Department for Energy Security & Net Zero

Electricity market design – evidence from international markets

Final Report

DESNZ research paper number: 2023/054

June 2023

Any enquiries regarding this publication should be sent to us at:
remamailbox@energysecurity.gov.uk

Contents

Acknowledgements

The Department for Energy Security and Net Zero (DESNZ) commissioned Ove Arup and Partners Limited (Arup) to undertake a review of international experiences with a range of electricity market designs.

This report was written by Ove Arup and Partners Limited. The views expressed do not necessarily represent those of DESNZ or any other government department.

Executive summary

The Review of Electricity Markets Arrangements (REMA) set out a compelling case for change in GB's wholesale electricity market. Concerns around increased balancing costs, capacity adequacy, operability and the sheer scale of investment required have been highlighted. REMA is considering policy reforms in five key themes: wholesale markets, capacity adequacy, mass low carbon power, flexibility and operability.

Other markets have tried, or use today, differing wholesale market designs, approaches to capacity adequacy policy, and alternative approaches to managing operability challenges. Whilst it is important to note inherent differences, the experiences of these international markets offer a useful set of experiences for any potential GB reforms. This report aims to assess the experience from relevant international markets and draw inferences for GB, using the following methodology:

- **High-level data collation and assessment** – case study markets were selected based on their similarity to the GB -as indicated by peak demand, generation mix, and the degree of intermittency.
- **Rapid Evidence Assessment (REA)** – the available literature for each case study market, including cross study literature was assessed. In several cases, there was a lack of evidence from the literature. In these cases, this was supplemented by analysis with discussions with market experts and examined data supplied by system operators.
- **Synthesis** – high level inferences were drawn on what could be applied to the GB market.

Based on the assessment detailed above, the following markets were selected as case studies:

Option	Selected case study markets
Nodal pricing	CAISO (California), ERCOT (Texas), PJM (northeast US states)
Zonal pricing	Nordic markets (Sweden, Norway, and Denmark), Italy, and Australia (NEM)
Centralised reliability option	Ireland (I-SEM) and Italy

Strategic reserve	Australia (NEM), Germany, Sweden[1]
Supplier obligations	US markets (treated as a group), Poland and Italy
Co-optimisation	US markets (CAISO, ERCOT, MISO, ISO-NE)

Table 1: Options in scope of the study and selected case study markets.

The main body of this report considers the evidence base against the following dimensions:

- **Policy rationale and intended outcomes**
- **Implementation** – pathway to implementation, timeline/costs
- **Whole system costs** – including impact on prices, balancing costs, and the impact on consumers.
- **Investment in low-carbon technologies** – including deployment of low carbon technologies and impact on the cost of capital.
- **Security of supply** – effects on capacity adequacy or system reliability.
- **Market liquidity** – does the design enable sufficient trading and allow for the forward buying and selling of power?
- **Consumer impacts** – both in terms of cost and participation/responsiveness to price signals (where applicable).

Wholesale market design – nodal markets

Policy rationale and implementation

The rationale for introducing nodal pricing in the case study markets was principally driven by concerns around constraint and re-dispatch costs. These are similar to concerns set out in the REMA case for change, echoed by the by National Grid Electricity System Operator (NGESO), around increasing constraint and re-dispatch costs.

NGESO has estimated five years to implement a nodal market in GB. This would match the quickest time for the case study markets, considering the need to develop a centralised dispatch model. Nodal pricing was typically introduced as part of wider

[1] A brief analysis of Poland and Belgium was also carried out on the grounds that these markets moved away from a market design which is under consideration in the REMA programme. Arup explored the reasons behind the change of market design.

packages of reform in US case study markets, including greater temporal granularity and optimisation with ancillary services. The evidence on the costs of implementation of a nodal market design are difficult to transfer to GB and appear also in absolute terms compared to system change costs estimates in GB. However, there was credible evidence from the literature that the case study markets were easily able to recoup their costs in benefits.

Impacts in international markets

In general, the evidence assessed suggests that nodal pricing has delivered cost savings in the case study markets, as well as wider market benefits.

The evidence assessed suggests that US consumers have seen cost savings after the introduction of nodal pricing due to reduced constraint costs, including in ERCOT and CAISO (which have similar generation mixes to GB). This supports the theory that nodal markets transfer rents from generators to consumers.

The degree to which consumers are exposed to nodal prices varies in the US markets explored. In most of the cases studied, only generators face nodal prices, whereas suppliers face zonal or average prices. CAISO, PJM and ISO-NE have a consumer base largely removed from direct exposure to nodal prices whereas ERCOT stands out as the market which exposes its consumers the most to nodal prices. Less exposure to nodal prices reduces the potential for effective demand side response. The risk of consumer price spikes can also be reduced by the use of Financial Transmission Rights (FTRs) and mandatory hedging obligations.

The evidence reviewed did not suggest that the implementation periods for nodal markets' transition dramatically hindered investment. The evidence from CAISO suggests that investment continued despite a prolonged implementation period. However, the investment challenge delivering of net zero in GB is of a significantly greater scale and pace, significantly reducing the strength of any inferences. The evidence from case study markets also did not suggest that nodal markets per se hinder the development of renewables and low carbon generation. ERCOT and CAISO have seen significant growth in renewables and low carbon generation. However, PJM (and other US nodal markets) tend to be dominated by thermal generation not compatible with net zero. The evidence suggests that the generation mix is largely driven by the efficacy of policy support mechanisms, rather than the market design.

The evidence suggests the US markets have reasonably good levels of liquidity, with PJM often cited as one of the most liquid electricity markets in the world. The use of FTRs and policies mandating hedging appear to have supported liquidity.

There is no evidence to suggest that nodal markets have adversely affected security of supply. However, capacity adequacy has generally been driven by the capacity markets in the US states. It is more likely that whatever the market design a capacity adequacy policy is required.

Implications and lessons for GB

The evidence from international experience suggests a nodal market design is likely to address some of the concerns around constraint payments in GB, and result in transfer from generators to consumers. The impact on investment is unclear, but likely to be driven by the efficacy of policy support mechanisms than wholesale market design per se. The design can be complex, and a centrally dispatched system is required. Therefore, it is likely to take a minimum of 5 years to implement and have significant implications for GB market participants.

Given the scale of forecast constraint costs in GB, it is likely that a move to nodal pricing would enable a similar transfer of benefits from generators to consumers (i.e., rent transfer) as observed in US markets. Based on UK constraint costs forecasts, and the experience of the US, the recent estimates of consumer savings from nodal market implementation of £30 billion by 2030[2] and £69bn between 2025-2040[3] appear credible. The evidence also suggests efficiency improvements from moving to a nodal market design, however this is inherently difficult to quantify and difficult to disentangle from other market design choices.

Wholesale market design – zonal markets

Policy rationale for intervention

Many markets implemented a zonal approach largely due to a lack of transmission capacity causing constraints and increasing balancing costs. Several markets, such as CAISO and ERCOT moved to a zonal design in an effort to reduce constraint costs before settling on a nodal market design.

Impacts in international markets

There is no evidence to suggest that zonal markets harm investment in low carbon generation compared to a uniform price market or indeed a nodal market. Norway, Sweden, Italy, and Denmark have all had significant growth in renewables. However, there is evidence that this growth in renewable generation increases constraint costs in a zonal market when there is insufficient transmission capacity (as noted by the Australian Energy Market Commission (AEMC). This is partly owing to large wind and solar farms locating away from the large demand centres.

Zonal pricing also requires accurate zonal boundaries, which often need to be redrawn. The need to change zone boundaries can create uncertainty for investors.

[2] Energy Systems Catapult and Octopus (2021).
[3] FTI (2022).

The experience of zonal markets in Australia and in US markets suggests a zonal market design may not alleviate concerns with constraint payments. However, European countries have experienced lesser concerns with constraint costs, perhaps because of greater transmission capacity. Modelling work by Aurora (2020) looking at a simple three-region model in GB suggests that zonal markets can result in reduced system costs between 2025 and 2050 of up to £50bn. However, given the experience of other countries, benefits of zonal markets need a thorough interrogation.

The evidence also suggests that, without adequate mitigations and regulatory oversight, zonal pricing can increase market power for some firms to exploit their power behind a constraint (“increase-decrease” gaming). The potential for this type of behaviour already exists in GB.

Implications and lessons for GB

Modelling has suggested significant system savings if GB implemented a zonal market. In addition, a zonal market is likely to be far simpler to implement than a nodal market. However, the experience from other markets suggests that zonal markets may not resolve all the issues with constraint costs. The need to continually examine and redefine boundaries creates further complications. The risk of market power being exploited in a zonal market and the impact on transmission charging also needs further examination.

Capacity Adequacy

Security of supply is a key objective of energy policy, and market mechanisms for its delivery vary significantly internationally. GB, alongside the US, uses a market-wide capacity mechanism, but European markets use a range of alternatives. REMA is considering the extent to which these alternatives could deliver more effectively; this study focuses on Centralised Reliability Options (CROs) and strategic reserves, and their implementations internationally.

Centralised Reliability Options

In a CRO market design, the system operator purchases reliability options (call options) from the capacity providers via an auction. A reliability option is a financial contract that entitles the option holder, the system operator, to receive difference payments from the sellers, the generators, if the price in the electricity market exceeds a predefined strike price (auction clearing price). The system operator pays the sellers of the option, a price that can then be interpreted as a capacity payment.

Theoretically, CROs promise to limit market distortions and to act as a hedge for consumers against higher bills in addition to maximising the incentive for contract capacity to generate during scarcity events. There are a limited number of markets that have used CROs, this report reviews the experience of Ireland and Italy in implementing a CRO. It should be noted that the Irish market is significantly smaller than that of GB and drawing inferences applicable to GB is challenging.

Irish (I-SEM) CRO

The Irish CRO was established in 2017 and, consequently, there is not a large amount of academic literature on the efficacy of the approach. However, the available literature does suggest:

- The CRO has secured a larger amount of capacity at lower cost than the previous schemes.
- In only the second year of the scheme, the auction outcome had to be superseded because of concerns of lack of power in Northern Ireland leading to Kilroot coal fired power station being awarded a direct contract after being unsuccessful in the auction.
- The CRO in Ireland has not delivered new low carbon generation technologies,
- The CRO has favoured fossil fuel power generation, in particular gas-fired generation.

Italian CRO

The Italian CRO, implemented in 2019 was motivated by concerns that decarbonisation efforts can bring a security of supply risk. This is because with reduced running hours there is a risk prices won't rise high enough for dispatchable power generators to recoup costs when they do run.

We conclude it is too early to draw firm conclusions on Italy's CRO as, given its recent implementation, the literature evaluating its impacts is very limited.

Applicability to GB

Given the limited evidence base discussed above, extracting relevant inferences for GB is challenging. However, the Irish experience suggests the policy is not well suited to bringing forward new low carbon technologies, and the need to override the auction outcome to ensure there is generation located in the right areas suggests it is not well suited to markets with significant transmission constraints.

A longer time frame is needed to properly assess the efficiency of CROs in Ireland and Italy.

Strategic reserve

A strategic reserve works by having a set amount of capacity kept outside the market for use when there may not be enough supply to meet demand. This can have the potential benefit of only paying a few plants to remain open, potentially reducing a market's capacity adequacy costs. However, commentators are wary of introducing such an approach in GB due to the potential for an ever-increasing strategic reserve, as plant exit the market in the hope of landing a strategic reserve contract.

Australian (NEM) strategic reserve

The Australian market introduced a strategic reserve in the late 1990s. It was initially intended as a transitional mechanism, planned to be phased out in 2000. However, it was extended four times and officially made permanent in 2016.

The cost in 2017/18 was approximately $51m Australian dollars and appears relatively cheap. This is significantly less than the CRO reviewed above, or our own CM, whose costs are much greater. There is little evidence a slippery slope has been created. However, some have argued it has distorted market prices and is dominated by fossil fuel generation.

German strategic reserve

Germany announced their strategic reserve in July 2016, and it came into operation in October 2019 with 2GW of capacity. Germany identified and quantified the security of supply risks and deemed that the reserve was needed to ensure security of supply during the ongoing reform of the German electricity market and to manage the phase-out of nuclear electricity generation. The reserve is intended to be temporary. This is planned for 30th September 2025.

To limit the risk of a ‘slippery slope’, the German mechanism prohibits plant from re-entering the market once they enter the reserve. This has proved effective in limiting this concern. However, some argue this has inhibited investments and the Research Institute of Industrial Economics suggests this has negatively affected the mechanism from a resource adequacy perspective.

Applicability to GB

The evidence presents a mixed picture on strategic reserves. In Europe, there is a growing tendency to move away from the mechanism in favour of capacity markets. Some policy appraisals suggest a significant cost benefit in this. However, in the countries where they have been used, strategic reserves have been a cheaper year-on-year option than capacity markets. In addition, there does not appear to be any evidence of the ‘slippery slope’ that has previously concerned GB policy makers. The academic literature suggests that strategic reserves may not be sufficient in securing the required amount of resources, especially in a system with a high penetration of renewables (Billimoria et al., 2018).

A strategic reserve would likely be straightforward to implement in GB, given the ESO’s recent Supplemental Balancing Reserve, and evidence from case study markets suggests they could be a cost-effective measure to keep existing plant online as the system decarbonises. However, it is not clear whether this would meet all of GB’s policy objectives. The evidence from other markets suggest that strategic reserves do not provide investment incentives, especially for low carbon generation. However, they may offer the ability to keep existing plants operational to cover year-ahead capacity risks. Further work is required to fully assess the suitability for GB, but strategic reserve could offer targeted support for existing plants in times of heightened risk.

Low carbon flexibility and operability - supplier obligations

Supplier obligations have been widely used in energy markets across the world. To make the case study approach manageable the case studies have been grouped into US and European markets.

Supplier obligations in US markets (CAISO, ERCOT, MISO, ISO-NE)

In the US, the supplier obligation has been introduced through the Renewable Portfolio Standards (RPS). This sets a minimum requirement for the share of electricity supply that comes from designated renewable energy resources. The policies are focused on electricity generators and vary state to state.

The RPS in the US market has resulted in significant increases in renewable capacity in the systems. ERCOT has observed significant increases in wind capacity and similar trends can be observed for solar generation within the CAISO market. Additional policies have helped renewables connect to the grid. There are also mechanisms to support the balancing needs of such intermittent generation sources. Further, both ERCOT and CAISO enjoy many hours of sunshine, and in the case of ERCOT consistent wind speeds suggesting less intermittence relative to GB.

There is little evidence on the cost effectiveness of the supplier obligations. However, research suggests RPS has, so far, delivered an increase in renewables penetration at a low percentage of consumer bills. However, the compliance cost of the scheme is increasing, and it is likely that this will increase further reflecting a similar experience with GB's renewables obligation.

Supplier obligations in European markets (Italy and Poland)

Italy introduced its supplier obligation in 2002 under the Green Certificates mechanism. Poland implemented its Quota Obligation and Green Certificate scheme in 2008. Both countries observed significant increases in renewable generation capacity and investments into low carbon technologies. However, there is little evidence in the literature which examines the cost effectiveness of the policy. It is highly likely that part of the success of obligations stems from high subsidies for renewables.

Applicability to GB

The experience of US and European markets suggests that supplier obligations have been an effective policy in bringing forward the deployment of renewables technologies. However, there is a lack of evidence as to the cost of the obligations. A lingering concern is that they overcompensate investors and may not be suitable for the mature renewables market we have in GB today.

Low carbon flexibility and operability – Co-optimisation of ancillary services

Centralised dispatch enables the system operator to co-optimise energy and ancillary services; both services can be scheduled within the same process. This is not possible in markets with self-dispatch, where energy and ancillary services are procured separately, and can potentially result in imperfect optimisation of resources. Markets that operate a central dispatch regime typically co-optimise ancillary services with energy. There is evidence that co-optimisation can reduce ancillary services costs, as well as supplying consistent price signals to providers of both energy and ancillary services.

The academic and policy literature is overwhelmingly supportive of the co-optimisation of ancillary services and energy. However, owing to the inherent difficulties in quantifying the benefits, there is little empirical research to back this up. Nonetheless, there is strong consensus that co-optimised electricity markets lead to efficiency gains.

Applicability to GB

NGESO highlight the numerous difficulties in co-optimising energy and ancillary services in GB. The more elements that are required to be co-optimised, the harder it is to solve the optimisation problem, significantly increasing the difficulty of building the algorithm and optimiser. Nonetheless the ESO believe that, in a decarbonised system, centralised dispatch with co-optimisation can offer efficiency benefits.

The evidence from other markets suggests that the co-optimisation of ancillary services works well in nodal markets with centralised dispatch. Arup sees no reason why this would not also work well in GB with a nodal market and centralised dispatch. However, it is likely that co-optimisation would need to be a greatly simplified set of services, based on short term operating reserves.

Introduction

Research context

The existing GB market design was established in 2001 (NETA) when the system relied predominantly on fossil fuelled generators. These were broadly located in areas where the grid did not constrain their operation. Weather dependent renewable energy generation accounted for less than 1% of the electricity mix.

As climate change concerns became a key priority, the UK government responded with the introduction of the Electricity Market Reform (EMR). As part of the EMR, the Capacity Market and the Contracts for Difference (CfD) schemes were introduced. These targeted security of supply and strong investment in renewable energy assets (>£18bn in the last 4 years). However, the EMR did not change the core of the electricity market design which still relies on NETA.

Under the current trading arrangement generators can make their own dispatch decisions. They receive a single national price, irrespective of location, which is revealed via the day-ahead auctions and over the counter trading. This design assumes that generators can serve load anywhere in the country, in effect assuming away transmission constraints and transmission losses. To ensure system frequency, the electricity system operator (ESO) is charged as acting as a ‘residual balancer’.

The existing market design has delivered a significant increase in renewables penetration, mostly through the uniform price, the ‘Connect and Manage’, and subsidy schemes. Renewable energy accounted for nearly 40% of total generation in 2021 with intermittent generation accounting for 26% (DUKES, 2022). This dramatic growth has had significant impacts on system costs and is placing the GB design under strain. The System operator is required to take increasing amounts of actions to manage system constraints at significantly higher costs. Moreover, the loss of conventional fossil fuel generation also led to a loss of traditional system services, such as Black start and System Inertia. This is leading to increased costs for ancillary services. Constraint costs have seen an 8-fold increase, rising from £170m in January 2010 to £1.3bn in January 2022 with further rises projected (ESO, 2022).

On top of this, the country remains committed to challenging Net Zero Targets. Such targets are set out in the recently released British Energy Security Strategy (April 2022) setting a target of delivering up to 50GW of offshore wind (12.7GW in Q1-22) and up to 10GW of Hydrogen by 2030. These targets need to be delivered without compromising security of supply or consumer affordability. To ensure the GB electricity system is compatible with the ever-changing electricity market whilst delivering on its ambitious Net

Zero target in an affordable and secure way, the UK government has announced a Review of Electricity Market Arrangements (REMA).

REMA themes in scope

REMA is a wide-ranging review of the GB Electricity Market Arrangements. The present report focuses on five market areas (known as “REMA Themes”) and furthermore refines its focus on specific design options within these market areas (“REMA Options”): wholesale market (nodal and zonal pricing), capacity adequacy mechanism (centralised reliability option and strategic reserve), low carbon investment (supplier obligation on decarbonisation), flexibility (Clean Peak Standard CPS) and operability (co-optimisation of ancillary services with the wholesale market).

The section below provides a summary of each REMA Themes and its corresponding REMA options in scope of the study.

Wholesale market designs

The current GB market, there is little locational incentive for generators, albeit transmission charges attempt to send ex-post (i.e., after dispatch) locational signals. Certain market designs, that operate in other countries, can address this challenge.

Nodal pricing

This is largely motivated by the aim to increase efficiency of dispatch signals for generators. Efficient dispatch means that a generator considers the whole system costs of generation actions, as opposed to just its own production costs when its price is calculated. In a nodal market design, the transmission system is divided into many smaller, "nodal" locations where electricity is produced or consumed. Each nodal location has its own electricity price, and those prices are set based on the supply and demand at that location. The outcome of a nodal market is that there is no longer uniform electricity price set by the marginal source for the whole of system. Instead, there are many different nodal prices reflecting the marginal source of supply at that nodal.

Zonal pricing

Zonal pricing is an alternative to nodal pricing, which is used in the European energy market. Zonal pricing involves dividing the network into specific zones, which reflect major transmission constraints between different countries or within a country. Each zone has a single price assuming no network constraints within that zone. If zonal pricing is used on both the supply and demand side, suppliers pay for energy at the same price they receive for selling energy within a single zone. If the energy price differs between two zones, suppliers pay the difference between the zone where the energy was generated and the zone where it is supplied, reflecting the cost of network congestion between the two zones. System operators determine and communicate the available capacity values between zones.

Capacity adequacy mechanisms

Internationally there are a variety of designs that have been implemented to ensure sufficient, cost-effective flexible generation is commercially viable to meet security of supply standards. These designs include capacity markets, capacity payments, centralised/decentralised reliability options, targeted tenders, strategic reserves, and equivalent firm power auctions. In this study, we explore further two capacity adequacy market design options, Centralised Reliability Options (CRO) and Strategic Reserves (SR).

As part of the REMA consultation, DESNZ conducted an option assessment where it considered a range of capacity mechanisms and other options including those that have been proposed by stakeholders (including in the recent Capacity Market Call for Evidence) and those that have been deployed internationally (as illustrated by the figure below). This includes a wide range of options: from incremental changes to the existing Capacity Market to more radical options which propose to replace it completely.

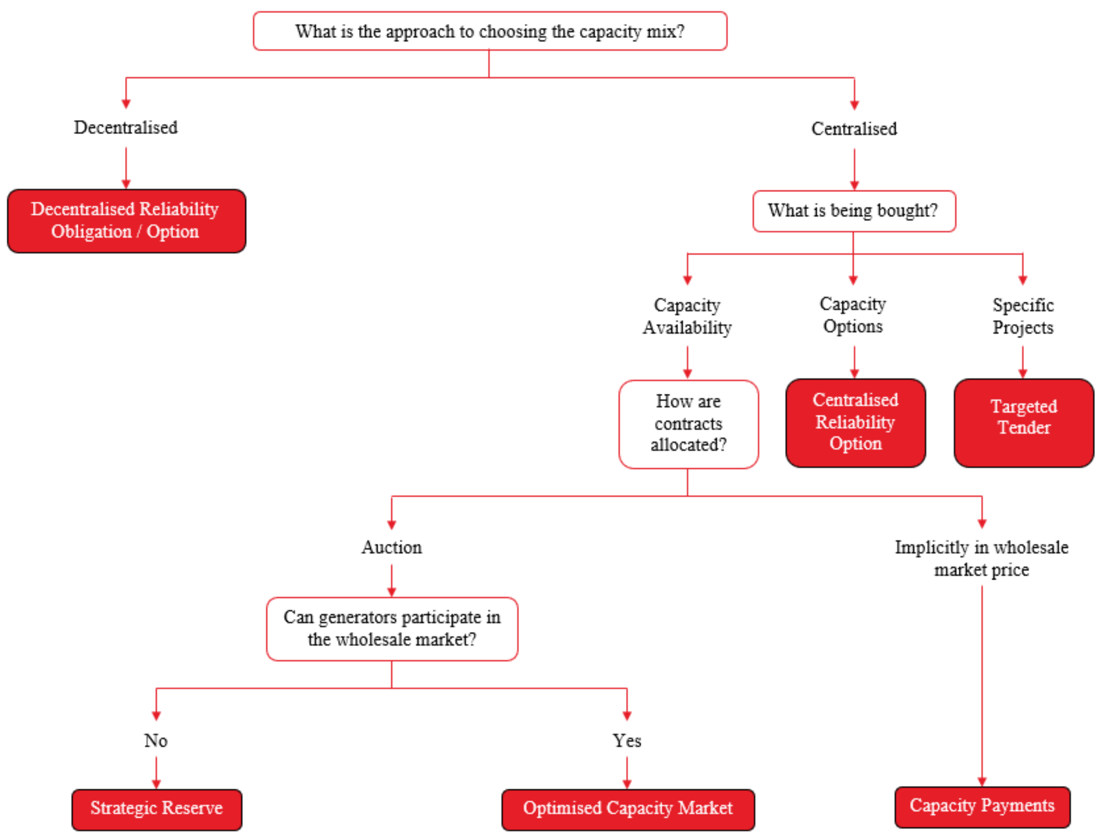

Figure 1: Capacity adequacy – policy decision tree Source: REMA

Out of the options that DESNZ have considered in the REMA consultation, there are three market designs which DESNZ believes could have potential advantages over current arrangements, and which DESNZ proposed taking forward for further exploration: an Optimised Capacity Market, a Strategic Reserve (as an emergency intervention only) and a Centralised Reliability Options design. In this report, Arup is focusing on the last two market designs: Centralised Reliability Option and Strategic Reserve.

Centralised Reliability Option (CRO)

Electricity markets in Ireland, Italy and Colombia run a version of this design. In a CRO design the ESO runs auctions and sets the amount of capacity required to meet peak demand. Successful generators get paid a premium to be available during times of scarcity in which wholesale prices exceed a strike price. In exchange the ESO has the right to buy electricity from the contracted generators at the strike price. Contract holders that are not available during system stress events get penalised thus creating an incentive to be available at times of scarcity. The reliability option mechanism provides an incentive to maximise the availability of power units by generation units during peak demand. The reliability options transfer the responsibility for ensuring capacity adequacy in the power system from the regulator (or TSO) to generation companies. Variations of this option, including a price cap, could address the rising costs currently seen in the GB Capacity Market.

Strategic Reserve

The REMA consultation sets out proposals for a strategic reserve model. Countries which have implemented such a model include Sweden, Germany, and Australia, whilst Poland and Belgium have replaced strategic reserves with new capacity market mechanisms. The model involves auctioning a set amount of reserve capacity above what the market is expected to supply. Providers are remunerated for availability whilst they also receive an activation payment. Capacity in strategic reserves is dispatched only when there is a risk of negative margins (i.e., demand is above supply). This can be more cost effective than a CM as it only procures capacity for peak times. However, it does not incentivise investment in low carbon generation. This mechanism is also flexible – if sufficient incentives come from the market, the volume of the strategic reserve is maintained at a minimum level. Additionally, power generation companies have no difficulties in estimating the expected revenues as they are set in the bilateral agreements.

GB has considered strategic reserves before as part of the policy development process that led to the CM. A major concern with a strategic reserve has been that it creates a ‘slippery slope’, where an increasing number of existing plants exit the market in the hope of gaining a reserve contract resulting in an ever-larger need for reserve and significant cost. A further concern is the distortionary effect of strategic reserves on market prices. They can have a dampening effect on scarcity prices which can reduce incentives to invest. GB also has experience in deploying a strategic reserve: the supplemental balancing reserve (SBR) developed by the ESO for 2014-2017 procured around 1.8GW of power kept outside of the market to be used only when required. This provides a useful learning experience. It could also be argued that in winter 22/23 GB had both a CM and a strategic reserve. This is because prior to the winter, government negotiated with two coal plants to remain open when they had planned to shut. An undisclosed fee was paid to these plants, but this is expected to be costly with the running costs of old power station at roughly £20 million annually. However, given the significant changes in generation mix and the enhanced net zero targets a fresh look is required. It is important to see whether the previous concerns around a ‘slippery slope’ and price distortion still hold.

Low carbon investment and flexibility

Implementing a supplier obligation has been identified in the REMA Consultation as one of several options for improving investment signals for flexible generation (e.g. Clean Peak Standard in Massachusetts) and fostering mass low carbon power. Supplier obligation market designs have international precedent, notably in the US. The GB experience with the renewable obligations highlighted the difficulty in using obligations as a bankable investment signal for large capital-intensive investments. However, the experience in Massachusetts suggests they could be an effective tool for incentivising flexibility responses that do not require large investments but do need a clearer incentive to change dispatch or consumption decisions. There is merit in considering obligations specifically targeted at responses that do not require large capital investment. We use the US, Poland, Italy as case studies to investigate how the implementation of a supplier obligation affects provision of operability services and how well the expected peak demand and flexibility requirement aligned with actual peak demand and actual flexibility requirement.

Operability

The REMA Consultation covers options proposed for delivering operability and ancillary services to NGESO. It outlines changes which could be made, such as giving the ESO the ability to prioritise zero or low carbon procurement or giving carbon reductions equal weighting to cost effectiveness in procurement. Co-optimisation of ancillary services with the wholesale market is one of the options under consideration in the REMA Consultation, with key questions being whether a centralised and co-optimised design has been more efficient than the current de-centralised approach and whether this would be effective in the GB market.

Research objectives and research questions

This report aims to provide an overview and synthesis of existing literature and evaluates evidence relating to REMA options where international precedent exists. It aims to provide insights on lessons learned from international experience. Further to this, the aim of this research project is to assess whether any of the successful market design elements of other markets could also be effective in GB. Table 2 presents the REMA themes and REMA options under scrutiny in this report.

The overarching research questions, which underpin the study’s overall aims and objectives, are set out in the Table below.

Research Theme	Research Questions
Market design choices - rationale	What are the market arrangements and conditions (e.g., capacity mix) for the identified case? How do these arrangements compare to current GB arrangements and the expected characteristics of the GB electricity system in the future? What was the rationale for intervention and the intended outcomes of the policy in question?
Market design impacts and efficacy	What does the existing evidence tell us about the impacts of the policy in question? This should include, where available, impacts on: • Decarbonisation and deployment of low-carbon technologies • Security of supply • Capacity mix (including firm and flexible, low vs high-carbon assets) • Whole system costs • Cost of capital and investment • Market liquidity • Consumers (both in terms of cost and participation/responsiveness to price signals, where applicable)
Implementation	What was the pathway to implementation? Was the policy delivered as intended (including to original timelines/costs)? How has the policy evolved since its introduction? What lessons from implementation could be applied to GB?

Table 1: Overarching research questions classified by research theme.

In addition to the overarching questions set out above, this report explores the following option-specific questions:

Theme	REMA Option	Option-specific questions

Wholesale market	Nodal pricing (including nodal pricing models with full and limited/no demand side exposure)	How does the network infrastructure and transmission access regime compare with GB? Which technologies are more and less likely to set the price at individual nodes? Did gas and electricity prices decouple to any extent? What was the pathway to implementation – including timelines, motivation for reform, assessment of enabling factors, grandfathering arrangements. How was the technical challenge of moving to algorithm-based dispatch overcome? How do suppliers manage nodal price risk? How are costs passed on to consumers, and where costs were not passed on, what impact did this have on the overarching impacts (system costs, technology deployment, etc.)?
Wholesale market	Zonal Pricing	As above.
Capacity Adequacy	Centralised Reliability option	What are the delivery penalties under the mechanism?
Capacity Adequacy	Strategic Reserve	What are the delivery penalties under the mechanism? For countries that have moved away from a Strategic Reserve model, what were the drivers for the change?
Low carbon investment and flexibility	Supplier obligation on decarbonisation and/or on flexibility (incl. Clean Peak Standard)	How did the implementation of a supplier obligation affect provision of operability services? How well did the expected peak demand and flexibility requirement align with actual peak demand and actual flexibility requirement?
Operability	Co-optimisation with	How substantial have the efficiencies been compared to non-co-optimisation?

	wholesale market	

Table 2: Option specific research questions

Methodology

The study consisted of three key steps, outlined in the Figure below, and used a Rapid Evidence Assessment (REA) approach as its basis. First, Arup constructed a dataset to inform case study selection. This dataset included details of the market design, policies and characteristics of International Energy Agency (IEA) member countries. Second, Arup then assessed the available literature for each case study market and policy. Third, Arup then assessed the extent to which case study learnings could be applied to the GB market; this assessment was informed by the dataset constructed in step one. Arup then formed policy recommendations based on the outcomes of the REA and subsequent synthesis step.

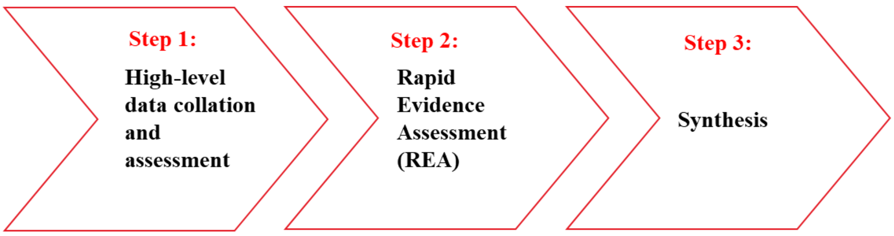

Figure 2: Methodology approach

Note that this document is accompanied by a sister document (the technical annex) where the reader seeking more detailed information on the markets studied in this paper will be able to dive into the specifics. This document focuses on the cross-market findings and the applicability of the various market designs to the GB power system.

Study approach

Step 1: High-level data collation and assessment

Step 1 A: Capturing high-level data on market design, policies, and market characteristics:

Before commencing with the Rapid Evidence Assessment (REA), we conducted an initial study that identified the key electricity market designs and arrangements across International Energy Agency (IEA)[4] members. This initial screening included the challenges being faced in each market. For each market, using our

[4] See Appendix: IEA Countries for further details on review of key market design parameters of the IEA countries.

knowledge and expertise, we have drawn comparisons to the GB market and highlighted which markets are most similar in terms of design, operation, and challenges; this comparison included the following market characteristics:

Market characteristics
Dispatch regime
Settlement period (IDM)
Pricing rules
Pricing design
Total demand
Peak demand
Electricity network decarbonisation target
Carbon pricing
Degree of interconnection (%)
Degree of intermittency (%)
Liquidity (churn factor, forward markets)

Table 3: Market characteristics reviewed for each market.

Step 1 B: Using the above market characteristics to determine case studies, Arup selected markets with close similarities to GB to draw stronger inferences on applicability to our own market; the so called ’doppelganger approach’:

The initial review focused on liquidity, extent of decarbonisation, emission and renewables targets, network constraints, capacity mix, trading exchanges, and key regulation and policy. Arup subsequently narrowed the focus of the REA to those markets most relevant to GB in terms of current structure and potential future changes.

However, the selection logic had to accommodate specific limitations which led to the selection of markets on a de-facto basis. Limitations to the selection approach included:

- Limited number of international precedents (e.g., Centralised Reliability Option).
- Markets moving away from a market design under consideration in the REMA consultation (e.g., Poland and Belgium moving away from Strategic Reserve). Thus these markets were selected on a de-facto basis as these markets were highly relevant markets to explore in order to answer the REMA option specific question on Strategic Reserve.
- Pragmatic considerations: economies of scale were considered when selecting certain markets which had several market designs under consideration implemented in their jurisdiction (e.g., for the operability theme, US markets were favoured due to their nodal market design and because several of these US markets implemented a nodal design and ancillary services co-optimisation conjointly as part of wider market reform).
- More generally, there are inherent difficulties in using a ‘doppelganger’ approach for comparison. There will always be specific intricacies relevant to each country which can be used to argue against its similarity to the GB power market.

Step 2: Rapid Evidence Assessment (REA)

Arup performed a Rapid Evidence Assessment of the academic and policy literature to draw out the key conclusion and evidence:

Arup conducted a systematic and repeatable literature search using the sources outlined in the Figure 3. Arup developed a conceptual framework which defined the parameters of the review including criteria for including and excluding studies. To ensure repeatable searches a key list of search terms was developed. Arup used the following criteria to focus the REA on the most relevant studies.

- **Date**: literature dating back to key market restructuring events, with a focus on more recent literature
- **Topics**: Market designs and arrangements identified in REMA with international precedent
- **Geography**: Limited to countries where relevant designs and arrangements have been considered and/or implemented
- **GB-relevance**: countries with similar market design / arrangements and capacity mix to GB (now or in future) will be given priority.
- To ensure the quality of the underlying evidence used in the report, Arup used studies sourced from peer-reviewed journals, system operator literature, reports from reputable consultancies and excluded general press articles, undergraduate/graduate dissertations.

Additionally, expert interviews were conducted and structured around four key themes (implementation experience, system cost impacts, security of supply impacts, and applicability to GB).

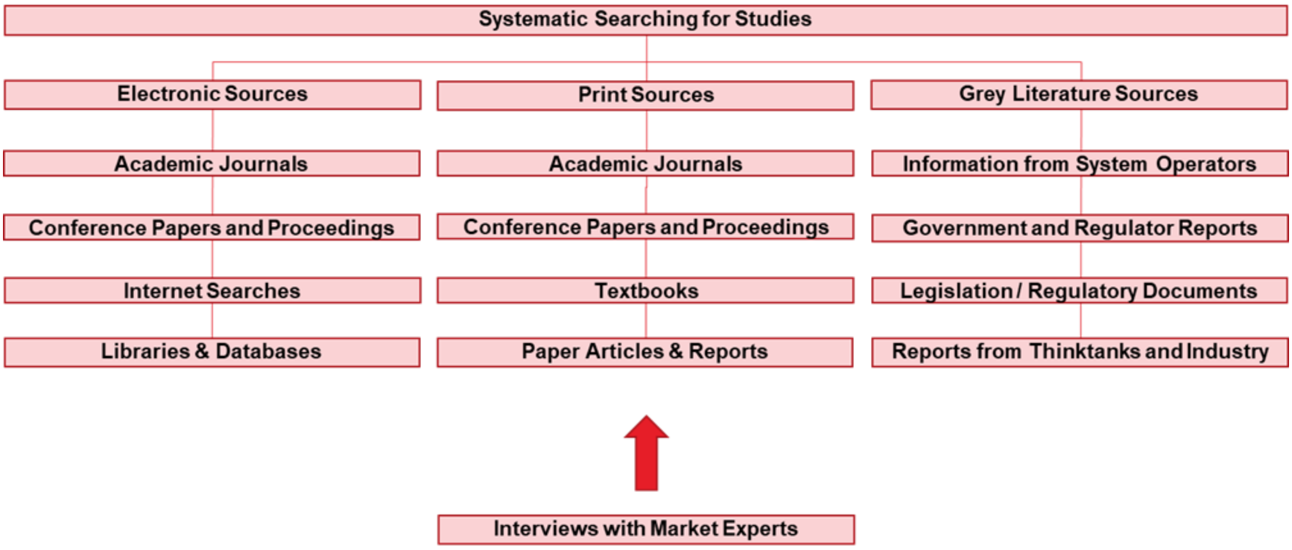

Figure 3: Literature review sources

Wherever there was insufficient literature to draw conclusions on certain types of impact, the report makes this clear consistently. Due to the nature of the research topic, there has been a lack of sufficient literature to address all research questions outlined by DESNZ. Where evidence of the causal effect is unavailable Arup has assessed market characteristics and inferred whether this can be attributed to the market design. Instances in which such conclusions have been drawn are made clear consistently throughout the report.

Step 3: Synthesis

The findings from the literature search were then assessed regarding their applicability to GB. This included considering the international impacts of the market design and how likely it is that these impacts would be replicated in GB. This was done through assessing similarities to GB market characteristics. Conclusions and policy insights were drawn from the synthesised findings and analysis of GB applicability. These cover the main findings relevant to the research questions and their implications for the findings of the REMA Consultation. Limitations and caveats of the literature review are also highlighted. Recommendations for future research are also provided.

High-level rationale for case studies selection

The approach above led to the selection of the following markets for the case studies. The full selection rationale is available in the annex document. Below is a presentation of the high-level selection rationales of the case studies for each market design.

Nodal markets

The criteria that Arup used to select nodal case studies were total and peak demand levels, generation mix and market structure. These criteria were used to find the most comparable power systems size-wise and generation-wise as these are the typical parameters used to describe power systems in academic literature and by practitioners.

The initial market review of all the member countries of the IEA provided a long list of nodal markets as a starting point. The selection criteria were then used to derive the selected nodal case studies. To be selected, a market only needed to meet one, not all criteria.

Figure 4: Nodal Markets Case Studies Selection[5]

Zonal markets

The criteria that Arup used to select zonal case studies were total and peak demand levels, generation mix and interconnection with GB. From the list of potential zonal market options, Australia and Italy were immediately identified as markets of interest. Australia faces similar constraint cost challenges to GB and it considered some of the market design options identified in REMA. Italy was identified given its size and generation mix similarity to GB.

Figure 5: Zonal Markets Case Studies Selection

Centralised Reliability Options

There is a very limited number of real-world examples. International examples within Europe include Italy and Ireland. Examples outside of Europe[6] include Colombia. Colombia was excluded on the grounds of its demand being c.4 times less than GB and its generation mix being too dissimilar.

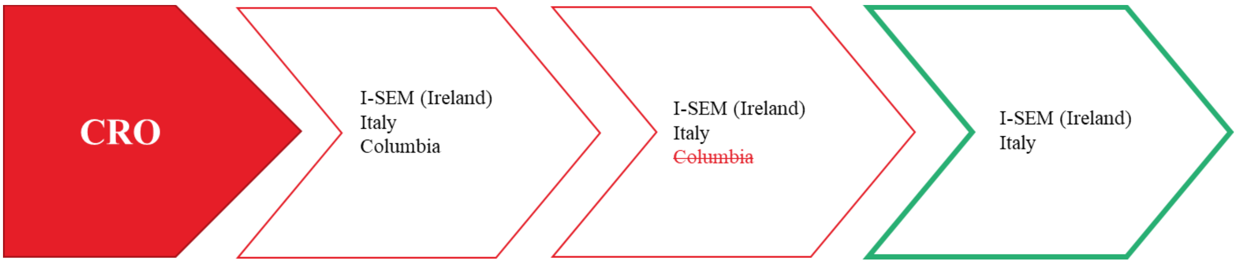

Figure 6: CRO Markets Case Studies Selection

Strategic Reserve

Based on market research, Arup established a long list of markets which had adopted this market design as well as countries having moved away from this market design. From the list of potential options, Croatia and Lithuania were discounted as options given, they are considered too different from GB's market in size, generation mix and security of supply challenges. All other examples were explored.

Figure 7: Strategic Reserve Markets Case Studies Selection

Supplier Obligation

The case study markets were restricted owing to the economic literature focusing on European and US markets. Markets were chosen on the similarities in size and carbon ambitions to GB. Therefore, Belgium, Poland, Italy, and the US markets were selected as key case studies.

[6] Note on the capacity adequacy market design of ISO NE, certain academics such as Mastropeitro et al. (2017)- question its classification as a proper Centralised Reliability Option (CRO) market design due to heavy reliance on administrative pricing rules, the design of the demand response program, and the lack of transparency and accountability. Other academics do classify it as a CRO. After discussions with the Chief Economist of ISO New England (Matthew White), Arup concluded it is not close enough to a pure CRO design to be considered one.

Co-optimisation of Ancillary Services

There are a limited number of options owing to the fact that for optimisation to be an option, the market under consideration must also be centrally dispatched. Typically, most centrally dispatched markets co-optimised markets are nodal in design. This resulted in CAISO, ERCOT, MISO, ISO-NE being identified as key case studies because of their size and relative generation similarities to GB. These markets were treated as group in the relevant chapter.

Case studies: market arrangements overview

In this section, Arup provides an overview of the selected market case studies that will be explored throughout the literature review. The main objective of this section is to answer the following research questions:

- What are the market arrangements and conditions for the identified case?
- How do these arrangements compare to current GB arrangements and the expected characteristics of the GB electricity system in the future?

The map below provides some locational context regarding the case studies researched. Table 4 illustrates the key market arrangements associated with these selected case studies, followed by analysis of the generation mix of each market case study (Figure 8).

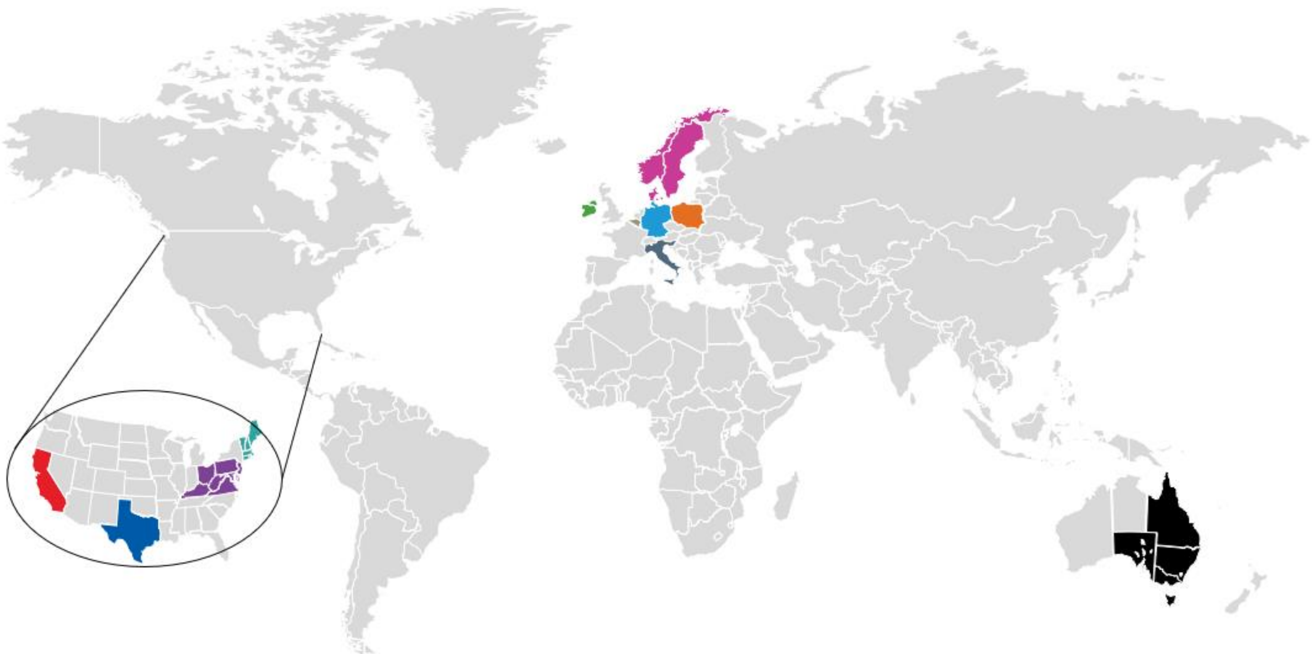

Figure 9: Case studies in this literature review

- CAISO
- PJM
- ERCOT
- ISO-NE
- Nordic markets
- I-SEM
- Germany
- Belgium
- Poland
- Italy
- NEM

Market Characteristic	GB	US Markets (CAISO, PJM, ERCOT, ISO-NE)	Zonal Nordic countries (Norway, Denmark, Sweden)	I-SEM	Germany	Belgium	Poland	Italy	Australia (NEM)
Dispatch regime	Self-dispatch	All central dispatch	Self-dispatch	Central dispatch	Self-dispatch	Self-dispatch	Central dispatch	Central dispatch	Central dispatch
Settlement period (IDM)	30-minute settlement	All 5 minutes	60-minute settlement	30-minute settlement (IDM)	15-minute settlement	15-minute settlement	60-minute settlement	60-minute settlement	5-minute settlement
Pricing rules	Marginal unit sets the price	All LMP	Marginal unit sets market price	Algorithm sets market price	Marginal unit sets market price	Marginal unit sets market price	Marginal unit sets market price	Marginal unit sets market price	Marginal unit sets market price
Pricing design	Uniform	All Nodal	All Zonal	Uniform	Uniform	Uniform	Uniform	Zonal	Zonal
Total demand	294 TWh	CAISO: 211TWh PJM: 806.5TWh ERCOT: 481.8TWh ISO-NE: 118.7TWh	139 TWh	30.3 TWh	504.5 TWh	21.6 TWh	174.6 TWh	289.3 TWh	189 TWh
Peak demand	59 GW	CAISO: 43.9GW PJM: 149GW ERCOT: *No data* ISO-NE: 25.8GW	Norway: 25.2GW Denmark: 6.4GW Sweden: 25.7GW	5.4 GW	81.4 GW	13.6 GW	27.4 GW	49.6 GW	32.8 GW
Electricity network decarbonisation target	Net zero power sector by 2035 and 2050 for the whole economy (October 2021).	CAISO: 50% renewable energy to retail load by 2030. PJM, ERCOT & ISO-NE: no public targets	Norway: reduce GHG by 90-95% by 2050 Denmark: 100% renewable power by 2035. 100% renewable	80% electricity production from renewables by 2030 (2019).	Decarbonise its electricity sector by 2035 (January 2020).	Supports EU carbon neutrality target by 2050.	Agreed PEP2040 which targets 32% renewable energy in the power sector by 2040 (2020).	Increase renewables' share of electricity to 72% by 2030 and to 100% by 2050 (2021).	82% renewable energy by 2030 and net zero by 2050 (2022).

			consumption by 2050. Sweden: 100% renewables by 2050						
Carbon Pricing	UK ETS + The Carbon Price Floor (CPF)	CAISO: California cap and trade, PJM & ISO-NE: RGGI ERCOT: *no data*	EU ETS Sweden: carbon tax	UK ETS EU ETS Ireland: carbon tax	Germany ETS EU ETS	EU ETS	EU ETS Poland: carbon tax	EU ETS	Australia Carbon Exchange
Degree of interconnection (%)	6%	*No data available*	Norway: 14.2% Denmark: 33.3% Sweden: 22.1%	3.9%	9%	14.3%	21.4%	2.3%	0%
Degree of intermittency (%)	28%	CAISO: 25% PJM: 4% ISO-NE: 7% ERCOT: 29%	Norway: 6% Denmark: 61% Sweden: 17%	36%	31%	20%	1%	15%	8%
Liquidity (Churn factor, forward markets)	2.5	PJM: 2.88 *No data available for CAISO, ERCOT & ISO-NE*	Nordic: 2.5	0.5	8.5	<0.5	1.5	2	7

Table 4: Characteristics of case study markets

Generation mix comparison

The chart below compares the generation mix of the selected case studies and GB. The generation mix has been simplified into four categories: renewable energy sources (intermittent and non-intermittent) and dispatchable energy sources (low and high carbon intensity). Intermittent renewable energy sources include wind and solar. Non-intermittent renewable energy sources include biomass, hydro and geothermal. Low carbon dispatchable energy sources include nuclear and waste. High carbon dispatchable energy sources include gas and coal.

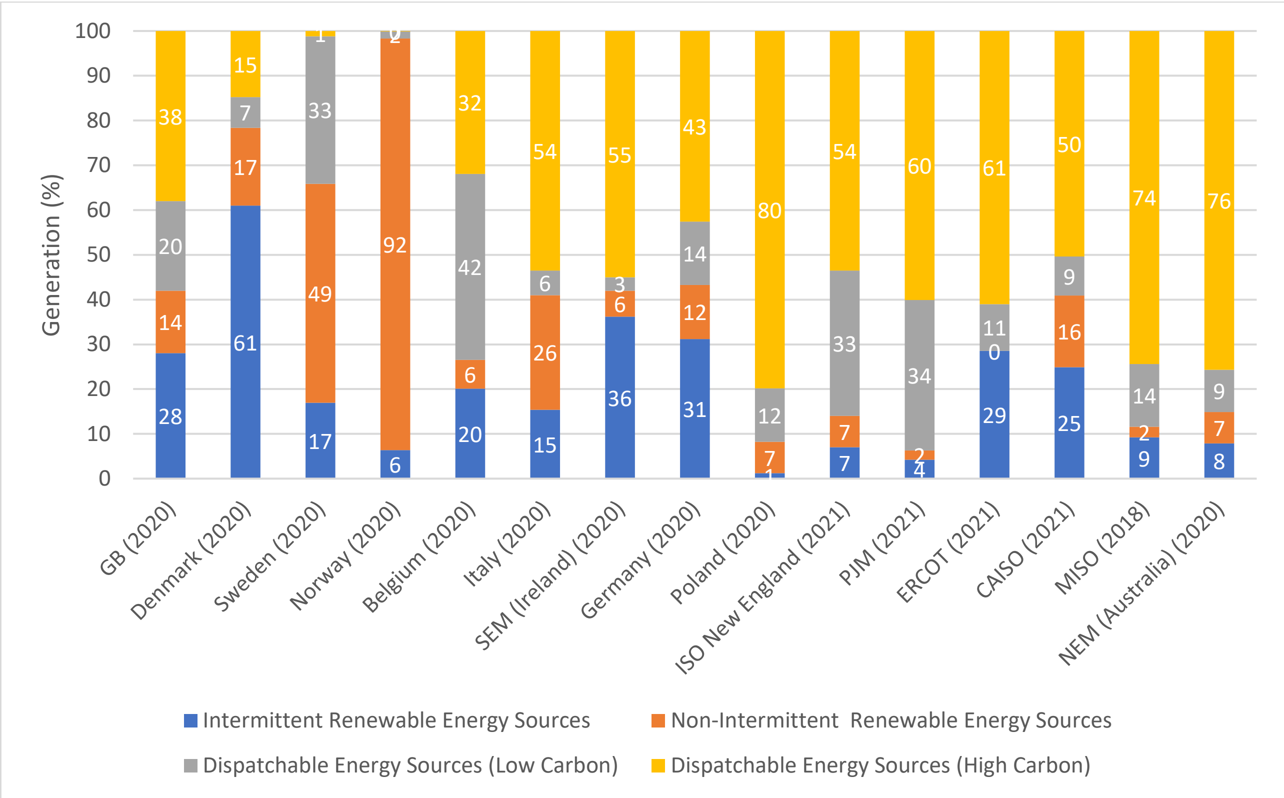

Figure 10: Case Studies generation mix comparison with GB.

Which markets will have comparable characteristics to the future Net-Zero (NZ) GB power system? This is a particularly challenging question to answer. However, markets which aim for high-level of renewable generation / net zero (NZ) target will share high-level similarities with the future NZ GB power system. The key differentiator will be which set of technologies will be underpinning the low (or zero)-carbon dispatchable generation. This question will be answered -in part- by the policy choices made and the natural resources available these markets.

The following section presents the cross-market findings for the nodal and zonal wholesale market designs and discusses their respective applicability to GB.

Wholesale market designs

Nodal pricing

This chapter explores the international experience of nodal market design with the aim of addressing the research questions set out previously. It covers four case studies: CAISO, PJM, ERCOT, and ISO-NE.

Cross-market findings on nodal market design

The following section brings together the findings the nodal market design from each case study (CAISO, PJM, ERCOT and ISO-NE) and uses these to draw aggregate conclusions.

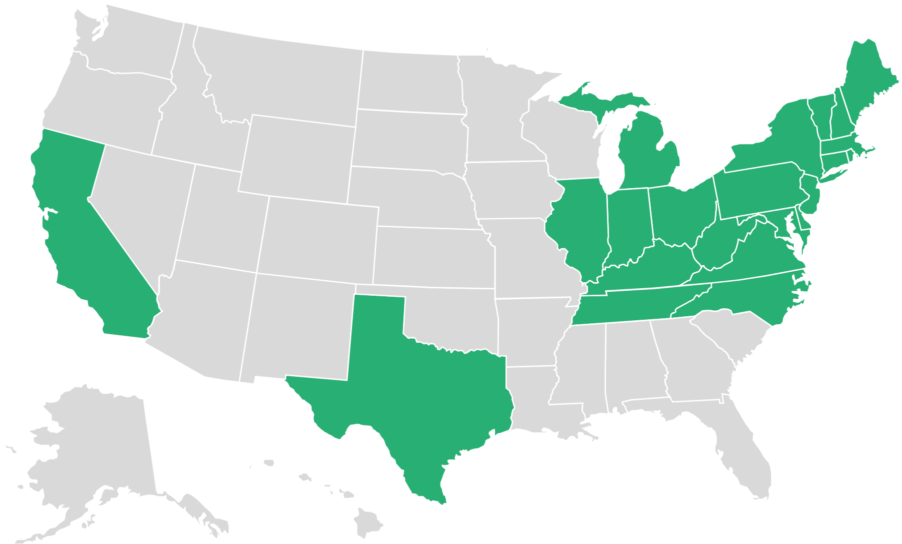

Figure 11: Overview of the US market case studies

Market design rationale

During the late 1990s and the start of the century, liberalised US markets began to switch to a nodal pricing design. The CAISO switched following an energy crisis which was worsened by the then current market design. PJM, ERCOT, and ISO-NE all highlighted high network congestion and subsequent high cost of managing this as a large driver of the rationale behind the switch to a nodal market.

The rationale is explored by Wolak (2021) who conducted a comparative review of the recent evolution of the wholesale market design in the US and Europe. The study notes that:

- Virtually all US wholesale electricity markets started as either a single pricing zone or zonal market with a day-ahead forward market. This meant that generation, load schedules, and prices were determined at the zonal or single-pricing zone level.

- These markets relied on congestion management or redispatch processes run by the system operator to ensure final schedules were physically feasible given the transmission network's configuration.
- Market operators found that transmission congestion was significant and increasingly expensive to manage in zonal or single-pricing zone markets. As a result, they adopted a more spatially granular pricing model, which led to the development of nodal market designs.
- PJM Interconnection was the first US market to adopt a nodal market design in 2003, and others followed, including New York ISO, California ISO, and ERCOT.
- European markets have managed to maintain their zonal/single zone market designs due to their history of state-owned utilities and significant investments in transmission capacity from the 1970s until the start of the restructuring process.
- However, as the share of intermittent renewable energy has increased in many European countries, the cost of making final schedules physically feasible has grown significantly. This has led some European countries to consider adopting more granular approaches to electricity pricing.

Both ERCOT and CAISO first switched to a zonal market design from a uniform price one. In both instances policy makers switched to a nodal design for similar reasons. Both markets saw increasing constraint costs and growing levels of generator re-dispatch, which appeared to be exacerbated by the zonal design. In addition, there were concerns that market participants were gaming between the wholesale market and balancing services, enabling them to exploit market power (see e.g. the Californian energy crisis in the early 2000s).

These findings resonate with the current situation in GB. There are growing concerns with constraint costs in GB and re-dispatch. There are also increasing concerns of generators gaming between the wholesale market and the balancing market. Such concerns have led Ofgem to introduce a licence condition to mitigate this.

The switch to a nodal market design in the US markets were part of packages of reforms that fundamentally changed how the markets were structured and designed. The US markets ended with relatively similar designs all based on economic principles of using granular price signals (in location and time) to ensure cost reflective prices and efficient dispatch. In addition, to address concerns with Transmission Owner monopoly power, Independent System Operators were established. The four main pillars of market reform in the US were:

- **Independent System Operators** - These entities were changed with efficient running of the network, planning the network, and connecting new assets to the grid. Some experts have highlighted this independence as being crucial in developing the confidence of market participants.
- **Centralised Dispatch** – These were based on algorithmic software – nodal markets require large quantities of constrained optimisation equations to be performed simultaneously.

- **Shortened settlement and gate closure periods** - Generally, the US markets moved to 5-minute settlement periods, some after a dalliance with 15-minute settlement periods, such as CAISO. In addition, many of the markets also employed a real-time market (closing 30-minutes ahead of real-time), or an intra-day market to work alongside day-ahead dispatch.
- **Retail price regulation, and the use of FTRs**– In response to the energy crisis in California some US markets mandated the hedging of power between 2-3 years out. There is also reasonably strong retail price regulation in the US, with many domestic tariffs capped and rules on what prices must be offered to consumers. To support hedging, all the case study markets include FTRs (though not all are mandatory).

System impacts

The evidence from policy evaluations on US markets and comparative studies all suggest evidence of the benefits from moving to nodal markets. Whilst most of the benefits identified have been the transfer of congestion rents form generators to consumers, some studies have also shown improved overall efficiency. It should be noted that estimating the costs and benefits of a move to nodal pricing is inherently difficult because it so difficult establishing a robust counterfactual.

International empirical studies indicate that, following the implementation of nodal pricing and central dispatch, generators incurred substantially lower operational costs through more efficient fuel use and ramping - c.2.1% in CAISO (Wolak, 2011) and 3.9% in ERCOT (Triolo & Wolak, 2021).

Neuhoff et al. (2011) as well as the Brattle Group (2017) are key studies which have explored the implementation costs and annual consumer benefits of the market reforms conducted by several US ISOs. Arup assessed the methodology used in these two studies and deem them reasonable. In all cases the annual consumer benefits at least equalled the one-off implementation costs as demonstrated by the figure below. For PJM, the analysis suggests the annual consumer dwarfed the one-off implementation costs by ten times.

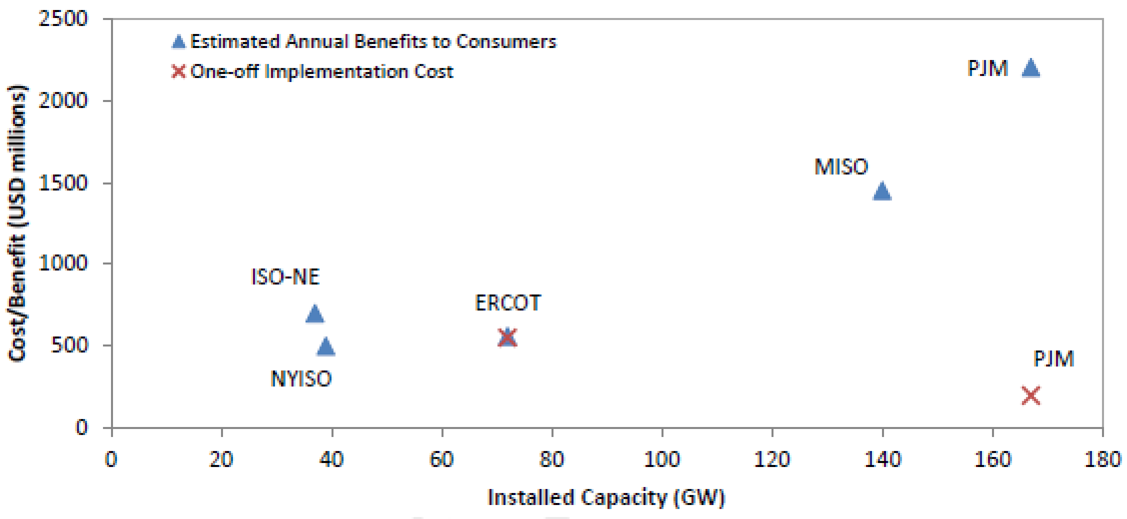

Figure 12: Annual benefits and one-off implementation costs vs. installed capacity. Source: Neuhoff, 2011

According to Neuhoff, implementing nodal market designs in the US takes advantage of existing institutions, skills, and experiences, resulting in lower implementation costs due to the need for specialized IT software, hardware, and staffing costs. Furthermore, US nodal markets combine congestion management with system-wide optimization at both the day-ahead and intraday stages, respecting transmission constraints to avoid costly readjustment and keeping the system in balance[7].

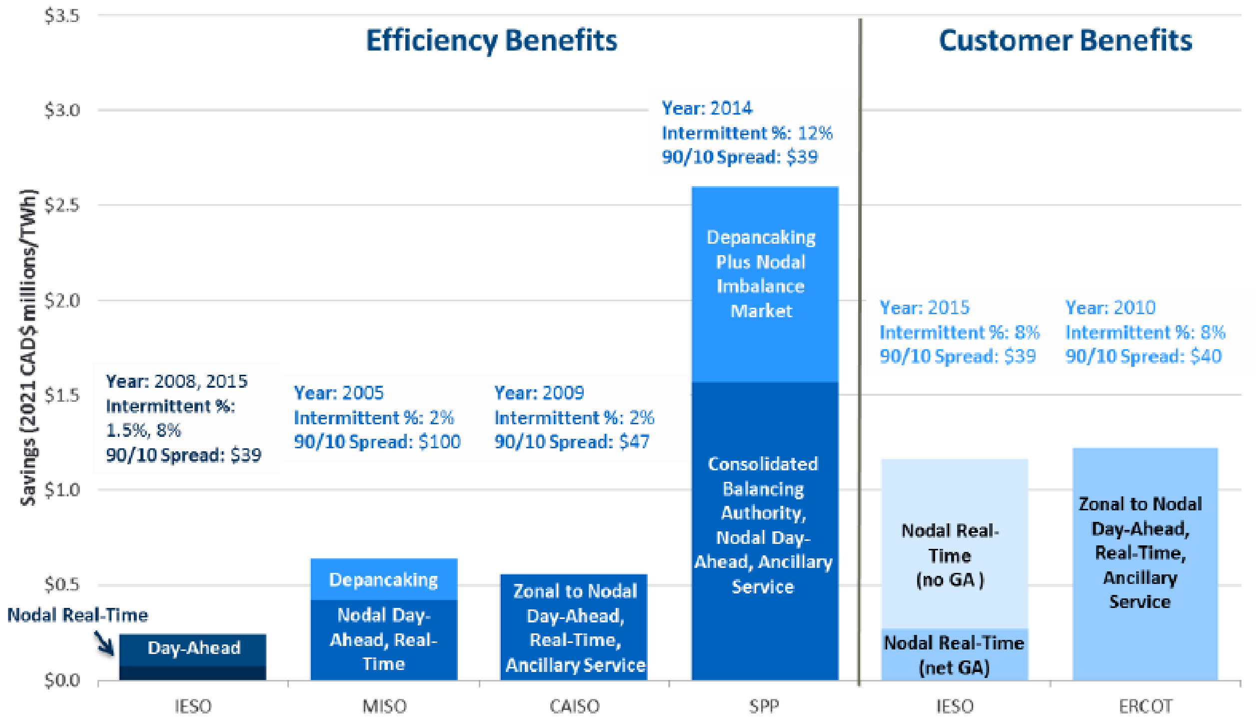

Figure 13: Estimated Benefits of Major Market Re-design Expressed in Savings per TWh of Market-internal Load (Brattle Group, 2017)[8]

The Brattle Group (2017) compared the benefits of the market reforms conducted by CAISO, ERCOT, PJM, SPP, IESO (Ontario's system operator) and MISO (Midcontinent Independent System Operator), shown in the Figure above. The authors' aims were to assess whether such benefits could be transferred to the Ontario market for future reforms in 2021. The research used previous studies' estimates of the benefits and attempts to make these comparable to Ontario. To account for the differences in market size, the study normalised the magnitudes of the benefits estimated by the annual load (TWh of annual energy consumption) in each of the markets. In addition, they estimated system-wide and customer savings for those studies on a

[7] Nodal-Pricing-Implementation-QA-Paper.pdf (climatepolicyinitiative.org)
[8] Figure expressed in 2021 prices assuming 2% inflation.

dollar per TWh basis, expressed in 2021 Canadian Dollars. The authors of the study highlighted the benefits of the CAISO and ERCOT reforms.

The CAISO study describes how the previous zonal market resulted in large inefficiencies with intra-zonal congestion management. Specifically, prior to reform, generation resources in CAISO system were self-committed on a day-ahead basis, but often needed to be ramped down at high costs. The CAISO energy market enhancements in 2009 included implementing a nodal real-time and a financially binding day ahead market. Those enhancements helped manage intra-zonal congestion much more effectively than the prior system due to more efficient day-ahead unit commitment, dispatch, and settlement. The study of the 2009 CAISO enhancements identified a USD $105 million annual reduction in production costs[9]. According to CAISO's 2009 Annual Report on Market Issues and Performance, ancillary service costs to consumers also reduced; these went from 1.4% of wholesale energy costs (USD $0.74/MWh) in 2008 (under the prior design) down to 1.0% (USD $0.39/MWh) in 2009.

As in CAISO, ERCOT recognised the inefficiencies associated with its intra-zonal transmission congestion management under its zonal energy market. Pre-2010, ERCOT managed intra-zonal congestion by instructing generating units to ramp up or down from original scheduled output levels in response to the system's needs. Units that followed these instructions were compensated with out-of-merit energy payments. Therefore, the generators in ERCOT did not have incentive to consider the state of the transmission congestion when self-scheduling because they were paid to lower their generation output in real time. The 2010 market reforms (moving from zonal to real-time nodal pricing and day-ahead markets) addressed these issues. The reform created large estimated savings for consumers. An ex-post study[10] of market prices estimated approximately a 2% reduction in consumer costs across the system. This reduction accounted for decreases in wholesale power prices, uplift costs, and congestion payments. These reductions in costs partially reflected the consumer savings associated with higher congestion rent being returned to customers under the nodal market compared to the prior zonal market. Unlike the other studies, the analysis of ERCOT's reforms did not include an estimate of system-wide efficiency gains. Rather, it only estimated impact on ERCOT customer costs.

However, it is important to note that doing an ex-post evaluation of the move to nodal market and estimating the benefits is inherently difficult. The principal difficulty is due to the need to establish a robust counterfactual – the outcomes had the market not moved to a nodal design. We have cited studies above where significant efforts have been made to control for the numerous variables that can affect the impacts. However, this remains an inherently difficult challenge, especially is trying to assess the impacts over a long timeframe.

The studies examined explore the benefits of the US reforms as a package (not the introduction of nodal markets in isolation). It is almost impossible to disentangle the benefits of

[9] Wolak (2011), "Measuring the Benefits of Greater Spatial Granularity in Short-Term Pricing in Wholesale Electricity Markets," American Economic Review: Papers and Proceedings 2011, 101: 3, May 2011.
[10] Zarnikau (2014).

the nodal market reforms from those aimed at great temporal granularity (shorter settlement periods, gate closure, real time balancing energy markets). This should be borne in mind when considering the benefits from US markets.

Nonetheless, the available evidence appears to support the theory that nodal markets transfer welfare gains to consumers from generators.

Market liquidity

In Europe, there is some concern that the large number of nodes in nodal pricing regimes creates a high level of complexity. This hinders the attempts of market players to find counterparties for longer term energy contracts. GB policy makers have had similar concerns that nodal markets will suffer from low levels of liquidity.

However, this has not been the experience of the US nodal markets. In the US, trading hubs emerged and are based on a relatively stable average price across a set of multiple nodes. Trading at and between these hubs is very liquid in the forward exchanges (e.g., the New York Mercantile Exchange (NYMEX) and Intercontinental Exchange (ICE)). In PJM's case, standardised power contracts for both day-ahead and real-time market delivery are defined for 20 trading hubs. NYMEX provides a monthly contract for PJM's Western Hub (800MWh total), which is likely to be the most liquid forward electricity market in the world.[11]

The US nodal markets tend to have good levels of liquidity up to three years ahead. This has, in part, been facilitated by the use of FTRs that enable congestion costs to be financial. In the case of PJM, forward market liquidity has developed organically due to market participants needing to hedge risk. In the case ISO-NE, forward trading has been supported by the regulations on suppliers which mandate them to hedge 80 percent of their output three years ahead. To note in GB, the market is not very liquid at the six months ahead stage and dries up more than 1 year ahead.

Consumer exposure

The figure below illustrates different degrees of consumer exposure to nodal prices. This ranges from "full" nodal pricing, where both generators and consumers are exposed to nodal prices (ERCOT), a mixed approach, where consumers pay the weighted average of the nodal prices in the region (PJM, CAISO), and nodal dispatch but zonal settlement (Ontario).

[11] Nodal-Pricing-Implementation-QA-Paper.pdf (climatepolicyinitiative.org)

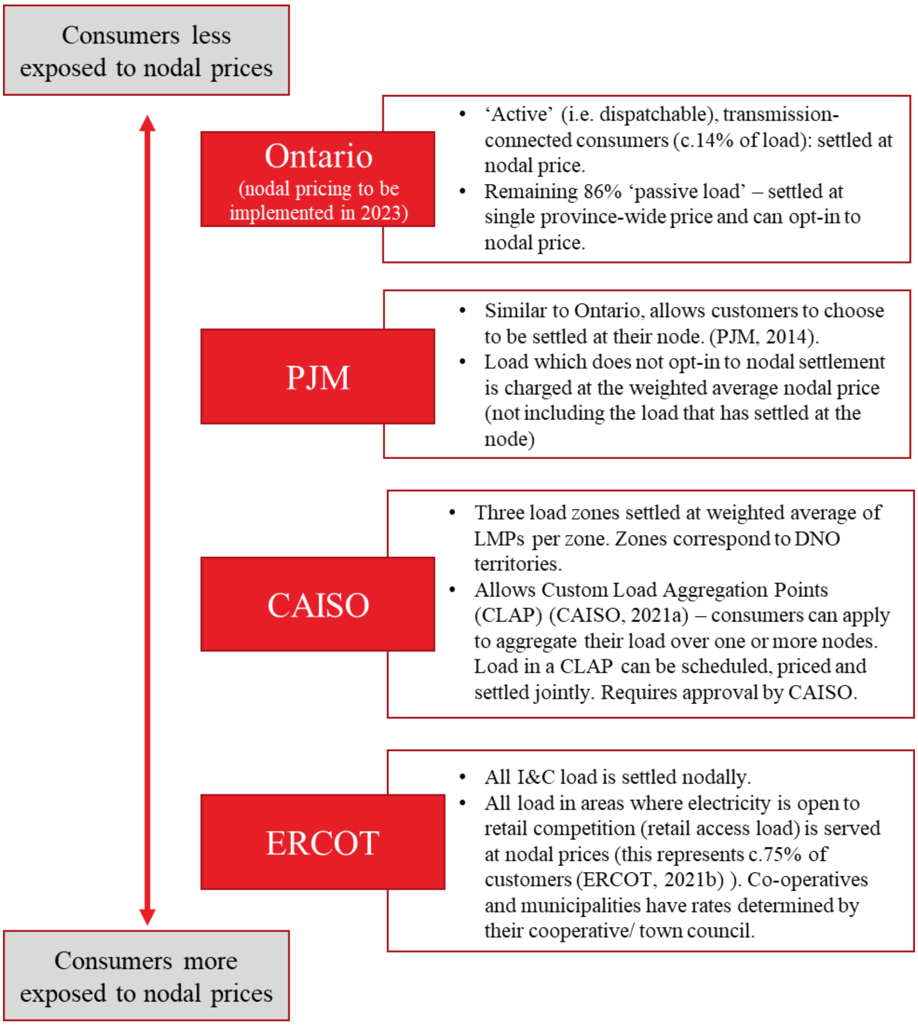

Figure 14: Degree of consumer exposure to nodal prices (Ontario, PJM, CAISO and ERCOT)- Source: NGESO MRNZ Phase 3

Neuhoff (2011) showed that no US market has full nodal pricing at the retail level. Instead, prices are often (demand weighted) averaged at zonal level (as shown in the figure above). CAISO is shown to have the least variability between domestic retail prices with only a minimal number of zones. PJM has the greatest number of zones and the greatest variation between them. This use of zonal pricing for retail consumers lessens the potential for effective demand side response but provides some support for consumers. The recommendation from the Federal Energy Regulatory Commission (FERC) is to favour disaggregation of zones.

ISO Operating Region	Number of Nodes / Buses	Aggregated Retail Pricing Regions	Wholesale Pricing Hubs
California ISO (CAISO)	3,000	3 regions with 23 sub-regions: covering the majority of one State	-
Midwest ISO (MISO)	1,300	7 pricing regions: covering all or most of 13 States.	6 trading hubs
ISO-New England (ISO-NE)	900	8 pricing regions (match 6 State borders – Mass. has 3 regions)	1 trading hub
PJM Interconnection (PJM)	6,000	18 pricing regions: covering 13 States and Wash. DC.	20 trading hubs
New York ISO (NYISO)	-	11 pricing regions: covering one State.	Weighted region prices used

Figure 15: Granularity of various ISO load regions Source: Neuhoff, 2011.

To protect consumers against spot market fluctuations, many of the US nodal markets have policies that mandate the forward purchasing of power by suppliers to protect consumers against spot market volatility. The clearest example is ISO NE, which mandates that suppliers hedge 80% of their demand 3 years out. The moves to mandate hedging were a reaction to the Californian energy crisis in which spot market volatility played havoc with unhedged retail suppliers[12]. The fact they faced absolute price caps meant they could not pass on the increases in wholesale costs. This situation has similarities to the experience in the GB market, which regularly saw unhedged suppliers go bust following the introduction of the ‘price cap’.

All the US markets are physical spot markets at the day-ahead, intraday real-time only. This means that forward trading hedges are not physical, but financial hedges based on CFDs against the spot price. In addition, the US nodal markets also use FTRs to help suppliers manage financial risk and hedge. FTRs do not offer the holder the right to physically transport electricity. Instead, FTRs give the holder access to financial compensation equal to the congestion and/or loss rent associated with the locational price differences. They are also referred to as Congestion Revenue Rights (CRRs) in some of the US markets. FTRs are an important central component of nodal pricing systems, but the number of years for which they are valid and whether they are defined as obligations or also available as options, and the principles that are used for such free allowance allocation, differ across markets.[13]

The Energy Policy Act of 2005 requires the FERC to conduct an annual survey of the demand response (DR) and advance metering in the US. Various iterations of this report have suggested progress is not fast as desired. According to the 12th edition of the report “since 2009, demand resource participation in wholesale markets has increased by approximately 6% but has been outpaced by an approximately 16% increase in peak demand”. It goes on to add that “ironically few are used to deliver time-of-use (TOU) or real-time prices (RTP)”. It seems to suggest that the US markets have not used greater locational and temporal granularity to try and stimulate demand side response, and this is likely to be for political reasons.

[12] Source: Interview with the chief economist of ISO-NE (Matthew White)
[13] Nodal-Pricing-Implementation-QA-Paper.pdf (climatepolicyinitiative.org)

Nodal market design: applicability to GB

The following section explores what inferences we can draw for the GB market from the experiences of nodal markets. Whilst it is not possible to fully generalise the outcomes observed to GB given differences in market design and characteristics, the experience from other countries suggests some evidence in support of a nodal market design.

Market design rationale

The evidence from US markets that moved to a nodal market suggests they were principally driven by concerns with increasing constraint and re-dispatch costs. Generators dispatch against the desired market outcomes. In some markets, there were also concerns market participants were able to game the existing design (be it uniform or zonal). Such behaviour pushes up prices and this was especially so when CAISO first implemented a zonal market.

This resonates with the experience in GB. As highlighted in the case for change by DESNZ and the ESO, constraint costs have risen dramatically and are forecast to further increase significantly in the future. Re-dispatch levels have also risen significantly; the ESO often must un-do and then create the desired outcome to balance supply and demand across the network. However, it should be noted that constraint costs in GB have risen dramatically since the connect and manage policy 'has encouraged renewables to connect to grid irrespective on transmission issues. This was not the case in the US markets.

In addition, the GB market has increasing concerns with generators withholding power from the open market to gain an excessive benefit. Ofgem is now intending to implement rules to prohibit excessive benefit from such situation. It is no coincidence the language is that of competition law because the fundamental concern is with the exploitation of market power.

It appears that the GB market is suffering from similar symptoms that led policy makers in the US to choose a nodal market design. This conclusion has also been drawn by the Florence School of Regulation with respect to European markets, not just GB. Their paper suggests that European markets suffered the problems of constraints costs inherent with uniform pricing design later than the US liberalised markets. This is likely to have been a function for European markets initially due to their greater network capacity.

It should be noted though, GB's objectives with market reform are wider than the reducing constraints costs, with key objectives around enabling low carbon investment.

Implementation and design

There is a wide range of implementation periods taken by the US ISOs to set up a nodal market, ranging from 1- 9 years. However, this range is affected by markets that initially implemented a zonal design, before moving to a nodal design. The markets that were clear about their design choices were able to implement a nodal design more quickly, with timelines ranging from 1 to 5 years.

FTI echoes this view (with a higher minimal time though) noting that transition to nodal designs predominately depends on the efficiency of the stakeholder engagement and usually takes between 4-8 years[14]. The PJM nodal market took approximately a year to implement, whereas CAISO market reform was initially proposed in 2000 but it was not officially implemented until 2009.

NGESO[15] has estimated a timeframe of five years to implement a nodal market design in GB. This estimate is consistent with the experience of other US markets that did not follow the incremental approach. Arup considers five years to be sufficient time for the ESO and market participants to design and implement a nodal market in GB. It is more likely that the limiting factors will be the need for political and stakeholder support, and passing necessary legislation. It is likely that, as in Ireland and Australia, generators will lobby hard against the proposals.

Nothing physically must change with a nodal market, but a new economic model market design must be set out. Elements of the model include:

- **Nodal market data** – To generate nodal market price data, the ESO would need to know all generation points on the transmission system. This data is already available as each is identified with its own Balancing Market Unit identification. It also needs the shadow price of each node. The shadow price is the cost of the generation unit, plus transmission losses and constraint cost. Transmission losses are calculated as part of the system design by Elexon[16]. Constraints costs are currently calculated by the ESO via its tagging and flagging or system actions[17].
- **Design and consultation on new market parameters** – Parameters such as settlement period length, gate closure, imbalance penalties and incentives, market clearing, and scheduling must be decided. This should be relatively straight forward for the ESO, it is already considering these design parameters. International experience suggests early and sustained stakeholder engagement results in more successful implementation. The ESO has a strong track record in working with stakeholders in delivering system changes.
- **Centralised dispatch software to allow algorithmic dispatch** – The ESO would need optimisation software that performs the mathematical constrained optimisations. However, it should be noted that the ESO has faced challenges in the past when it comes to delivering IT projects. For example, struggling to deliver a digital Electricity Balancing tool despite a £100 million spend. There are limited number of vendors that sell this optimisation software, and the ESO can learn from the experience of other US ISOs in procuring this. Matthew White, the Chief Economist of ISO-NE, suggested that

[14] FTI (2022): Net Zero Market Reform: Phase 3 - Assessment of market design options
[15] NGESO (2022): Net Zero Market Reform Phase 3 Assessment and Conclusions
[16] Transmission Losses - Elexon BSC
[17] SMAF-20111223-FINAL (nationalgrideso.com)

not much additional information was needed to run an LMP based market and that optimisation software would be available from external companies.

- **Set-up day-ahead and intraday market clearing and settlement** - Nodal markets require a 'market operator' to physically run the day-ahead and intraday markets as well settlement of transactions. This does not have to be the System Operator, but this has generally proved to be the case. The ESO already has a subsidiary that performs settlement (Elexon), and it has significant experience running the BM. Therefore, suggesting it could transition to the role of market operator easily.

As mentioned, the nodal market model was part of package of reforms. We consider that a Centralised Dispatch approach is inherent in a nodal market design, so do not include this here. If GB was to implement a nodal market it is likely to, at the very least, also explore:

- **Retail market reforms** – US markets mandate hedging for suppliers. For example, ISO-NE mandates hedging of 80% of demand 3 years out. GB should consider this to help ensure deep and liquid markets. This is also likely to have other benefits too. It would have prevented situations in which suppliers go bust due to their lack of hedging, such as when wholesale price rose dramatically. It would have resulted in a much lower price shock induced by Russian's invasion of Ukraine.
- **Price regulation** – All the US markets have some degree of price regulation, and it is likely GB would need to consider this too.
- **Independent System Operator** – A key element of the US reform packages was an independent system operator charged with running and planning the network. GB has taken these steps with the establishment of the FSO. This is an encouraging sign should GB choose to implement a nodal market.

Cost

There is a wide range in costs of implementation for nodal markets in the US, approximately $100-$500 million for the PJM and ERCOT. These costs appear low, against the expected benefits (discussed below), and focus on the costs to ISOs for implementation. It is difficult to transfer these costs over to GB. However, there should be some doubt that such low implementation costs would be replicated in GB. Previous estimates of much smaller changes have indicated much higher costs. For example, Frontier Economics for a report ENTOS-E estimated the impacts of shortening settlement periods alone to between £0.8-£1.5 billion.

Consumer impacts

The evidence suggests that US consumers have benefited from the nodal market design. The evaluations undertaken have suggested significant consumer savings from reduced constraint management. This corresponds to the theory that nodal markets transfer congestion rents from generators to consumers. It should be noted that conducting such evaluations are inherently difficult. This is because of the need to set out the counterfactual states - what would have happened in the same market had it not been nodal.

It is also important to note that the US markets have a heavy degree of retail price regulation and protection. It is also clear that consumers do not actually face nodal prices (although there a few cases of commercial customers doing so). In several of the variants of the nodal design, only generators face nodal prices, whereas suppliers face zonal or average prices. While this lessens the potential for effective demand side response from consumers, it provides safeguarding for consumers.

There are further consumer safeguards because of the opportunities created to hedge future risk. FTRs are a different approach to that of GB and Europe. They do not offer a physical hedge, but rather a financial hedge that compensates suppliers financially if they have FTRs and there is congestion. Neuhoff (2011) points out "that FTRs are a central component of all nodal pricing systems, but the number of years for which they are valid, whether they are defined as obligations or also available as options, the shares of FTRs that are allocated for free, and the principles that are used for such free allowance allocation, differ across regions."

One of the concerns voiced previously in the GB is that a nodal market will lead to a postcode lottery of electricity prices. This has not been the case in the US. The US markets have tended to not let consumers face full nodal prices but rather zonal prices, and with fairly large zones. Whilst this does result in some distributional difference, these can be managed by policy choices.

Other design options could be explored whereby consumer groups are differentiated e.g., Industrial and Commercial (I&C) facing full nodal prices. However, to really stimulate effective demand response economic efficiency would suggest designing demand zones at very disaggregated levels. The impact on consumers in terms of geographical distribution is fundamentally a choice for policy makers.

Some nodal markets have had good success in promoting DSR (PJM, CAISO) however, this has not been the case with all US markets. The FERC has placed significant emphasis on demand side response in the US. However, evidence suggests that peak demands have been increasing and growth in DSR is outpaced by demand growth[18]. The fact that US markets have such strong consumer price regulations has lessened the incentive for DSR. There are very limited uses of the tariffs. Where there has been more success in developing DSR, it has largely been the result of other policies.

Impact on investment and deployment of low carbon technologies

The evidence from the US markets suggests that nodal markets were not a barrier to generation of investment in low carbon technologies in the case studies considered. CAISO and ERCOT have both seen significant growth in renewables and low carbon technologies. There have been measures to help integrate intermittent renewable technologies with real-time

[18] DSR annual report, 12th edition (FERC).

markets in CAISO and ERCOT, such as helping generators manage their risk (because they can adjust position depending on the prevailing wind or solar conditions).

There is no evidence to suggest that the market reform transition hindered investment. The growth in RES capacity post nodal market implementation was largely supported by investment support mechanisms in place in the US markets studied. The evidence from CAISO suggests that investment continued despite the prolonged implementation period. A major concern expressed by stakeholders in GB is to that a prolonged period of market reform will create an investment hiatus. However, this was not the case in the US and these concerns have not been raised by developers there. It is generally accepted that market reforms create uncertainty for investors who generally like to have price and volume clarity. However, we think that if GB maintains its strong policy mechanisms to support low carbon generation, for example CfDs, investment can be maintained. The CfD would de-facto minimise the risk of an investment hiatus during a market design transition due to its revenue stabilisation design. Risks can also be mitigated with a clear and transparent reform pathway. Furthermore, in terms of investment, it is often the level of support relative to other markets that is the key in attracting new investment.

A further concern expressed by stakeholders in GB is that a nodal market will increase the cost of capital for renewable investors. The evidence in the US suggests this is not the case. Feedback from US experts suggests renewable developers are adept at using CfD using nodal prices. Arup sees no reason why a nodal market would raise the costs of capital in GB per se assuming CfDs remain in place. However, it is likely that a nodal market will adjust where the investment is made and change the attractiveness of some generation projects. This is because some generation assets behind constrained areas would see reduced payments in the Balancing Mechanism. This is essentially because zonal or unform pricing gives an implicit subsidy to generation projects in constrained areas because they are paid to not to produce. In a nodal market they would not be. However, that is one of the keys aims of nodal pricing; it gives the clearest signal as to where new generation investment should be located. Over time, Arup would expect to see a nodal market improve investment for flexible technologies such as batteries and potentially other types of storage (compressed air). This is because these technologies can relatively easily move location and choose to locate in higher priced nodes.

Impact on liquidity

A common concern among European policy makers is that nodal markets will suffer from low competition at nodes and low liquidity (Neuhoff 2011). However, this contrasts with the evidence and experience of the US markets which have enjoyed good level of liquidity, up to 3 years ahead of delivery. In GB, prior to the recent energy crisis, you could perhaps hedge up to a year out, but currently are more likely to be able to buy or sell power for only six months ahead. It is not clear from the evidence whether nodal markets themselves have improved liquidity but rather it is the use of FTRs and policies mandating hedging that have supported liquidity. For example, ISO-NE mandates suppliers hedge 80% of their demand for three years ahead. Should GB implement a nodal market, the evidence reviewed here suggests FTRs should play a vital role. The mandating of hedging several years out also appears to be an attractive policy that supports liquidity and risk management.

Impact on security of supply

There is no evidence to suggest that nodal markets have adversely affected security of supply. The market reforms were, in part, a reaction to an energy crisis which saw rolling blackouts in California and New York. These were not genuine capacity adequacy issues, but rather the result of spiralling of the poor market design that enabled market manipulation. In general, there has been an improvement in the security of supply in US markets. However, capacity adequacy has generally been driven by the capacity markets in the US states. It is more likely that whatever the policy design, a capacity adequacy policy is required. The nodal market design appears to work well with US capacity markets. This suggests that they could work well with GB's Capacity Market too.

Some of the evidence suggested that the greater information and the centralised dispatch approach has helped improve the ability to manage frequency on the system. However, this is likely to be, in part, due to many of the generators in US markets having Automatic Generation Control (AGC) technology. This technology allows the ISO's to remotely adjust the generators output to manage frequency. In ISO-NE, for example, CCGTs and pumped hydro provide this service. This gives ISOs additional frequency control measures and provide confidence in algorithmic dispatch. However, generators in GB do not have this technology fitted and GB is an outlier in this respect. If nodal markets are implemented in GB, we would expect the ESO to investigate the use of and mandating of AGC technology.

Zonal pricing

This chapter explores the international experience of a zonal market design with the aim of addressing the research questions set out previously. Arup considered five case study markets: Australia (NEM), Italy, Sweden, Norway and Denmark.

Cross-market findings on zonal pricing

Market design rationale

The rationale underpinning the implementation of a zonal design in several markets was to address transmission constraints issues. For example, prior to the introduction of the Nordic zonal market, Sweden had significant transmission constraints leading to bottlenecks in the system. This was exacerbated by production being concentrated in the north of the region while demand was concentrated in the south. Similar issues were experienced in the Norwegian market. Denmark also experienced network congestion from bottlenecks, however, these were naturally created due to the several islands within the country's geography. To address these issues, the three markets created the Nordic transmission system, and all implemented their own zonal markets.

Both Italy and Australia (NEM) implemented their zonal design between the late 1990s and early 2000s, at the same time as the nodal framework was beginning to gain traction.

The Italian wholesale power market is separated into geographical bidding zones which reflect transmission constraints in the country. As pointed out by the Italian TSO, Terna, “right from the start the national territory was modelled in the energy markets in the form of market zones. This happened partly for the geographical conformation, and partly to differentiate the purchase prices according to the balance between electricity generation capacity and demand which varies from zone to zone (providing opportune "price signals")”.

Australia (NEM) has operated a zonal market since the market initially became operational. The NEM initially considered both a zonal and nodal design. They settled on zonal due to a lack of real-world supporting evidence supporting nodal markets at the time. Specific concerns with the nodal approach included its complex nature, concerns over market management, and a lack of backing of certain stakeholders - notably the generators.

However, as the evidence from nodal markets has emerged and issues with the NEM system have become apparent, the Australia government still decided against a nodal design and instead went for a zonal design plus an optional market: the Congestion Relief Market (CRM) . The NEM has recently been highlighted in the REMA Consultation as an example illustrating the uncertainty of the extent to which zonal pricing delivers increased dispatch efficiency. This issue illustrates the zonal designs lack of success in recent history and is one of the reasons the NEM was still considering - until recently - a switch to a nodal market. A Cost Benefit Analysis of the proposed switch to a nodal market design in the NEM estimated the expected social benefit to be between £283-£649 million per year ($382-$877 million per year), illustrating the opportunity cost from operating a zonal market.

Implementation and design

The implementation experiences of international zonal markets have slightly differed. There is no mention of implementation issues in the Australian market within the literature, which may suggest it occurred smoothly. It was only once in operation that problems with the zonal system became apparent.

Italy conducted technical trials prior to its market implementation. These were a year long and aimed to establish an infrastructure and mechanism that ensured a transparent and competitive market. It could be implied that these trials helped contribute to the successful introduction of the design. Since becoming operational, Italy has increased the number of its bidding zones.

The Nordics experienced some issues with system expansion. The Swedish market delayed the implementation of an additional two bidding zones by four months. However, this was due to the initial proposed timeframe being overly ambitious given the need for extensive stakeholder engagement. This implies implementation is more successful with an extended timeframe, due to the complex nature of the system.

Zonal markets can provide strong locational signals however this is dependent upon how accurately the zonal boundaries reflect transmission congestion. Therefore, markets that operate a zonal design will have to constantly monitor the definition of these boundaries. Such monitoring has occurred in the Italian and Norwegian markets. This element could be an important consideration for a GB zonal market.

Decarbonisation

Like much of Europe, since the 2008 Climate and Energy Package (which set out the EU's target of 20% renewable energy by 2020), Italy has experienced significant renewable growth. However, evidence from the Italian market suggests this resulted in increased intra-zonal congestion. Sapio (2015) investigated the impact of renewables on congestion in Italy, finding that increased renewable production increased intra-zonal congestion[19]. Wind energy had the largest impact. With continually accelerating decarbonisation targets this should be a major consideration for whether to implement a zonal market. As more and more renewables come online, congestion on the network changes and increases, suggesting the ever increasing need to change zones and increase constraint costs.

Security of supply

European network operators must consistently re-distribute the network and the scale of this intervention has been systematically increasing over recent years[20]. However, there is a lack of evidence in the literature on how a zonal design specifically affects security of supply in the Italian and Nordic markets.

However, the Italy case study highlights the importance of reviewing periodically the remit of the zones. New zonal configurations make it possible to maximise the efficiency of the market. According to Terna, it reflects more "accurately the criticalities of the grid, enabling thus operators to optimise negotiations avoiding grid security problems, that is for example overloads, voltage collapses, instability".

In recent history, there has only been one major blackout in the Nordic markets. This was in 2003 between Denmark and Sweden, where the Southern part of Sweden and Eastern part of Denmark were blacked out. This was due to a failure in a nuclear plant which started due to a fault in an internal valve.[21] This was then followed five minutes later by a very severe 'double busbar' fault. The impact of these two events together caused large amounts of disruption to the power system. These two events happening five minutes apart was very low probabilistically and caused compounding disruption to the network, which is why the blackout was so severe. It is widely accepted this blackout was not due to the zonal market design.

---

[19] The effects of renewables in space and time_ A regime switching model of the Italian power price | Elsevier Enhanced Readers
[20] P. Borowski (2020): Zonal and Nodal Models of Energy Market in European Union
[21] The black-out in southern Sweden and eastern Denmark, September 23, 2003. | IEEE Conference Publication | IEEE Xplore

Investment in low carbon capacity

There is no evidence to suggest that zonal markets harm investment in low carbon generation compared to a uniform price market or indeed a nodal market. Norway, Sweden, Italy, and Denmark have all had significant growth in renewables. However, there is evidence that this growth in renewable generation increases constraint costs in a zonal market when there is insufficient transmission capacity (as noted by the Australian Energy Market Commission (AEMC)). This is partly owing to large wind and solar farms locating away from the large demand centres.

Zonal pricing also requires accurate zonal boundaries, which often need to be redrawn. The need to change zone boundaries can create uncertainty for investors.

Whole system costs

Zonal market designs tend to ignore the limitations of transmission network and its complexity. This can significantly affect the cost of providing energy to the end user[22]. Borowski (2020) highlighted the systemically growing need and scale of inventions by network operators to redispatch the system. These redispatch activities have resulted in huge costs for Europe's system operators. In particular, cross-border re-dispatching is contributing to large increases in energy prices.

Empirical research from both NEM and the Italian market found that strategic behaviour from generators was present in the markets and resulted in increased system costs. Katzen & Leslie (2020) found this resulted in $2.8 billion worth of misprices generator revenues in the NEM and Graf et al. (2020a) found it significantly increased the cost of power generation in the Italian market. This suggests zonal markets may require controls or governing bodies to monitor and mitigate the impact of generator behaviour.

Through the experiences of international markets, evidence suggests there are some elements of the zonal design that result in worsened outcomes for consumers. For example, Graf et al. (2020a) found that the zonal market incentivised generators to strategically offer in tight situations. This resulted in increased energy prices and these effects were passed on to consumers.

Evidence from the NEM also highlights a similar impact. The consultancy firm Oakley Greenwood found that the strategic bidding stemmed from intra-zonal congestion not being priced. This is a characteristic of a zonal market and contributes to counter-price flows which distort settlement residues mechanisms and result in inefficient market outcomes and negatively impacts consumers.

Evidence from Italy also found that the zonal design incentivising gaming of the redispatch process is driving increased balancing costs. Evidence from the Italian markets shows the risk

[22] Zonal and Nodal Models of Energy Market in European Union. [Energies | Free Full-Text | Zonal and Nodal Models of Energy Market in European Union (mdpi.com)] Borowski, 2020

of participants gaming between the wholesale and balancing services markets to resolve the network constraints, which can be increased by abuse of market power. Similar concerns have driven both the PJM and CAISO markets to transition from zonal to nodal pricing in the early 2000s[23].

Furthermore, Borowski (2020) highlighted zonal markets inability to consider physical phenomena and the effect on costs of electricity production. He further concluded that this gap between the market and the physical capabilities of the system will be worsened as the energy sector continues to transition towards decarbonisation. "The zonal model of the market averages spatially differentiated price signals within large price areas, without showing energy values in a given location. Consequently, system users, and in particular investors, do not know the true value of energy in a given location (node) but only the average value of energy in the market area, which is not conducive to the optimal location of new sources".[24]

Market liquidity

Whilst it appears that there is a gap in the literature on the discussion of an explicit impact of a zonal design on liquidity, Nordic countries' zonal markets (except Finland) and Italy- have respectively the fourth and fifth highest liquidity levels in Europe[25] (just above 2 on average between 2016-2020). Similarly, the churn rate in Australia NEM has been mostly above 2 since November 2010[26]. These findings suggest that a zonal market design does not appear to hinder liquidity.

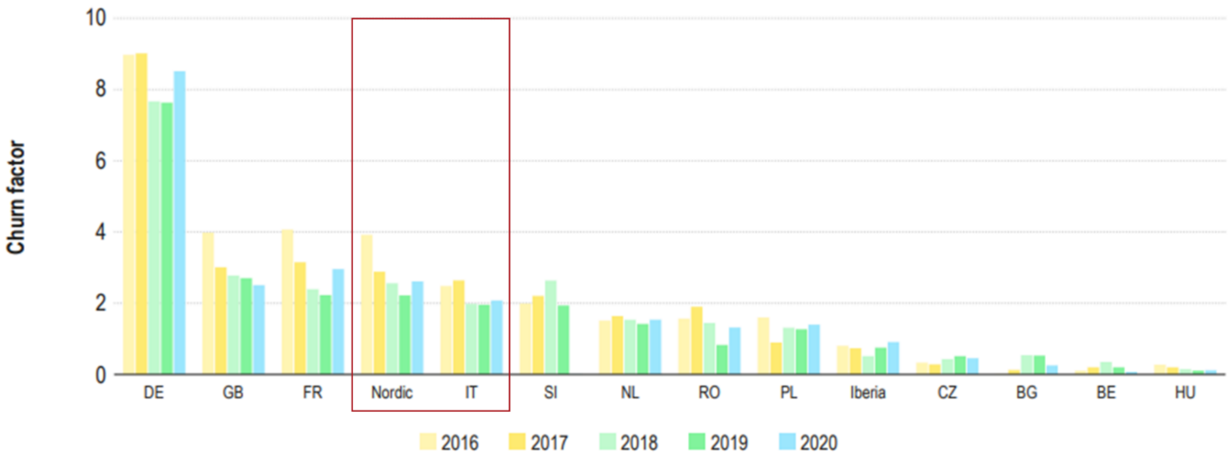

Figure 16: Churn factors in major European forwards markets – 2016–2020 Source: ACER Market Monitoring Report 2020

---

23 benefitsofnodaldesignercot.pdf (amazonaws.com)
24 Zonal and Nodal Models of Energy Market in European Union. Energies | Free Full-Text | Zonal and Nodal Models of Energy Market in European Union (mdpi.com) Borowski, 2020
25 ACER Market Monitoring Report 2020 – Electricity Wholesale Market Volume
26 AER State of the Energy Market, 2022

Zonal market design: applicability to GB

Rationale for zonal design

The markets that have moved to zonal pricing have largely been motivated by a desire to better manage transmission constraints. This is a similar issue to that facing GB and why a zonal market design is being considered as part of REMA.

It is important to note that several markets that have tried zonal design, notably ERCOT and CAISO in the US, ultimately moved to a nodal design. In Australia, the Government recently decided against moving to a nodal design and maintaining the existing zonal design. This was despite modelling suggesting significant benefits to consumers from moving to a nodal design[27].

A zonal market is undoubtably simpler to implement than a nodal market and this simplicity (implementation of about 18 months, compared to a minimum of 5 years for nodal) could prove part of the rationale for implementing a zonal market.

Investment in and deployment of low carbon generation

There is no evidence to suggest that zonal markets harm low carbon generation compared to a uniform price market or indeed a nodal market. Norway, Sweden, Italy, and Denmark have all had significant growth in renewables as well as Australia. However, there is evidence that this growth exacerbates constraint costs in a zonal market (as evidenced by the rationale supporting the recent re-design of the Italian bidding zones). Given the ambitions for significant growth in offshore wind, and the forecasts for constraints in GB, policy makers should be mindful of the potential impacts of implementing a zonal market given GB's constrained network.

The evidence from the case study markets, in particular the Nordic markets, suggest there may be a need to make continual adjustments to zones. Whilst we have not found any evidence of this, we suspect that continual adjusting of zones could negatively affect investment. This is because with changing zones investor certainty of the overall zonal area is reduced. This suggests it may impede an investor's ability to assess which zone to locate in. It also implies continually changing transmission charges, which would further increase an investor's ability to predict costs of potential assets. However, we would still expect renewables development to continue apace in GB given the policy mechanism in place (CfDs). A zonal market would though likely require the need for a top up price for CfDs, or a locational CfDs. This is because areas of high wind output, such as Scotland, would have a much lower price than the high demand areas with low generation, such as the south of England.

[27] Instead, the creation of a new market was proposed- the Congestion Relief Market (CRM). The CRM allows parties affected by congestion to trade voluntarily with each other, to make the best possible use of the available grid.

System impacts

Modelling by Aurora[28] for the Policy Exchange, albeit on a simple stylised three-region model, suggests that a zonal market can deliver up £50bn for consumers up to 2050. Preliminary modelling results from FTI for Ofgem also suggest a significant benefit to GB consumers for moving to a zonal market – at £0.2 - 2.2 bn per year up to 2050. These suggest significant consumer savings, albeit less than the estimate for a nodal market of £3.9 - 9.6bn per year consumer savings.

It is highly likely that a zonal market would produce greater incentives for generation to locate nearer demand, and also for demand (e.g., commercial/industrial) to locate closer to generation compared to a uniform price market. This is likely to provide efficiency benefits for the system. The zonal prices are also likely to resolve some of the issues with constraints, but not all as there are likely to be within zone constraints. Therefore, it is likely that zonal markets would bring consumer and efficiency benefits. However, the experience from the case study markets (Australia, Italy, Nordic, CAISO and ERCOT) suggest that zonal markets still leave a concern with constraint costs which may lead a perennial need to change zones or consider a fully nodal market. This may lead to a constant need to consider reform.

The experience of other markets having tried a zonal design can lead to concerns with zonal constraints being manipulated by generators with market power. Given that GB has had long standing concerns with temporal locational market power, it may be necessary to consider additional market power mitigations for a GB zonal market design. The Transmission Constraint Licence Condition (TCLC), the mechanism in GB design to mitigate location power concerns originally had two ‘circumstances’ (behaviours to be prohibited). The first ‘circumstance’ specifically intended to prohibit portfolio generators manipulating constraints by increasing generation in one area, knowing they would have to be brought on in another. However, this ‘circumstance’ was removed when REMIT came into force, in belief that it was covered and superseded by REMIT. However, in a zonal market it may be necessary to send a clear and explicit signal that manipulation of zonal transmission constraints is clearly prohibited.

Conclusion

Modelling work in GB suggests the potential for a zonal market to produce both efficiency and consumer benefits. A zonal market is also a much simpler design change that a nodal market and could be implemented in significantly less time. However, the evidence from other markets indicates significant challenges still remain with constraint costs and market power concerns and the need to continually reassess zone sizes and configuration. This suggests that a zonal market may not be the optimal market design in the long run and the potential drawbacks need a thorough investigation.

[28] Impact of locational pricing in Great Britain (policyexchange.org.uk)

Capacity Adequacy Market Designs

This chapter explores the international experiences of the capacity adequacy market designs with the aim of addressing the research questions set out previously.

Centralised Reliability Option

Cross-Market Findings on CRO Market Design

Centralised reliability options promise two advantages over other types of capacity mechanisms. Firstly, it ensures the availability of capacity contracted via the capacity mechanism during scarcity. Secondly, the reliability option mechanism limits any energy market distortion[29] due to its implementation and provides the consumer a hedge from high prices. Academic studies have assessed the ability of reliability options in delivering the two promises by analysing the reliability option designs in Italy and Ireland. The conclusions were that that they deliver on the first promise but only partly on the second.

Security of supply

For both Italy and Ireland, it is too early to conclude whether their reliability options will deliver security of supply in a net zero electricity system. The CROs appear to be dominated by fossil fuel technologies, and it is not clear how they would deliver the new low carbon dispatchable power required to in the next zero scenarios for 2030, and 2035.

Similarly, is not clear that if implemented in GB that a CRO is a bankable enough policy to bring forward the low carbon dispatchable generation required to meet supply and keep costs down.

Capacity mix impacts

There is little evidence on the changes to capacity mix brought about due to CRO implementation. In the case of Italy, this is because the first auction was held in 2019 with delivery not till 2022. Therefore, there is a gap within the literature as the effects have not yet been studied. Lazarczyk & Ryan (2019) studied Ireland’s capacity auctions, finding that during

[29] Capacity mechanism designs such as capacity payments and capacity obligation require defining an artificial product (i.e., capacity) which can distort energy-only market. As Oren, (2005) explains: "Generators receiving capacity payments can be more aggressive in pricing the energy they produce. This in turn may suppress energy prices, making it impossible for generators to recover their capacity costs from inframarginal profits on energy, thus perpetuating the need for the capacity revenues.” In principle, reliability options do not have these distortions, because it is a risk-sharing mechanism. The buyer is insured against high prices as well as the availability of generation. In return, the seller of this option earns a reliability premium based on its opportunity cost that should factor in any loss of revenue during the period of energy market prices above the strike price. The energy market itself continues to function ‘business as usual’, and any capacity that is uncontracted in the reliability option auction but clears the market would receive the market-clearing price.

the first RO auction most capacity was auctioned to gas and steam generators. They also found that almost all the new capacity was auctioned to demand side units. However, the paper fails to comment on the long-term impacts of the Irish RO. Analysis of both the Italian and Irish auctions results clearly shows the RO market design has mostly forested the development of fossil fuel generation.

Capacity adequacy costs

Lazarczyk & Ryan (2019) found that the first two Irish RO auctions successfully auctioned €333 million and €345 million capacity, respectively. These were both significantly less than the 2016 annual capacity payment, which was €515 million. This evidence suggests that the introduction of Ireland's CRO lessened whole system costs. While there is a lack of evidence on the effects to the Italian system, the theory underlying an CRM suggest that it will follow a similar trend.

Arup believe that a reliability option in the UK is likely to be cheaper than the existing CM if used in the very short run. This is because it would not result in paying all capacity in the market, as the CM does. However, it is unclear as to the magnitude of this costs saving especially as we get closer to net zero and the number of available dispatchable power plants decrease. If the RO is unable to incentivise new plant onto the system, which we doubt it would, then there is not enough dispatchable power, then the option RO auctions could have limited liquidity and suffer from market power concerns. Arup suspect that the RO would suffer similar issues to the existing CM in that ultimately neither policy is really intended to bring forward new technologies, but rather cover the missing money for existing technologies.

Applicability to GB

There is limited evidence from international experience on which to judge the applicability to GB. However, the experience of the Irish market suggests a mixed applicability to GB. It is likely that, at least in the short run, the option could be cheaper than our existing CM. However, the early Irish experience suggests the policy is not well suited to bringing forward new low carbon technologies. These will be crucial for GB in meeting its security of supply objectives in a net zero market. Additionally, the need to override the auction outcome to ensure there is generation located in the rights areas suggests it is not well suited to markets with significant transmission constraints, as we have in GB. A longer period is needed to properly assess the efficiency of CROs in Ireland and Italy. However, security of supply concerns are getting more pressing, as indicated by decreasing liquidity in the CM auctions, and there is unlikely to be the time to fully evaluate the CRO. The early evidence suggests that the CRO is likely to suffer similar issues to the existing CM in GB but may reduce the cost because it is not market wide.

Strategic Reserve

In Europe, 14 markets have implemented a Strategic Reserve market design. Arup selected the following markets which have currently a Strategic Reserve design in place: Sweden, Germany, and Australia (NEM). Arup also analysed the reasons behind Poland and Belgium having replaced their strategic reserves with new capacity market mechanisms. The model involves auctioning a set amount of reserve capacity above what the market is expected to supply. Providers are remunerated for availability whilst they also receive an activation payment. Capacity in strategic reserves is dispatched only when there is a risk of negative margins (i.e., demand is above supply). The following chapter explore the cross-market findings on the Strategic Reserve market design and its potential applicability to the GB power market.

Cross-Market Findings on the Strategic Reserve market design

Market design rationale

Ensuring that supply is sufficient is critical for energy systems, especially under accelerating decarbonisation targets which have led to an increase in penetration of renewable energy generation. Strategic reserves are one option used to secure adequate supply of dispatch resources. These have been adopted in many energy markets worldwide including Germany, Sweden, Finland, Australia and Belgium. Strategic reserves are typically implemented with a view that that are a temporary measure to maintain the security of supply during a key transition period for a power market.

Both the Swedish market and the NEM introduced strategic reserves in the late 1990s. The NEM introduced their mechanism to provide a ‘safety net’ for capacity in their system. This was originally introduced as a temporary solution; however, the mechanism was eventually made permanent in 2016. The Swedish market originally introduced the strategic reserve to address the increased risk of power shortages as a result of the closing down of the Barsebäck nuclear plant’s second reactor.

Germany, Belgium, and Poland introduced strategic reserves in 2019, 2014, and 2014 respectively. Germany’s motivation was underpinned by their decision to phase out nuclear generation, whereas Poland’s strategic reserve was implemented to help the security of their supply during the country’s decommissioning of their 3.4GW of coal capacity. Both mechanisms were introduced with the underlying knowledge that a longer-term solution will be needed as a in the future after this. Belgium’s strategic reserve was introduced to ensure that there was adequate supply during the colder months following concerns of shortages during these tighter periods following the decision to phase out nuclear by 2025.

However, Belgium and Poland are both currently considering moves away from their strategic reserves. Belgium is currently undergoing the process of replacing its reserve with a reliability option, under the rationale that this will provide adequate security of supply and will not distort any market signals or competition. Poland decided that moving away from a strategic reserve and introducing a capacity market would be best in their market. This is due to increasing concerns that the market design that is currently implemented cannot deliver sufficient security of supply. A study by Compass Lexecon concluded that the current strategic reserve was inadequate and not the most cost-effective capacity measure available. The analysis found that a capacity market would result in more than €350 million in consumer savings per year compared to the strategic reserve.

The experience of these markets, especially Belgium and Poland, question the long-term applicability of a strategic reserve.

Impacts on investment in and deployment of low carbon technologies

For most of the case studies analysed, the strategic reserves were introduced to help ensure adequate capacity during the transition to a decarbonised economy. Despite this, there is little analysis available on the causal effect of a strategic reserve on the decarbonisation and the deployment of low carbon technologies.

Both Germany and Sweden have had to impose additional environmental standards for their respective strategic reserves. Germany has an Energy Performance Standard (EPS) that is required for its strategic reserve, which caps carbon emissions within the capacity reserve. Sweden has a similar environmental requirement on its reserves.

Security of supply

Strategic reserves are a measure used to address the issues of ensuring adequate supply of dispatch resources and therefore increasing confidence on security of supply. However, research by Billimoria et al. (2018) suggested that strategic reserves may not be sufficient in securing the required amount of resources, especially in a system with a high penetration of renewables. With the accelerating global targets of net zero, these findings suggest that strategic reserves will not be an efficient capacity adequacy measures in the transitioning energy markets.

Evidence from the Australian market shows that the mechanism can be successful to a certain extent. Prior to the summer of 2017/18, the RERT had never been dispatched and contracts had only been entered three times. During the summer the RERT was required to cover significant amounts of supply, for instance over a single six-hour period a total of 390MWh was required to be dispatched. While this illustrates the mechanisms' ability to procure adequate dispatch, wider research suggests there are more efficient mechanisms that can also do so.

Whole system costs

Evidence from the international markets has shown that a strategic reserve can result in significant costs. For example, a 2017 study carried out by FTI Compass Lexecon found that the “German SR could increase net costs for consumers by € 800-1,300 million per year” between 2020 and 2040. The research also found that under high penetration of intermittent renewables the mechanism resulted in cash flow volatility. Similar findings can be observed in

Australia. During the summer of 2017/18, the RERT (Australia's strategic reserve) cost the market approximately $51 million (AUD), with a large amount of this being passed onto consumers through indirect costs despite the mechanism being designed to minimise these costs. Further research also suggests that strategic reserve may not be the most cost-efficient mechanism to ensure security of supply. Research by Compass Lexecon on the Polish strategic reserve found that by switching from strategic research to a capacity market would result in €8 billion savings for consumers over the 2017-2040 period.

Cost of capital and investment impacts

In theory the introduction of a strategic reserve could help incentivise investments into capacity provision and for more players to enter the market provided that that there is sufficient demand i.e., the reserve is used often. However, this has not been the case for the markets we have studied. For instance, stringent rules on plants being unable to return to the market after participating in the reserve have minimised incentives for investment. The Research Institute of Industrial Economics suggest that this is overly restrictive and is resulting in an inefficient policy. In contrast to Germany, generators in Sweden operate in the strategic reserve during winter and then can return to competition in the energy-only market during the remaining months of the year. However, this has been highlighted as an issue as these generators can exhibit market power and distort prices during the summer months, deterring entrants and consequently investment.

Sweden also found that stringent regulation impacts market investment, however, the lack of investment stemmed from the strict environmental policies. The environmental regulation on strategic reserve plants is stricter than those that participate in the market, so have deterred entrants due to their inability to meet the standards and therefore worsened competition within the reserve.

There is potential for strategic reserves to promote investment, however this is significantly dependent on the design on the mechanism.

Applicability to GB

The evidence from international experience presents a mixed picture on strategic reserves. In Europe, there is a growing tendency to move away from the mechanism in favour of capacity markets. Some policy appraisals suggest a significant cost benefit in this. However, in the countries where they have been used, strategic reserves have been a cheaper year-on-year option than capacity markets. In addition, there does not appear to be any evidence of the ‘slippery slope’ that has previously concerned GB policy makers. The academic literature suggests that strategic reserves may not be sufficient in securing the required amount of resources, especially in a system with a high penetration of renewables (Billimoria et al., 2018).

Given GB’s experience of the Supplemental Balancing Reserve (SBR), which was effectively a strategic reserve, it is clear a reserve could be implemented quickly in GB. However, it is not clear whether this would meet all of GB’s policy objectives. The evidence from other markets suggest that strategic reserves do not provide investment incentives, especially for low carbon generation. However, they may offer the ability to keep existing plants operational to cover year-ahead capacity risks. Further work is required to fully assess the suitability for GB, but strategic reserve could offer targeted support for existing plants in times of heightened risk.

Low carbon investment and flexibility

Unprecedented large-scale investment in renewables is required to meet net zero targets. Simultaneously, large scale investment in flexible technologies is needed to enable this increased uptake of intermittent generation. Currently, renewable developers can access wholesale revenues. This alone is not sufficient to meet the required investment in renewables (hence the need for CfDs to provide the necessary price certainty for investment). Similarly, today flexible technologies aren't appropriately exposed to operational market signals, nor have opportunity to procure long term forecastable contracts thereby resulting in higher financing. To achieve mass deployment of both technologies, and to address other market failures, market interventions are required.

One possible intervention is supplier obligations. At its most basic definition, a supplier obligation is a decentralised, market led approach whereby government imposes an obligation or target on suppliers to achieve.

As it relates to renewables, and as set out in the REMA consultation, government has considered imposing an obligation on suppliers by setting a maximum carbon intensity of electricity, which aligns with carbon budget six and net zero, that suppliers can sell to consumers. To meet the carbon intensity threshold, suppliers can contract with a renewable generator directly or via an intermediary. Through the REMA consultation, government also proposes to set an obligation on suppliers to drive the uptake of flexible technologies, including demand side flexibility.

There are benefits and drawbacks to this market intervention. First, the supplier obligations proposed by government allow for competition between a range of technology types as the targets are relatively technology agnostic. This commercial pressure may drive technological and business model innovations resulting in least cost decarbonisation. Lastly, this is a market led approach so reducing the risk that government makes ineffective decisions about the future electricity mix. The key drawbacks to this method relate to financing and deliverability. In recent years, the supplier market has been relatively volatile resulting in several suppliers having to exit the market. Generators and flexibility technology developers will be exposed to this volatility as they will have to either contract directly with a supplier or via an intermediary, resulting in higher financing costs for projects. Second, this approach places a large amount of responsibility on suppliers to manage the delivery of significant investment required to meet net zero targets and they may not be best placed to do this.

The chapter below explores the cross-market findings and the applicability of the Supplier Obligation market designs to the GB power system.

Cross-Market Findings on the Supplier Obligation Market Design

Market design rationale

In recent history decarbonisation and net-zero has been an ever-accelerating target for countries across the globe. With the majority aiming for net-zero around 2045, Governments have been implementing policies to help achieve these goals. A sector particularly targeted by such policies is the energy sector. In the EU, the sector accounts for more 75% of the Unions GHG emissions, and similar patterns can be observed in the US[30].

A particular approach that certain countries and markets have opted for are supplier obligations. This is an obligation on electricity suppliers to supply a required amount of their total electricity from renewable sources[31]. Supplier obligations have been implemented by the US market operators and in some EU countries.

In the US, the supplier obligation has been introduced through Renewable Portfolio Standards (RPS), which set a minimum requirement for the share of electricity supply that comes from designated renewable energy resources. The policies are focused on electricity generators and vary state to state. They have been introduced by states under the notion they are an efficient, cost-effective, and market-based approach to achieving both renewable electricity targets and wider net-zero goals[32].

In Europe, some countries have also adopted a similar approach. Under EU legislation, member states must commit to reducing their GHG emissions and, as an extension of this, must achieve their renewable target goals. This has significantly encouraged member countries to implement a variety of support mechanisms to help encourage renewable energy development. Italy, Belgium and Poland have all implemented their own version of a supplier obligation to help adhere to EU regulation and achieve their renewable energy targets.

Decarbonisation

The introduction of supplier obligation schemes has led to positive effects regarding the decarbonisation effort in both the US and in Europe. In Poland, significant increases in renewable generation were observed following the introduction of the quota obligation scheme, especially in wind capacity. By 2010, more than 12GW of wind permit was permitted to be connected to the grid and 65GW had applied for connection. Italy also experienced an increase in renewable capacity, observing an immediate increase within the first year following the policy implementation. However, unlike Poland, analysis of the impacts of the Italian mechanism suggested that the effect was not technology specific and increases in capacity

---

[30] European Commission: Renewable energy targets. https://energy.ec.europa.eu/topics/renewable-energy/renewable-energy-directive-targets-and-rules/renewable-energy-targets_en accessed 26th January 2023
[31] BEIS (2022): Review of Electricity Market Arrangements. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1098100/review-electricity-market-arrangements.pdf accessed 26th January 2023
[32] EPA (2022): Energy and Environment Guide to Action - Chapter 5: Renewable Portfolio Standards. https://www.epa.gov/statelocalenergy/energy-and-environment-guide-action-chapter-5-renewable-portfolio-standards accessed 26th January 2023.

were observes across all technologies. The RPS in the US market has also resulted significant increases in renewable capacity in the systems. ERCOT has observed significant increases in wind capacity and similar trends can be observed for solar generation within the CAISO market. While supplier obligations will have assisted in the increases in renewable capacity across the specified markets, Arup cannot conclude that this was solely due to the SO mechanisms. Decarbonising the grid has been focused target globally and, as a result, countries have tended to implement multiple polices to help encourage increased renewable generation. Therefore, the exact effect of the supplier obligations cannot be isolated and some of the evidence suggests that compliance costs are increasing the further systems decarbonise.

Security of supply

The experiences of the US and European markets suggests that supplier obligation policies do not significantly impacts the security of supply. For example, there is no mention of events involving a risk to security of supply in Italy or Poland following the implementation of their supplier obligations. The CAISO market illustrates those mechanisms, such as the EIM, can be successfully used to help with the integration of increased renewables, therefore helping ensure security of supply.

Capacity mix

Evidence from the Polish market suggest that RES support schemes, including the quota mechanism resulted in the development of the RES sector. Between 2005 and 2014, RES installed capacity increased 6-fold and the market structure began to evolve. While previously hydro power was the dominant RES technology, the introduction of the scheme promoted the growth of technologies such as wind[33].

Whole system costs

There is a lack of evidence in the literature regarding the impacts of supplier obligation in the European markets. However, the evidence for the US markets suggests that the implementation of the RPS scheme did not have a significant impact on system costs. Research found that RPS compliance costs constituted less than 2% of average retail rates in most U.S. states over the 2010–2013 period. Although, the research also concluded that there is substantial variation across years and states suggesting the exact impact is unknown.

Consumer impacts

There is a debate within economic literature on the impact consumers will face following the implementation of a supplier obligation mechanism. Some research suggests that introducing RPS policies could lead to a 3% increase in electricity prices, and this was the underlying line of thought for the many US states that did not want to implement and RPS. As the case with any quota, the enforced renewable share could result in increased operating costs for the

[33] Wedzik et al. (2017): Green certificates market in Poland – The source of crisis

generators which is potentially passed onto the consumer through the pricing mechanism. While a supplier obligation may increase investment in renewable technologies, it will result in inefficient market outcomes and negatively impact consumers due to increased retail costs.

Implementation

The implementation of supplier obligations in Europe and the US is not widely covered within the literature, however, there appears to have been little to no issues with implementation.

The Italian market has been praised for its use of a pilot study before full implementation and this was one of the reasons attributed towards its success. The mechanism also built-in transitions periods, allowing stakeholders up to two years to adjust their processes.

It can also be inferred that the Polish quota system occurred smoothly due to a lack of mentions of any issues within the literature. However, the Polish mechanism did experience issues in accurately forecasting capacity increases which led to inefficiency quota levels being set. In 2008 they decided the sustain a fixed quota of 10.4% however under forecasted the increase in capacity. This resulted in a structural oversupply over the subsequent years, in 2012 there was an oversupply of 3.4 TWh[34].

Applicability to GB

The experience of US and European markets both suggest that supplier obligations have been an effective policy in bringing forward the deployment of renewables technologies. However, this does not mean they are the policy best suited to the GB market.

GB has used supplier obligations previously and concluded that government backed CfDs are a preferred method for delivering investment in renewable technologies. Under the renewables obligation, GB policy was like that of the US markets. However, the policy was found to be more expensive than CfDs and several suppliers were unable to properly manage their obligation costs.

Additionally, there were concerns that investors found the obligation difficult as a bankable investment tool due to the uncertainty of its value. Given the previous experience with obligations, we would question why GB would want to move away from CfDs as a bankable investment policy for low carbon technologies, albeit that market maturity suggest developers can now manage more merchant risk. However, the experience from other countries does suggest supplier obligations can work. If used in a GB context, given GB's previous experience, perhaps using them to target flexibility technologies that don't require large financial investment could be a useful approach; for example, a supplier obligation to have a certain amount of DSR.

[34] G. Skarżyński (2016): The Green Certificates market in Poland – origin, evolution, outlook. http://psew.pl/en/wp-content/uploads/sites/2/2017/03/232_psew_eng.pdf accessed 26th January 2023

Operability: Co-optimisation of Ancillary Services

Co-optimisation of Ancillary Services

Historically ancillary services have been cheap to procure, and many of the services were a complimentary by-product of the type of generation; for example, if gas fired turbine is use generate electricity it will continue to spin and produces electricity even after an event, this inertia gives the system operator crucial seconds to get fast response reserves to kick in. However, as the generation mix has changed many of these services are no longer a complimentary by product and the system operator is having to purchase different types of services at increasing costs. In addition, the system operator is also having to assess the outcomes of the GB's self-dispatch market and make sure there are enough ancillary services to cover what it needs to keep the system within its frequency tolerances. However, in the US markets which have moved to centralised dispatch, the approach is to simultaneously solve the right amount of generation dispatch with the right amount of ancillary services. This is described as co-optimisation, and theoretically it should provide a more efficient solution than trying to solve the two separately. The following section explores the case studies and evidence from the literature on the co-optimised US markets (CAISO, ERCOT, MISO and ISO-NE).

Market design rationale

The US markets adopted co-optimisation of electricity and reserve, or ancillary markets, at varying times. As with the other market design changes in the US, co-optimisation was generally part of broader set of market design changes to better get market prices to reflect system needs. However, co-optimisation has increasingly been motivated by a desire to better manage intermittent renewables generation. We briefly consider the date and motivations for optimisation.

- ISO-NE's implementation of the Pay-For-Performance (PFP) market design was a major overhaul of its previous market design, which had been in place since the early 2000s. The PFP design was developed over several years in collaboration with stakeholders, including market participants, regulators, and consumer advocates, and it was subject to extensive testing and evaluation before it was implemented.
- ISO-NE implemented its co-optimized electricity and reserves market, called the PFP market design, on June 1, 2018. The PFP market design aims to includes a new two-settlement process for the energy and reserves markets, which aims for prices to better reflect the needs of the grid in real time. Part of the motivation for this change was better help the integration of renewable energy.

- ERCOT implemented co-optimisation of energy and ancillary services in its market in March 2011. The co-optimisation was part of a larger set of changes to the ERCOT market design as discussed earlier. Co-optimisation was seen as a way to facilitate the integration of renewable generation into the grid. By co-optimising energy and ancillary services, ERCOT hoped to manage the variability and uncertainty of renewable resources and maintain grid reliability and stability more effectively.
- MISO (Midcontinent Independent System Operator) implemented co-optimisation of its electricity and ancillary services markets in, 2017. The co-optimisation was part of a larger set of changes to the MISO market design, which aimed to improve grid reliability and reduce costs for consumers by better aligning market signals with the actual needs of the grid.
- CAISO (California Independent System Operator) implemented co-optimisation of its electricity and ancillary service markets on November 1, 2014. The co-optimisation was part of a larger set of changes to the CAISO market design, which had been encouraged.

Implementation

Several studies in the academic literature have investigated the challenges of designing and implementing co-optimised markets. They tend to highlight the complexity of co-optimisation and the need to including factors such as market structure, pricing mechanisms, and operational constraints. For example, a paper by Chassin et al. (2014) explores the challenges of integrating energy and ancillary services markets in the Western Interconnection, which includes CAISO and ERCOT[35]. The authors discuss the need to balance the different objectives of energy and ancillary services markets, and the importance of market design and governance in achieving this balance.

O'Neill et al. (2016) examine the challenges of co-optimising energy and ancillary services markets in the ISO-NE and MISO electricity markets, focusing on the difficulties of integrating renewable energy resources and demand response programs. The authors suggest that coordination between the different markets and system operators is essential for achieving effective co-optimisation.

Khan, Arnold, Baldick, and Xie (2018) explore the challenges of implementing co-optimised electricity markets. The key challenges presented by these authors include the need for sophisticated optimisation algorithms, changes to market rules and operational protocols, and the potential for market power abuse.

It should also be noted that that the co-optimisation of ancillary services does not include all such services we would describe in GB. Black start capabilities, reactive supply and voltage control, inertial response, ramp capability are all services normally labelled as ancillary. However, 'most U.S. power systems do not operate markets for these products but do maintain

[35] Chassin, D. P., Widergren, S. E., & Pratt, R. G. (2014). Co-optimization of energy and ancillary services in the Western Interconnection. IEEE Transactions on Power Systems, 29(1), 163-172

adequate resources through internal requirements and other procurement mechanisms'[36]. Therefore, referring to co-optimisation of energy and ancillary service in US markets, we principally mean reserves.

	Spinning reserves	Non-spinning reserves	Regulation
CAISO	Spinning (10)	Non-spinning (10)	Regulation-up Regulation-down
ERCOT	Responsive ("a few")	Non-spinning (30)	Regulation-up Regulation-down
ISO-NE	Synchronised (10)	Non-synchronised (10) Operating (30)	Regulation (5)
MISO	Spinning (10)	Supplemental (10)	Regulation (5)
NYISO	Spinning (10) Spinning (30)	Non-synchronised (10) Non-synchronised (30)	Regulation (5)

Table 5: Overview of the ancillary services offered by each ISO/RTO, the required response time in minutes for each product is indicated in parentheses when available.

Market impacts

An assessment of the evidence on the impacts of co-optimisation on the US markets supports the view that co-optimised markets can reduce the costs of ancillary services and lead to efficiency gains. There was no evidence suggesting that co-optimised markets create inefficiency or additional costs for consumers. However, it should be noted there is a lack of empirical analysis that quantifies the benefits, and a lot of the evidence is derived from theoretical simulations.

Multiple studies[37] have suggested that the implementation of co-optimisation in US electricity markets can lead to increased efficiency, lower costs, and improved reliability. The inclusion of opportunity costs in the price of ancillary services through co-optimisation can result in efficiency gains and lower procurement costs. The use of co-optimised markets has helped to

[36] Microsoft Word - Argonne_Survey_US_Ancillary_Services_Markets_20170111 (anl.gov)
[37] Including but not limited to Baldick et al. (2005), Jamalzadeh et al. (2008), Baldick et al. (2021), and the California Energy Commission (2019).

reduce the costs of electricity generation, increase the reliability of power supply, and improve the overall efficiency of electricity markets. Co-optimisation has also led to a reduction in uplift costs and improved the utilisation of resources on the grid. The introduction of co-optimisation has incentivised market responses and promoted innovation in market mechanisms. However, there is a lack of empirical analysis quantifying the benefits, and much of the evidence is derived from theoretical simulations.

Some evidence for U.S. markets suggests that co-optimisation may promote innovation of market mechanisms. For example, most U.S. markets implemented administrative penalty factors alongside their co-optimised ancillary services. This helped to incentivise market responses in five-minute dispatch intervals that were quicker and provided stronger pricing signals. The introduction of such mechanism incentivised dispatch instructions that minimised the system-wide costs. Another pricing innovation that emerged from the introduction of co-optimisation was ERCOT's Operating Reserve Demand Curve. This mechanism increased the price of energy and ancillary services as the system approached shortage conditions, and therefore incentivising resources to optimally allocate their reserves. Services were deployed when they are needed most, and it decreased the likelihood of an outage (Brattle Group, 2017).

The view in the US is that co-optimised markets are the optimal design due to the increased efficiency generated by this market design. The Federal Energy Regulatory Commission (FERC) considers co-optimisation can help reduce market inefficiencies and better align market signals with the actual needs of the grid. It now recommends co-optimisation as the model of choice.

Impact on investment in and deployment of law carbon technologies

Theoretically, co-optimised electricity markets should help support the deployment of renewables and storage technologies by enabling system stability. And this was, in part, the rationale for co-optimisation. The evidence tends to support this view.

The Brattle Group, reviewing all the ERCOT market design changes (i.e., nodal and shortened settlement periods), found that co-optimised energy and ancillary services markets in ERCOT has facilitated the integration of renewable resources, resulting in an increase in renewable penetration from less than 2% in 2001 to over 20% in 2018. They suggest that co-optimisation has reduced the need for curtailment of wind and solar resources, resulting in cost savings and increased revenue for renewable generators. In addition, co-optimisation has led to improvements in the accuracy of renewable energy forecasting, enabling ERCOT to better plan for the integration of renewable resources and maintain grid stability.

The analysis from the ISOs in the US suggests there is a strong belief that co-optimisation supports the market for batteries and other storage technologies.

> *"The integration of energy storage resources into MISO's co-optimized energy and ancillary services market has helped to create additional revenue streams for*

these resources and provide valuable services to the grid." - Midcontinent Independent System Operator

"The implementation of co-optimized day-ahead and real-time energy and ancillary service markets in ERCOT has resulted in greater opportunities for battery storage resources to participate and provide valuable services to the grid." Electric Reliability Council of Texas (2019)[38].

"The introduction of a co-optimized energy and capacity market has helped to create new opportunities for energy storage resources to participate in the market and provide valuable services to the grid." - ISO New England (2020)[39].

Security of supply

The evidence from the rapid evidence assessment suggests that co-optimisation is likely to have improved the reliability of the systems that have co-optimised their energy and reserve/ancillary markets. The implementation of co-optimised energy and ancillary services markets has improved the reliability of electricity systems, particularly during periods of high demand and extreme weather events. The evidence from various studies[40] suggests that co-optimization has helped to maintain sufficient levels of operating reserves and reduce the likelihood of blackouts, while also improving the overall efficiency of the grid. Co-optimisation has allowed for greater flexibility in the dispatch of resources and helped to reduce the need for out-of-market actions to maintain system reliability. The impact of co-optimisation has been observed in several electricity systems, including ISO-NE, CAISO, ERCOT, and MISO.

Conclusions

Generally, the move to co-optimisation was part of a wider package of market redesign, that included centrally dispatched nodal markets with greater temporal granularity. However, increasing the desire to better manage the risks of intermittency from renewables generation led us system operators to co-optimise. The rapid evidence assessment review suggests co-optimisation has helped reduce the cost of ancillary services provision and helped improve system reliability through improved forecasting and greater transparency. In addition, there is some evidence to suggest that co-optimisation has helped the deployment of low carbon generation assets. Batteries and storage technologies are indicated to have had greater ability

[38] 2019 State of the Market Report (no longer accessible).
[39] 2020 Regional Electricity Outlook.
[40] One example of this is the ISO-NE market, which implemented a co-optimized energy and reserves market in 2018 to better align market signals with the actual needs of the grid. This change was part of ISO-NE's efforts to improve grid stability and reliability in the face of increasing renewable energy penetration and other grid challenges. While it is difficult to say whether this change specifically prevented a blackout, it is likely that it has contributed to improved grid stability in the region. Another study by the Analysis Group in 2019 found that co-optimization has improved the reliability of the ISO-NE grid, particularly during periods of high demand and extreme weather events. The study found that co-optimization has allowed the ISO-NE market to better capture the value of resources that can provide both energy and reserves, which has improved the grid's ability to maintain reliability during periods of stress.

to generate revenues through co-optimisation. Intermittent wind and solar generators are suggested to have indirectly benefited through reduced curtailment. The benefits of co-optimisation are very difficult to quantify and separate out from other market design features.

The literature indicates that co-optimisation can be complex to implement. It is noticeable that the US markets co-optimise a limited number of reserve products, which is significantly lower than the long list of ancillary services we have in GB. The literature also suggests that there can be risks of market power, but this is only theoretically presented. There is an overwhelming consensus in US policy makers, which is supported by the literature, that co-optimisation is the preferred market design.

Applicability to GB

The current approach to ancillary services provision in GB is markedly different to that of the US market approach of co-optimisation. NGESO procures a significant number of ancillary services, see table below, and in a significantly different way to the day-ahead approach used in the US. Currently, NGESO procures its ancillary services with annual and monthly auctions. The services, such as STOR (Short-Term Operating Reserve) are then dispatched via the balancing mechanism. So, this, like constraints costs, the actual system is re-dispatched for its actual needs, after the market has settled is and dispatched according to generators' financial incentives.

Co-optimisation could only be considered as a move to centrally dispatched market. Arup interviewed NG ESO engineers to discuss the potential for co-optimisation. They highlighted there are huge opportunity costs of co-optimising and although the benefits are significant, the implementation presents significant difficulties. NGESO mentioned they had previously investigated co-optimising reserves, frequency response, black start, and inertia but found that the more elements you include the harder this is to solve the optimisation problem. It also significantly increases the difficulty of building the algorithm and optimiser. Nonetheless the ESO believe that, in a decarbonised system, centralised dispatch with co-optimisation can offer efficiency benefits.

The scale of the benefits is difficult to predict. However, given that ancillary services costs are an order of magnitude lower than constraint costs, and in the hundreds of millions, we would anticipate the annual efficiency gains to be in the tens of millions per year.

In recent years, there has been a growing need for NGESO to procure system services outside of network connected generation. Products and services have been developed, such as the Optional Downward Flexibility Management in 2020 and the recent demand side response products. This is suggesting a need to not only optimise transmission connected services, but with growing opportunities that lie on the distribution system.

The evidence suggests the NGESO should aim to include co-optimisation if they proceed with a centrally dispatched approach. However, they are likely to need to dramatically rationalise the number of ancillary services so reduce the complexity of the optimisation algorithms.

This publication is available from: www.gov.uk/desnz