancing is necessary to reward flexibility within the market.” ACER does not say that this is sufficient, and later, in discussing long-term contracts, they appear to restrict these to contracts between generators and consumers, but note they are unlikely to be sufficient to “trigger investment decisions because their prices usually reflect expectations regarding future short-term market prices and because of their own inherent limits, e.g. durations limited to a few years for standard products and complexity for tailor-made contracts.”

ACER thus places great weight on short-term market signals to signal scarcity by time and location, but it is a leap of faith that these will be sufficient to deliver suitable investment signals for the kinds of plant that would be the best placed to provide energy, reliability and systems services in each location at least cost. Given the huge delays in delivering XBID (for cross-border trading intra-day),15 removing price controls, harmonizing gate closure times, and coordinating balancing across borders, this is an optimistic and largely unfounded assertion, unsupported by any available evidence.

The argument for long-term capacity contracts is to address the absence of missing futures markets and to reassure investors that future policy changes will not expropriate their investments, as argued for Britain above. The British philosophy is aligned with ACER (2015) in that all other markets, including those for system services, and all transmission charges, are either spot markets or very limited duration contracts. Thus for Transmission charges, the rates by location can change annually, although the methodology is moderately stable. Contracts with the SO for various system services are normally annual. The question is whether it would be desirable to have longer-term contracts for system services, and the argument above is that in periods of transition at

15 See e.g. https://www.epexspot.com/en/market-coupling/xbid_cross_border_intraday_market_project
least when future service prices are hard to predict, there is merit in de-risking them through auctioned contracts.

6. Lessons for the transition to low-carbon electricity

One of the concerns reflected in the Energy Union document (EC, 2015) is that RES is not subject to market signals that are becoming increasingly relevant for location and also their provision of services. Their conclusion is to argue for pFiTs, in contrast to the analysis above which demonstrated that FiTs (or CfDs) dramatically reduced the WACC and hence the cost of delivering RES targets. In a full nodal pricing system (or with a large number of price zones as in Nordpool), an over-concentration of RES in one location (e.g. the sunniest or windiest place) would depress the LMP (or zonal price) in periods of high RES supply, more so if transmission export capacities were limited, less with strong interconnection (of the type that Denmark enjoys). If this local price were passed through (with a premium) to RES suppliers, they would soon find it unattractive to continue locating in export-limited zones and would diversify into lower resourced but higher priced locations. Less critically, if some RES can provide system services, they should be encouraged to offer them, although this may more simply be provided through contracts than spot prices. It has also been argued that RES should bear the full extra balancing and system service costs they impose – an argument not limited to RES.

The practical question is how to achieve the risk-reducing benefits of contracts which transfer the risk to those for whom it can be borne at lower cost (end-consumers, or those with complementary generation portfolios) while providing signals for efficiently managing those risks – the classic principal-agent problem. The solution would seem to be suitable technology and location-specific contracts, such as the Transmission Congestion Contracts or Financial Transmission Rights much used in the U.S. Standard Market Design. These specify a strike price per MWh at the node, with the counter-party paying or receiving the difference with the LMP, and can be provided as long-term contracts in return for the connection agreement and payment. Balancing contracts can similarly be provided, and in effect already are under the GB RO scheme through PPAs with incumbent utilities. The TSO might be a better-placed (and regulated) counterparty, but this may run up against unbundling requirements. Independent System Operators might escape that restriction, but would need access to other counterparty funds, perhaps recovered as at present in GB through Balancing Service Use of System charges that could be passed through to final consumers.

The main lessons to draw from the theory and examples presented above is that the transition to the low-carbon ESI will require likely considerable increases in fixed and considerable reductions in average variable costs, making prices more volatile and less predictable. Proper scarcity pricing over time and space, and efficient remuneration of system services similarly introduce new uncertainty into revenue streams, and signal the need for different bundles of generation and demand side attributes. Markets with risk
benefit from hedging contracts, which can (and have) dramatically reduce(d) financing costs. Transmission tariffs, which are subject to regulatory scrutiny, need considerable reform if markets are to remain liberalized and unbundled. The alternative (which may be better suited to some markets) would involve a single buyer (essentially the TSO) signing long-term PPAs which would have a capacity and energy element (and possibly payments for each system service). The TSO would then select the least-cost dispatch on the basis of variable (i.e. energy) costs. Consumers could similarly hedge through CfDs and/or with Reliability Options (Vazquez et al., 2002) as under consideration for I-SEM.
References
CCC, 2015. Fifth Carbon Budget. Committee on Climate Change, at 
dot.econ, 2015. DS3 system services auction design report, at 
Eirgrid/SONI (2010). Harmonised Ancillary Services and Other System Charges; Market Participant Perspective
Eirgrid/SONI (2011a). All Island TSO Facilitation of Renewables Studies, July
Eirgrid/SONI (2014). DS3 System Services: Portfolio Capability Analysis
SEMC (2014) DS3 System Services Procurement Design; Clarifications to SEM-14-059 Information Paper, SEM-14-075, 13 August
SONI, 2014. All-Island Generation Capacity Statement 2014-2023, at
http://www.eirgridgroup.com/site-
files/library/EirGrid/Generation%20Capacity%20Statement%202014.pdf
Taylor, S., 2016. The fall and rise of nuclear power in Britain, UIT. See
https://www.uit.co.uk/the-fall-and-rise-of-nuclear-power-in-britain
Weiller, C. M. and M. G. Pollitt, 2013. Platform markets and energy services, EPRG WP 1223,
**Acronyms**

CCGT Combined Cycle Gas Turbine

CiD Contract for Difference

CoNE Cost of New Entry

CP Capacity Payment

DN Distribution Network

DNO DN Operator

E&W England and Wales

ESI Electricity Supply Industry

ETS Emissions Trading System (for CO₂)

FiT Feed-in tariff: a fixed price per MWh of metered output

G Generation

I-SEM Integrated SEM

L Load

LMP Locational Marginal Price

LoLP Loss of Load Probability

MSQ Market Scheduled Quantity

MW Megawatt

MVA megavolt amps, and takes into account of both the resistive and reactive load.

pFiT Premium FiT

PPA power purchase agreement

QoS Quality of Supply

REC Regional Electricity Company

RES renewable electricity supply

RO Renewable Obligation

ROC RO Certificate

SEM Single Electricity Market of the island of Ireland

SMP System Marginal Price

SNSP system non-synchronous penetration – e.g. wind

SO System Operator

TEC Transmission Entry Capacity

TNUoS transmission network use of system

TSO Transmission System Operator

VoLL Value of Lost Load

WACC weighted average cost of capital