September 2024

Assessment of likely impacts of proposed reforms to the balancing mechanism within a national-price regime

FTI Consulting | Report for Octopus Energy

FTI CONSULTING

Table of Contents

Executive summary

Several commentators have proposed reforms to the balancing mechanism ("BM") as alternatives to a transition to a more locational wholesale electricity market design in GB. In this note, we describe and assess the main proposals suggested by commentators. We find that the suggested reforms to the BM are unlikely to address the current issues in the GB electricity market and, in many cases, may have negative impacts on consumers and the GB energy system as a whole if implemented within the current system of national wholesale pricing.

1. GB's current electricity market arrangements are generally agreed to be not fit for purpose.[1] The problems of the current system can be seen in many ways, but perhaps most obviously in high and increasing constraint costs. These are the costs that the System Operator (SO) incurs in dealing with bottlenecks in the transmission network. They have risen from £170m per year in 2010 to over £1.3bn in 2022, an eight-fold increase.[2] Ultimately, these costs are borne by consumers on their energy bills.

2. More transmission will have to be built in the future, but investing in the transmission network to the extent that constraint costs would fall significantly would be very costly, politically contentious and slow. A well-recognised solution is to divide the current electricity market into a number of price zones. This can help to ensure that buyers and sellers of electricity trade in a way that is more consistent with the electricity transmission network. With zonal pricing, trading energy across zones is only possible if there is enough capacity on the network. Because transmission constraints are reflected in the wholesale market, the SO would have much less need to take costly action to ensure that electricity supply and demand balance. In the longer run, electricity prices that reflect local conditions should help to attract large consumers (data centres, for example) to site closer to generation, reducing network costs further. The introduction of locational energy pricing in GB is currently under consideration as part of the government's Review of Electricity Market Arrangements ("REMA").[3]

3. Some commentators have proposed alternative reforms, whereby the current national wholesale market stays in place, but the way in which the SO manages constraints changes. In this note, we briefly examine the extent to which the proposed reforms might be effective – or not. We have categorised these proposals as follows:
   - Proposals that would **increase centralised control by the SO ahead of delivery**, with the aim of reducing the costs of managing any given level of system constraints.
   - Proposals that would **give the SO discretion to restrict or alter the scheduling** of pre-specified classes of assets in the wholesale market, without financial compensation, with the aim of reducing the level of constraints.

[1] See the DESNZ publication of the Review of Electricity Market Arrangements ("REMA") summary of responses to consultation (link).

[2] ESO, 2022. Net Zero Market Reform: Phase 3 Assessment and Conclusions, p.13 (link).

[3] DESNZ, 2024. Review of Electricity Market Arrangements: Second Consultation Document, p. 90-91 (link).

- Proposals that would **redesign how constraints are managed by creating a market for balancing the system**, with the aim of increasing competition between different types of assets and so reducing the cost of managing constraints.

4. We find that, while each of the proposals has the potential to produce some benefits, each of them also have negative unintended consequences.

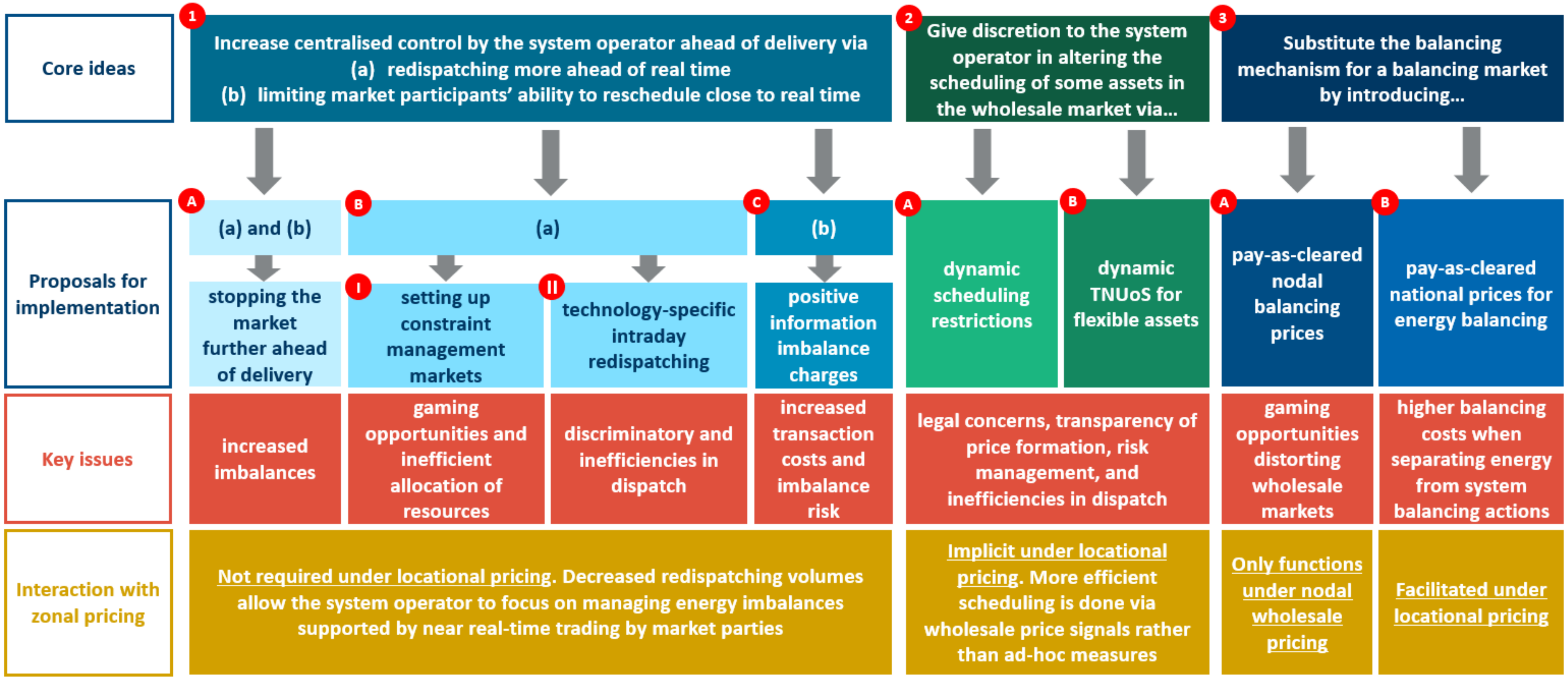
- **Increasing centralised control by the SO ahead of delivery.** For instance, the time window between the end of wholesale market trading and actual dispatch of electricity could be increased – this is currently one hour but could be extended to six hours. More control by the SO could help to reduce constraint costs, as the SO would have better information at an earlier stage about the likely constraints on the system and could take more proactive action to resolve constraints. But it would do so only at the expense of creating significant other problems. In particular, these proposals would reduce the opportunity for market participants to use well-functioning intraday markets to balance their own positions, which they are arguably better at than a central entity.
- **Giving the SO discretion in restricting or altering the scheduling of pre-specified classes of assets in the wholesale market.** These proposals could have more profound effects, since they would reduce the volume of constraints, not just the costs of managing them. For instance, the SO could be given more powers to prevent interconnectors to other countries from operating when doing so would increase transmission constraints. However, these proposals could also increase inefficiencies because the SO would be treating some classes of assets differently from others. For instance, more expensive assets might be selected to generate instead of cheaper ones, purely because of the different rules around how they are scheduled. By increasing the role of the SO in deciding what assets are scheduled rather than leaving such decisions to the market, these proposals could also reduce transparency of pricing and make risk management harder.
- **Redesigning constraint management, such as by developing a market for balancing the system.** For instance, if the SO decides to pay a wind farm not to generate, the price it receives could be based on the clearing price in a market for constraints (so-called ‘pay-as-cleared’) rather than, as at present, what it bids (‘pay-as-bid’). This could in principle make it easier for different types of assets to participate in the balancing market, thereby increasing competition and reducing system costs. However, we find that these proposals would only work under a locational wholesale market. If wholesale prices remain nationally-determined, changes of this type would create increased gaming opportunities, with market participants having an incentive to, for instance, withdraw from wholesale markets in order to increase their profits in a balancing market. As a result, these proposals would likely make constraint costs rise rather than fall.

5. Our assessment focuses on the short-term expected impacts of the proposals. The longer-term effects would depend on the extent to which they would create incentives for market participants to locate in areas where their actions would reduce rather than increase system costs. For instance, proposals that reduce payments to generators that are constrained off would be expected to create better incentives for generators to locate in beneficial areas than proposals that would maintain the current flows of payments within a national wholesale market.

6. On the basis of our assessment, we conclude that it is highly doubtful that the proposals we have considered would be effective in addressing the current, recognised, problems with the GB energy system. Moreover, in some cases, these proposals are expected to lead to significant unintended consequences. Figure 1 provides an overview of the proposals we have considered and their expected impacts.

7. In a system with locational energy pricing (such as zonal pricing), these proposed reforms would be unnecessary. In particular, because there would be far fewer network constraints to manage near real time, there would be much less need for intervention by the SO, and much more opportunity for price signals to work effectively in encouraging market participants to make decisions that reduce system costs – and ultimately the bills that consumers pay.

Figure 1: Taxonomy of BM reform proposals and unintended consequences



Source: FTI Consulting

1. Introduction

1.1. Designed and implemented nearly a quarter of a century ago, it is increasingly obvious that the current way in which the GB electricity market operates needs a fundamental re-design to enable it to serve the objectives of decarbonisation, energy security and price competitiveness; 80% of respondents agreed with DESNZ's assessment in the first Review of Electricity Market Arrangements ("REMA") consultation that "*current market arrangements are not fit for purpose*".[4]

1.2. The current electricity market design was conceived in a very different era, when large coal plants built in the 1960s and 1970s were gradually being displaced by new, high efficiency gas plants. Together, coal and gas fired generation met 71% of GB's electricity needs in the year 2000.[5] These large thermal-based plants had to be located where they could economically access their respective fuel. The output of these plants could be controlled relatively easily to meet fluctuations in national demand for electricity. Concerns about climate change were, at the time, merely murmurings by a small minority of environmental lobby groups.

1.3. Twenty-five years later, the country's electricity system faces a very different backdrop. Climate change is a major political issue, and the ambition to have an entirely carbon-free electricity system is a legally binding policy objective of government. Coal-fired power plants have virtually disappeared, while gas generation is now outstripped by renewables. Almost 30% of our electricity is generated from renewable – but intermittent – wind and solar generation.[6]

1.4. While providing great benefits in reduced pollution and emissions, renewable electricity brings new challenges too. Notably, unlike coal and gas plants, renewable generators tend to locate in areas with the highest wind speeds or best solar conditions, which are often distant from where electricity is consumed. Their output also cannot be controlled – rather, their output is dependent on the weather conditions. All of this means that the way in which the electricity market needs to work is very different from that which was conceived in the late 1990s.

1.5. Network build to connect the vast amounts of new renewable generation in Scotland to demand centres has been insufficient. However, under the current national wholesale market design, market participants still operate on the basis of a hypothetical 'copper plate' network, ignoring the physical realities of the GB transmission network, such as the difficulties of transmitting energy from wind generators in the windy seas of northern Scotland to southern England across a limited transmission network. This is a key driver of the dramatic increase in constraint costs observed in recent years.

4 DESNZ, 2023. Review of Electricity Market Arrangements: summary of responses, p. 17 (link)

5 DESNZ, 2023. Digest of UK Energy Statistics (DUKES): electricity, Electricity fuel use, generation and supply (DUKES 5.6) (link).

6 DESNZ, 2023. Digest of UK Energy Statistics (DUKES): electricity, Electricity fuel use, generation and supply (DUKES 5.6) (link).

1.6. Constraint costs arise when electricity that is bought and sold in the electricity market cannot actually be transported from where it is intended to be produced to where it is intended to be consumed because of bottlenecks on the transmission networks. To resolve the problem, the Electricity System Operator ("ESO") intervenes in the market by buying electricity in some parts of the country and selling it in others. These interventions take place after the wholesale market in the so-called balancing mechanism ("BM").

1.7. This cost – the cost of transmission constraints – has risen eight-fold since 2010, from under £170m to over £1.3bn by 2022,[7] and is recovered from customers through an additional charge. More transmission will have to be built in the future, but investing in the transmission network to the extent that constraint costs would entirely go away is costly and has shown to be a very slow process. Even if very ambitious and historically unprecedented transmission expansion plans are achieved, balancing costs are likely to stay high well into the 2030s, at between £2bn and £4bn per annum.[8]

The BM was never designed to minimise system costs and is not set up to do so

1.8. There is now broad acceptance that high and increasing constraint costs are a significant problem for the current energy system in GB. In response, some commentators have proposed reforms either to the BM, or to the role of the System Operator ("SO") in balancing the market near real-time delivery.[9] They argue that such reforms would eliminate the need for a fundamental change to the design of the wholesale market. In assessing these reforms, it is crucial to consider the underlying rationale for the existence and functioning of the BM in GB.

1.9. The original idea underpinning the existing market design, more specifically the New Electricity Trading Arrangements ("NETA"), was for the BM to act as a residual balancer after gate closure, with little impact on the functioning of the preceding wholesale market. In the early 2000's, in its first-year review of NETA, Ofgem wrote: "*A key difference between the Pool and NETA is that now NGC's electricity balancing role has been reduced to that of residual energy balancer, with most balancing being undertaken by participants themselves before Gate Closure. It was also anticipated that NGC's residual electricity balancing role would further decline over time as participants learnt with experience to refine the balancing of their own positions and as Gate Closure moved closer to real time.*"[10]

---

7 ESO, 2022. Net Zero Market Reform: Phase 3 Assessment and Conclusions, p.13 (link).

8 ESO, 2024. Balancing Costs: Annual Report and Future Projections Technical Report, May 2024 (link).

9 Frontier Economics, 2024. Analysis of reform options for status quo electricity balancing arrangements (link). Regen, 2023. Improving locational signals in the GB electricity markets (link). AFRY, 2024. National and Zonal electricity market designs for Great Britain (link). Cornwall Insight, 2023. Reform options for TNUoS and constraint management (link). The Energy Landscape, 2024. Exploring options for constraint management in the GB electricity system: the potential for constraint management markets (link). An overview of proposals is provided in the "Thermal Constraints Collaboration Project" set up by NG ESO (link).

10 Ofgem, 2002. The review of the first year of NETA: A review document. Volume 1 (link).

1.10. The idea at the time of NETA's introduction was that market participants owning portfolios of large, flexible thermal units would be better at scheduling their own fleet than the SO, which would have to rely on an incomplete set of operational parameters disclosed via a centralised market. The gains in scheduling were expected to be higher than the costs incurred due to market participants not having to consider operational network constraints in their decision-making. We are not aware of empirical evidence for this claim, but it is plausible given the initially small scale of balancing actions. In its first year of operation, only about 2% of electricity demand was bought and sold in the BM, of which the large majority was to solve energy imbalances, rather than grid constraints.[11]

1.11. In its report on potential reforms to the BM, Frontier Economics states that the BM has a fundamentally different role: "*The primary purpose of the BM is to ensure security of supply in a manner that minimises the total operational system costs… of all assets on the system*". This reasoning evidences a fundamentally flawed understanding of the BM's purpose. Under the current market design, as conceived in the 1990s and operating to this day, incentivising the minimisation of production costs is the role of the wholesale market– not the BM. Market participants should be incentivised to bid in competitive wholesale markets in a way that ensures energy requirements are met at the lowest possible cost. This rationale still stands today; not only in GB but in many other jurisdictions, introducing market discipline has led to important production savings in electricity generation.[12]

1.12. The current trend in constraint costs does not change the purpose of the BM but it enlarges its role to an extent that was not envisioned when it was established. This means that, as AFRY writes in its report on the topic, "*ESO is increasingly acting as a central scheduler, with a scheduling and dispatch process designed for a residual balancing role, resulting in inefficient operational outcomes.*"

High constraint costs are a symptom of an outdated market design

1.13. The BM clearly no longer plays the residual role that was initially expected. But some commentators have argued that high constraint costs are not necessarily an issue since they typically do not represent inefficiencies, but merely transfers between consumers and producers.[13] Consumers may pay more as a result, but this is counterbalanced by higher returns for producers, so there is no net loss for GB as a whole.

---

[11] Ofgem, 2002. New Electricity Trading Arrangements (NETA) – One Year Review (link).

[12] See for example: Cicala, S., 2022. Imperfect markets versus imperfect regulation in US electricity generation. American Economic Review, 112(2), pp.409-441.

[13] "*In broad terms, minimising dispatch costs will lead to lower customer costs. However, there are complex transfers between generators and customers which may mean, in some specific cases, minimising dispatch costs does not minimise customer costs. We do not consider these specific cases as part of this report. In principle, an idealised version of a national market (under either self or central dispatch) and an idealised LMP market should result in the same physical dispatch of power (if not the same commercial outcomes for participants). This is because the fundamental objective of each of these forms of markets is identical i.e. to satisfy demand over time at lowest cost subject to the constraints imposed by the physical production park and the network.*" From Frontier Economics, 2024. Analysis of reform options for status quo electricity balancing arrangements (link).

1.14. However, this is only true under a set of unrealistic ideal assumptions that do not hold in practice. As we outline in Figure 2, there are significant economic costs of the current wholesale market design. Here we focus on issues related to achieving a cost-minimising dispatch, i.e., taking generation capacity and demand as given. At the end of the note, we also discuss issues related to distorted investment incentives under current national wholesale market design.

Figure 2: Overview of reasons why a national wholesale market plus BM does not lead to cost-minimising dispatch

Source of inefficiency	Key underlying reasons	Examples of consequences	Result
Prices in national wholesale market plus BM fail to lead to an efficient dispatch	Wholesale and BM price signals are not incentive-compatible	Gaming opportunities between the wholesale market and the BM	Failure to use least-cost assets on the power system, with increased costs for GB consumers
	Limited horizon (1h) over which dispatch decisions are made	No seamless functioning of the BM due to thermal ramping requirements, pricing of interconnectors and battery skip rates	
	Bidding strategies in the BM are hard to monitor	Significant differences in bid-offer spreads across technologies	
	Entry barriers for most consumers (and some generators / storage) to participate in the BM	Consumers react to day-ahead wholesale market prices (e.g. via dynamic pricing contracts) and not BM prices	

Source: FTI Consulting

1.15. Prices in the national wholesale market plus BM fail to lead an efficient dispatch for at least four important reasons.

- First, the incentives provided by the wholesale market in conjunction with the BM are not always aligned with system benefits – or even system needs. As a result, the actions of market participants following their individual profit-maximising incentives in the wholesale market are often required to be unwound in the BM as a result of network constraints – with generators redispatched to generate more, or less, depending on which side of a network constraint they are situated on. Even more perverse incentives may arise; for example, large sources of demand in an export-constrained area could often reduce their electricity costs by waiting until the BM to schedule their load.[14] Such behaviour could mean that even less demand would be cleared in export-constrained areas in the wholesale market, further increasing the volumes of required redispatching (and hence consumer costs).[15]

14 In the BM, sources of demand in export-constrained areas would expect to pay a lower price than in the wholesale market, with a high probability of being selected (as the SO's main alternative option would be to curtail renewable generators with low marginal costs).

15 This strategy is referred to as inc-dec gaming in the literature and was observed before the transition to nodal pricing in US ISOs. In the academic literature one of the first papers to describe the inc-dec game was Stoft, S., 1999. Using game theory to study market power in simple networks. IEEE Tutorial on Game Theory in Electric Power Markets, pp.33-40. (link). Inc-dec gaming can only entirely be removed by introducing nodal pricing but zonal pricing can mitigate the problem. An account of experiences in the US with inc-dec gaming is provided by Hogan, W., 1999. Restructuring the electricity market: Institutions for network systems. Harvard University, April (link).

- Second, the SO is often unable to select those assets which would have the lowest cost to resolve constraints because of the limited time horizon over which it makes redispatch decisions. For instance, this makes it difficult to take advantage of least cost thermal assets which require ramping up and ramping down, or to use interconnectors, since doing so may require interactions with other SOs.[16] Dispatching batteries on a short horizon also presents a challenge because it is hard to anticipate for the SO at what particular point in the time the scarce energy in the battery would be of best use. It is possible that the uncertainty around the value of energy stored could be part of the driver of battery skip rates; i.e., the ESO prefers not to activate batteries, but other non-energy limited resources when reserves are plenty to ensure that those batteries could still be activated in later hours when reserves might be more scarce.
- Third, different assets appear to have different bidding strategies in the BM, with significant differences in bid-offer spreads across technologies. Bids in wholesale markets are easier to monitor because there are fewer potential arguments for market participants to inflate their bids above marginal costs. In general, if some technologies or particular units feature higher bid-offer spreads than others, the BM will typically fail to achieve efficient dispatch. For instance, an asset with lower costs may not be selected to resolve constraints because it has a higher bid-offer spread (and thus higher price to the SO) than an alternative.
- Fourth, consumers are typically unable to access the BM (even though increasing shares of consumers are able to respond to day-ahead wholesale prices through dynamic pricing plans). For example, the national wholesale price can be relatively high while wind is being constrained off in Scotland. Scottish consumers could have access to this otherwise-curtailed electricity at zero or even negative prices in the BM but are currently unable to do so. By doing so they would reduce the cost of resolving constraints.[17]

1.16. Today, the raison d'être of NETA no longer exists. The higher constraint costs are, the greater the inefficiencies are likely to be, and the higher the economic costs to GB as a whole. High constraint costs are therefore a key symptom of an outdated market design — even advocates for BM reform are suggesting a much bigger role for the SO in coordinating the progressively more complex interactions between the grid, supply and demand. The question now comes down to whether to solve this coordination challenge by:

- introducing a locational market design; or
- reforming the BM and expanding the role of the SO.

1.17. In this note we evaluate the merits of the latter approach.

16 Commercial issues, rather than technical constraints, are likely to be the main constraint on instructing changes to interconnector flows on a short notice (other than under emergency conditions when bespoke procedures are in place).

17 There are exceptions where consumers are active in the BM via aggregators, but most consumers remain price-takers.

Structure of the note

1.18. This note is structured as follows:

1. We provide a taxonomy of the current BM reform proposals.
2. We critically assess the three broad buckets of BM reform proposals. We show that the proposals are likely to be ineffective in addressing the current, recognised, problems with the electricity market. Moreover, we argue that there is a risk that some of these proposals can lead to significant unintended consequences.
3. We zoom out and discuss the core reason behind the enhanced BM proposals: keeping the existing financial transfers between producers and consumers in place. We argue that these transfers cause long-term and significant dynamic problems, by distorting siting decisions for new investments by consumers and producers and increasing further the already substantial costs of grid upgrades. These problems are only likely to increase without major reforms to wholesale markets.

2. BM reform proposals can be categorised into three buckets

2.1. Based on our review of recent position papers by other consultancies and think tanks commissioned by market participants, we have created a taxonomy of BM reform proposals (Figure 3). We group the BM reform proposals into three broad buckets (which are not necessarily mutually exclusive). These are:

- Bucket 1: Increasing centralised control by the SO ahead of delivery with the aim of reducing the costs of managing any given level of system constraints.
- Bucket 2: Discretion in restricting or altering the scheduling of pre-specified classes of assets in the wholesale market with the aim of reducing the level of constraints.
- Bucket 3: Redesigning constraint management, such as by developing a market for balancing the system with the aim of increasing competition between different types of assets and so reducing the cost of managing constraints.

Bucket 1: increasing centralised control by the SO ahead of delivery

2.2. This bucket contains proposals intending to take a step back towards the market design before NETA ("the Pool"), increasing the SO's centralised control ahead of delivery.

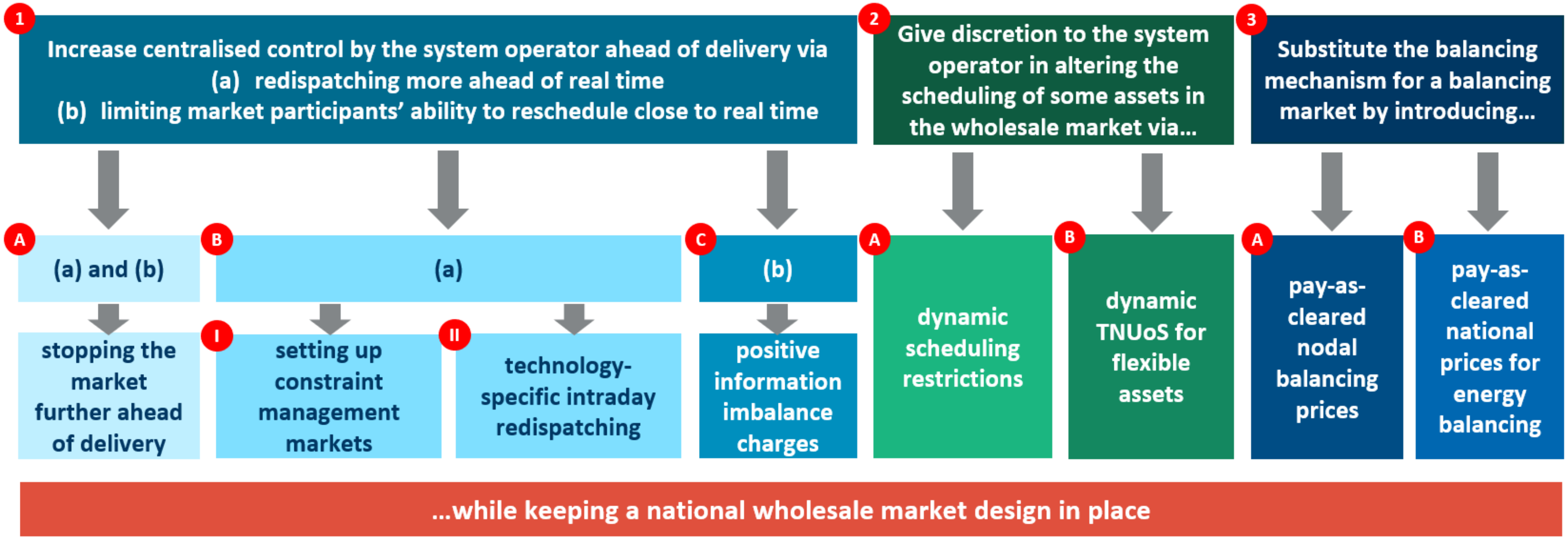
2.3. Proposal 1.A., brought forward by AFRY, Cornwall Insight and Frontier Economics, is to stop the market further ahead of delivery by moving gate closure time forwards, for instance extending the time between gate closure and delivery to six hours rather than the current one hour. By moving gate closure earlier, the potential advantage is that the SO can commence redispatching earlier and possibly improve the dispatching of some assets (e.g., batteries) or have access to resources that otherwise would not be available (e.g., thermal generators that require a long start-up time). The SO could also have greater visibility of the status of the grid at an early stage, and thus select more efficient balancing actions.

2.4. Proposal 1.B. is to start redispatching actions earlier but not necessarily move the gate closure earlier. This can be done via separate constraint management markets (CMMs) (Proposal 1.B.I) allowing potentially access to cheaper BM resources as proposed by AFRY, Cornwall Insight, Regen, and The Energy Landscape. Alternatively, technology-specific redispatching ahead of gate closure (Proposal 1.B.II) can be formalised as suggested by AFRY and Frontier Economics. Currently, some interconnectors are being redispatched ahead of gate closure, but this arrangement would be formalised.[18] Similarly, storage resources could be redispatched earlier to benefit from multi-step optimisation.[19]

[18] Currently the redispatch of some existing interconnectors (but not all) takes place intraday via ad hoc ESO run auctions. For more details we refer to Section 3.4.1 of Frontier Economics, 2024. Analysis of reform options for status quo electricity balancing arrangements (link).

[19] With multi-step optimisation the dispatch of storage is optimised over multiple hours (for example, a 'rolling horizon' of 6 hours that is re-optimised every 30 minutes) rather than within the one-hour gate closure time as currently. By having a longer optimisation horizon, the value of the state of charge across time can be calculated more precisely and the storage resource can be used more efficiently. In contrast, under a one-hour optimisation horizon there is limited information about the value of the state of charge before or after that hour, potentially leading to a suboptimal dispatch of storage.

Figure 3: Taxonomy of BM reform proposals based on a review of recent position papers



Source: FTI Consulting

2.5. Proposal 1.C., suggested by AFRY and Frontier Economics, include mechanisms to improve the quality of initial physical notifications ("IPNs") and/or final physical notifications ("FPNs"), and thus improve the information that is available for the SO when deciding upon which balancing actions to call upon. IPNs and FPNs are asset-specific dispatch information provided by market participants before and at the time of gate closure, respectively. Currently, market participants are only expected to follow "*Good industry practice*" when providing IPNs and FPNs. The informational content of IPNs could in principle be improved by introducing a formal procedure that market participants need to follow to justify differences between IPNs and FPNs. FPNs could be improved by introducing a positive imbalance information charge which would have to be paid for each MWh the asset-specific production/consumption, as metered in the BM, deviates from the FPN. Information imbalance charges are currently set at zero and have lain dormant for 25 years in case needed.

2.6. The proposals in Bucket 1 are not expected to affect scheduling in the day-ahead wholesale market directly as they do not affect firm access rights, and so are likely to have little or no impact on constraint volumes. The overall aim of these proposals is to enable the BM to function more efficiently, thereby reducing the costs of managing a given volume of constraints (see Figure 2).

Bucket 2: giving discretion to the SO to restrict or alter the scheduling of pre-specified classes of assets

2.7. The second bucket contains proposals that would give the SO discretion to restrict or alter the scheduling of certain assets in the wholesale market, without financial compensation, with the aim of reducing the level of constraints. In more technical terms, these proposals would imply to end firm access rights to the wholesale market for certain pre-specified asset classes.

2.8. Proposal 2.A. is to limit how certain assets can be scheduled in the wholesale market by allowing the SO to introduce dynamic restrictions based on system conditions. This proposal is not based on any of the commentators in GB but rather on experiences in other jurisdictions.[20] For example, several continental European countries have imposed limitations on the commercial capacity of interconnectors in the day-ahead time frame to reduce the necessity for redispatch after gate closure.[21] Another example is the non-firm transmission access rights for generators in the National Electricity Market (NEM) in Australia. If there is network congestion, including within the NEM pricing zones, generators face a risk of not being dispatched – being constrained-off the system – or, in some cases, being constrained on.[22] In recent years, intermittent renewables located in export-constrained zones are particularly affected by being constrained-off the system.

[20] In Ireland, Eirgrid has proposed non-firm access rights for new entrants as a temporary solution. Eirgrid, 2022. "SEM-22-068 Firm Access Methodology in Ireland" (link).

[21] See for a discussion: ACER, 2024. Opinion No 02/2024 of ACER of 10 April 2024 on the necessary developments for the fulfilment of the minimum cross-zonal capacity requirements (link).

[22] In contrast to the approach adopted in GB, those generators that are constrained-off are not compensated. While seeming potentially attractive (as customers do not need to pay constrained-off payments), this approach has some material adverse consequences as discussed in FTI Consulting, 2021. Forecast of congestion in NEM. Presentation for ESB (link).

2.9. Proposal 2.B. covers dynamic transmission network use of system charges (TNUoS) to incentivise certain assets to schedule in the wholesale market in a way that reduces constraint costs (relative to how they would be scheduled without the dynamic TNUoS charges). For instance, Frontier Economics has proposed that the SO could determine the level of hourly-varying highly locational TNUoS charges ex-post, with this TNUoS charge based on the difference between the national price and an ex-post estimate of a zonal price.[23] Knowing that the charge would be applied ex-post would lead market participants to take account of the expected value of the charge when purchasing capacity and nominating flows. Another option might be for these charges to be determined ex-ante.

2.10. Frontier Economics proposes to apply such dynamic TNUoS charges to interconnectors. However, there would seem to be no particular reason why the regime could not be extended to other asset classes, such as batteries or electrolysers, if policymakers deemed it desirable. A scheme with some similarities has recently been introduced in Germany via the amended Energy Act. Concretely, the mechanism allows transmission system operators to allocate a certain volume of electricity to eligible subscribers at a discounted price (the '13k price', after Section 13k in the Act) if doing so is expected to lead to a reduction in redispatching costs.[24]

2.11. The objective behind these proposals is to address the root of current problems by reducing BM constraint volumes.[25] They could in principle improve both static and dynamic efficiency by reducing the volume of BM actions required to address constraints and by reducing the incentive to locate generation in constrained parts of the network.

Bucket 3: redesign how constraints are managed by creating a market for balancing the system

2.12. The third bucket contains proposals that suggest changing the BM so that it functions more like a market. Prices received by BM participants would be determined by the bids of the marginal generator or consumer, rather than by each individual participant's bid (so-called 'pay-as-cleared' rather than 'pay-as-bid').

2.13. Proposal 3.A is the Frontier Economics proposal to turn the BM into a real-time extension of the wholesale market with nodal pay-as-cleared balancing prices. Under such an arrangement the prices received/paid by assets activated in the BM in a certain settlement period would be set by the marginal bid or offer that is accepted in their node.[26] We understand that imbalance prices, also referred to as the cash-out prices, would remain national.

---

[23] Frontier Economics suggests that an ex-post estimate of a zonal price could be derived from: (i) BM bids and offers, or (ii) using outputs from a nodal optimisation algorithm.

[24] For more background information see: Bundesnetzagentur, 2024. Start of consultation on the "Use, don't curtail 2.0" determination (link).

[25] An alternative proposal of this type that we do not discuss here in more depth would be for certain assets having to buy access to the grid, e.g., interconnectors. This would be like gas pipelines having to buy entry capacity.

[26] Frontier Economics clarifies that this would imply 'local' clearing prices in the BM for each group of unconstrained nodes in each period.

2.14. Proposal 3.B, introduced by AFRY, entails adjusting the current pay-as-bid pricing arrangement. Balancing actions to resolve energy imbalances ("unflagged") would change to pay-as-cleared pricing while balancing actions to resolve other constraints ("flagged") would remain pay-as-bid.[27] As such, balancing actions used to resolve energy imbalances would be encouraged to bid at their marginal costs. Also, prices for balancing actions to resolve energy imbalances would align with the way imbalance prices are currently computed, i.e., based on unflagged BM actions. Imbalance prices are paid by the unbalanced market participants causing the need for those balancing actions.

2.15. This bucket of proposals has two main objectives: (i) to encourage competition within the BM through improved price signals, thus leading prices to move closer to costs; and (ii) to enable BM prices to be 'bankable' for investors, for instance through enabling trading based on forward markets.

[27] This approach resembles the current default approach in the EU as prescribed by the Electricity Balancing Guideline - Commission Regulation (EU) 2017/2195 of 23 November 2017 establishing a guideline on electricity balancing (link). In practice, determining whether an action is to resolve system constraints or energy imbalances is often not straightforward.

3. Each of the reform proposals could have serious unintended consequences

3.1. In this section, we carry out an assessment of the expected impacts of the different types of reform proposal. We also cover likely negative unintended consequences which their proponents have thus far tended to ignore or downplay.

Bucket 1: more centralised SO control to manage constraints further ahead of delivery is likely to come at the expense of less efficient wholesale markets

3.2. The first bucket of proposals includes three ways to grant more control to the SO further ahead of delivery. The success of such reforms relative to locational pricing hinges on the answer to a fundamental question of market design: whether the market, or a centralised market operator, is better at conducting balancing activities in an efficient manner. Our view in this regard is that generally:

- market mechanisms are better at bringing together lots of diffuse information (e.g. on individual assets); whilst
- centralised mechanisms can be better at co-ordinating for a given set of information, as they will take account of the trade-offs among them.

3.3. In the context of the electricity market, locational pricing has the advantage here, as it retains the role of the market (ensuring allocatively efficient outcomes) whilst ensuring participants factor in the locational trade-offs (which are currently not exposed to market participants resulting in high constraint volumes in the balancing mechanism).

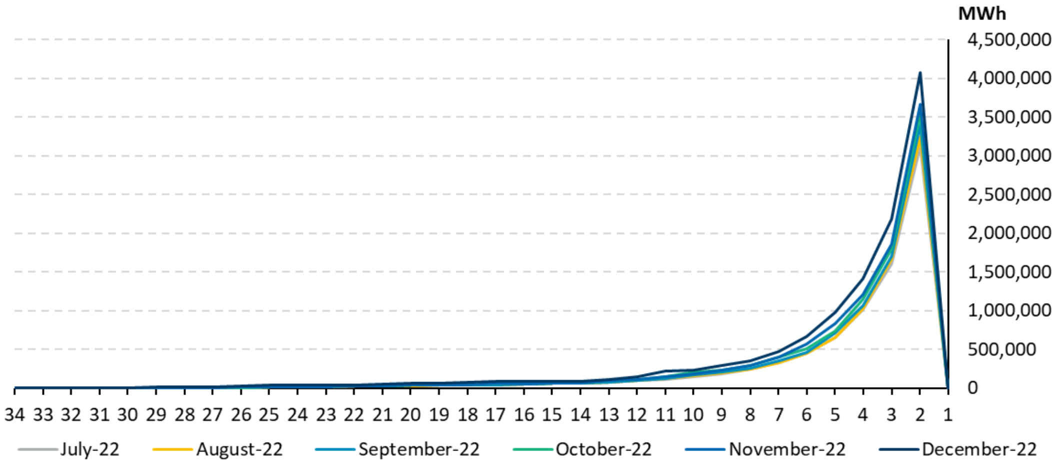
3.4. We discuss these proposals from most far-reaching in transferring control to ESO to least impactful.

Proposal 1.A: Stop the wholesale market further ahead of delivery

3.5. The claimed benefits of moving gate closure earlier are that the SO would be able to improve the efficiency of final dispatch by having a longer optimisation horizon (e.g., batteries) and by having access to resources that otherwise would not be available (e.g., thermal generators that require a long start-up time). Moreover, since market participants would not be able to reschedule their assets, the SO could have greater visibility of the status of the grid, potentially enabling it to select more efficient balancing actions.

3.6. However, these potential benefits come with an important trade-off. By moving gate closure earlier, market participants would no longer be able to reschedule their assets close to real time in the intraday market as more accurate information becomes available about their final consumption or production. This is especially important in the current context with increasing volumes of production from intermittent renewables for which production forecasts improve significantly in the hours before delivery. Figure 4 shows the volume of matched intraday trade in the EU within the hours before delivery. It shows that most of the trade happens within 3 hours of delivery, suggesting a strong preference among market participants to trade close to real time.

Figure 4: Total volume matched in the EU Single Intra-Day Coupling (SIDC) within hours before delivery.



Source: Market Coupling Steering Committee, 2023 (link)

3.7. One of the key ideas of liberalised electricity markets is to make market participants responsible for accurate forecasting of production and demand. Reduced forecast errors, potentially traded out via intraday markets and incentivised by imbalance prices that reflect the cost of imbalances, should lead to lower overall system imbalances and improved system security. In the short run, this means that the SO needs to intervene less to resolve imbalances. In the longer run, it means that lower volumes of reserves, procured as an insurance against potential imbalances, are required. This can result in important consumer savings as the costs of reserve procurement are typically socialised across consumers.

3.8. None of the papers pitching the idea to move gate closure earlier discuss to whom the responsibility to forecast production and demand accurately would be allocated in such a scenario. This is a vital question without a clear answer. On one extreme, market participants could remain responsible. This would lead to a significant increase in the imbalance cost risk for intermittent renewables. Storage providers would also find it harder to implement intraday trading strategies, for which the bids/offers in one hour depend on market realisations in the preceding hours. Both of these effects could be expected to increase average wholesale prices and ultimately consumer costs. On the other extreme, the SO could be made responsible for forecasting demand and production. This would imply an implicit subsidy for renewables and granting control over the intraday dispatch of batteries to the SO.[28] In that case, increased energy imbalances can be expected because:

i. the SO typically has less information about the assets and local circumstances than the market participants that own the assets; and

ii. it is difficult to incentivise the SO to forecast accurately as balancing costs due to forecasting errors can typically be passed through via transmission charges.[29]

[28] In the past, in many European countries, intermittent renewables have initially been exempted from balance responsibilities to encourage adoption.

[29] Regulators can try to incentivise accurate forecasting by the SO, but such incentive schemes are typically complex and imperfect. See, e.g., Ofgem, 2023. "The Electricity System Operator Reporting and Incentives Arrangements: Guidance Document" (link).

3.9. In short, **when trying to address one problem by moving gate closure further ahead of real time, another problem is created – increased expense and difficulty of managing balancing actions for energy imbalances.**

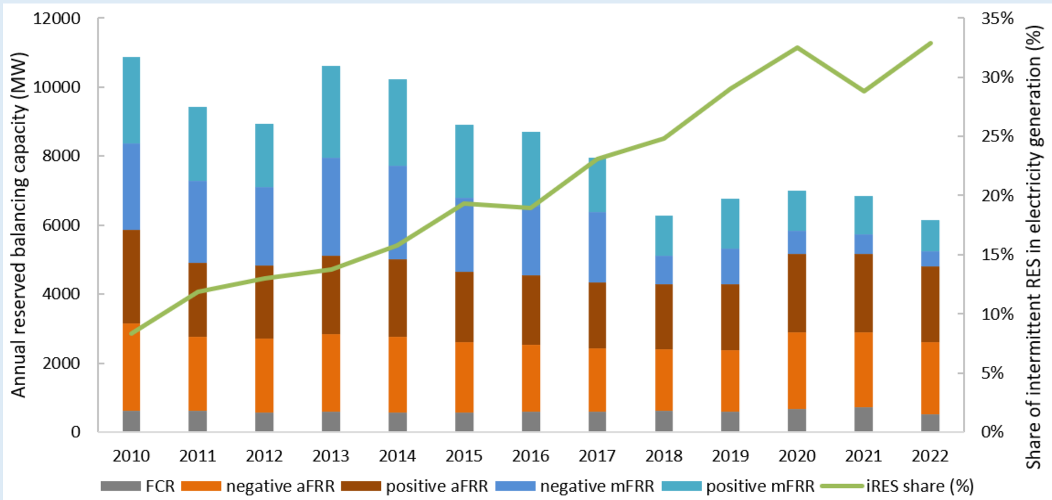
3.10. Box 1 discusses the evolution of reserve procurement in Germany, demonstrating some of the benefits available from later gate closure.[30] Conversely, moving gate closure earlier in GB could be expected to result in the opposite trend – increasing reserve procurement due to greater system imbalances.

Box 1: The German balancing paradox

Between 2010 and 2022, the volume of procured reserves in Germany fell by about half while the share of intermittent renewable generation tripled (Figure 5). Procurement of the slowest type of reserve, manual frequency restoration reserves (mFRRs), has fallen most. This reduction in the volume of reserves procured has led to important customer savings without any apparent impacts on security of supply.

The counterintuitive combination of increasing intermittent renewables and falling procurement of reserves is ascribed to two reasons: improved cooperation among national and international SOs;[31] and changes to the market design. Market design changes have enabled market participants to trade out their imbalances in intraday markets more easily, for instance by moving gate closure time closer to delivery and introducing 15-minute products in the continuous market. These have resulted in significant increases in volumes of intraday trade.[32]

Figure 5: Procured volume of reserves to resolve energy imbalances versus the share of generation by intermittent renewables in Germany.[33]



Source: FTI Consulting, data from Regelleistung.net (link) and Netztransparenz.de (link)

---

30 In Germany actions instructed by the SO to resolve energy imbalances (via the balancing market) and actions to resolve system constraints (via cost-based redispatching) are not necessarily optimised jointly as in the BM in GB. Here, we focus on actions instructed by the SO to resolve energy imbalances.

31 The common balancing power market among the four German TSOs as part of the German Grid Control Cooperation (Netzregelverbund) was already introduced between 2009 and 2010.

32 See Ocker, F. and Ehrhart, K.M., 2017. The “German Paradox” in the balancing power markets. Renewable and Sustainable Energy Reviews, 67, pp.892-898 (link).

33 FCR: Frequency Containment Reserves/ aFRR: Automatic Frequency Restoration Reserves/ mFRR: Manual Frequency Restoration Reserves/ iRES: Intermittent Renewables.

3.11. Experiences in Germany and other countries lie behind the recent EU Regulation proposing a further reduction of gate closure time.[34] It is stated that: *"Since variable renewable energy generators are only able to accurately estimate their production close to the delivery time, it is crucial for them to maximise trading opportunities via access to a liquid market as close as possible to the time of delivery of the electricity. The gate closure time of the cross-zonal intraday market should therefore be shortened and set closer to real time in order to maximise the opportunities for market participants to trade shortages and surplus of electricity and contribute to better integrating variable renewable energy sources into the electricity system."* Concretely, the Regulation mandates that from 1 January 2026, the intraday cross-zonal gate closure time shall not be more than 30 minutes ahead of delivery.

3.12. The trade-off in increasing energy imbalance volumes versus more cost-efficient management of system constraints only bites under the existing national wholesale market design. Locational pricing should significantly reduce the volume of balancing actions to resolve system constraints. As a result, there would be no need to move gate closure further ahead of real-time and the responsibility to forecast demand and production could be left to market participants who are arguably better at balancing their positions. Transitioning to a locational market design is likely to reduce further the socialised costs of resources to manage system imbalances, through better system visibility and reduced risk that procured resources are not deliverable in real time.[35]

Proposal 1.B: Organise constraint markets running in parallel with the wholesale market

3.13. Rather than moving gate closure earlier, an arguably slightly less far-reaching proposal is to reserve units for redispatching and/or start redispatching actions earlier than the current gate closure time, i.e., in parallel with intraday wholesale markets. This can be done via technology-agnostic constraint management markets (CMMs) (Proposal 1.B.I) or technology-specific redispatching ahead of gate closure (Proposal 1.B.II).

3.14. There are two overarching concerns with these proposals. First, the original point of gate closure (initially set at 4 hours) was to give the SO a chance to balance the system in a time window during which market participants could not change their intended output. The **major risk of resolving constraints ahead of gate closure is that market participants might at a later stage adjust their positions in a way that is unhelpful for the system** (e.g., recreating the constraints that were intended to be resolved by earlier redispatched units). To avoid this to happen, the possibility to deviate from day-ahead positions needs to be limited, hence **reducing activity in the intraday market**. In that sense, there is a link between proposals from Bucket 1.C aiming at better monitoring changes in schedules and proposals from this bucket. Second, the activation costs of 'regular' BM actions are not yet known when decisions would need to be made in the intraday timeframe about which units to redispatch (procured via a CMM or technology-specific). **Incomplete information would often lead to sub-optimal final dispatch, raising consumer costs.**

---

[34] European Parliament legislative resolution of 11 April 2024 on the proposal for a regulation of the European Parliament and of the Council amending Regulations (EU) 2019/943 and (EU) 2019/942 as well as Directives (EU) 2018/2001 and (EU) 2019/944 to improve the Union's electricity market design (link).

[35] When ERCOT transitioned from a zonal to a nodal market design in 2010, the volumes of Regulation Up and Down procured fell by nearly 50%, see Andrade, J., Dong, Y., and Baldick, R. 2018. "Effect of market changes on the required amounts of frequency regulation ancillary services in ERCOT" (link).

3.15. With regards to CMMs, there is no universal definition of a CMM as there is currently little experience of their operation; but in general, they involve a single buyer (the SO), are voluntary, and intend to be technology-agnostic. In specific cases, when also referred to as local constraint markets (LCMs), they focus on the procurement of distributed energy resources (DERs), including demand, which often face entry barriers to the BM. Market participants in a CMM are paid an availability payment when being reserved for constraint management purposes, and possibly an activation payment when eventually being turned up or down before or in the BM. Availability contracts can be auctioned off from years ahead down to several hours ahead of gate closure. The claimed benefit of CMMs is that the predictability of having day-ahead (or longer) commitments to participate in constraint management may open up a new group of potentially cheaper providers that would otherwise not be able to participate. However, there are well-known issues with CMMs which can be divided into two parts.

3.16. First, CMMs are prone to gaming. The same market participants that are awarded availability contracts in the CMM are those that can cause a need for a CMM in the first place. Ehrhart *et al.* study the incentives of market participants in reservation markets for constraint management purposes.[36] They find that the availability payment incentivises participants to change their energy consumption or generation behaviour in unhelpful ways, which eventually increases the volume of constraints.[37] They conclude that such mechanisms "[do] *not resolve network constraints, while causing costs for the compensation payments*". CMMs could only function effectively were it possible to filter participants and exclude "*undesirable participants*", i.e., those market participants that change their behaviour because a CMM is in place. However, this is likely to be very difficult in practice as elaborated upon in the same study.

3.17. Second, CMMs will typically lead to an inefficient allocation of resources between markets. Market participants bidding in CMMs will have to estimate the opportunity cost of their participation in the wholesale market. For example, a generator that has committed to be available via a CMM cannot offer its entire capacity in the wholesale market. Estimating opportunity costs requires accurate forecasting of wholesale prices, with any mistakes in wholesale price predictions leading to an inefficient allocation of resources and ultimately increased wholesale prices and total consumer costs.[38]

---

36 Ehrhart, K.-M., Eicke, A., Hirth, L., Ocker, F., Ott, M., Schlecht, I., and Wan, R., 2024. "Analysis of a Capacity-Based Redispatch Mechanism" (link).

37 For example, a battery located behind an import-constraint can decide to charge up in the day-ahead wholesale market knowing that it would be eligible to be reserved in a subsequent CMM to be potentially asked to discharge rather than charge (and receive an availability payment for it). In case the battery would not have decided to charge in the day-ahead wholesale market, it would not have been eligible to receive the availability payment. In addition, due to choosing to charge in the day ahead wholesale market, the battery aggravated congestion.

38 This is not an issue specific to CCMs but a general issue for reservation markets organised separately from wholesale markets, such as the current sequential procurement of reserves and response.

3.18. Another related proposal is not to organise a technology-agnostic CMM before gate closure time but instead only start redispatching assets from pre-specified technology classes ahead of gate closure. The standard example has been interconnectors. Frontier Economics proposes to generalise specific intraday arrangements allowing ESO to redispatch interconnectors ahead of gate closure time across all interconnectors. This proposal is, besides likely to lead to inefficiencies, also discriminatory. There is a risk for the creation of a patchwork of gate closure times running in parallel for different technologies; equally as for interconnectors, gate closure for batteries could also be moved forward so batteries can be more optimally dispatched across a multi-hour horizon.

3.19. As with earlier gate closure, there are significantly reduced potential benefits of distortionary CMMs or discriminatory, technology-specific early redispatching actions if locational pricing is in place. There would be much less need for such interventions, since improved scheduling in the wholesale market under locational pricing would dramatically reduce overall constraint volumes.

Proposal 1.C: Improve the ESO's visibility of market conditions by incentivising market participants to provide improved asset-specific dispatch information

3.20. Market participants provide asset-specific dispatch information through IPNs and FPNs before and at the time of gate closure, respectively. According to the Grid Code, they are expected to follow "*Good industry practice*" when providing IPNs and FPNs. Differences between IPNs and FPNs can be explained by updated forecasts of market participants about their expected consumption or production. Market participants can choose to act on these updated forecasts via rescheduling their own portfolio or via intraday trade.

3.21. Frontier Economics and AFRY propose the introduction of formal procedures that market participants must follow to justify deviations between IPNs and FPNs. Unlike the other proposals we consider in Bucket 1, this proposal is likely to have only relatively small benefits or costs. The benefits include potential incremental improvements in the SO's ability to identify likely constraint problems, and thus allowing the SO to tackle them at an earlier stage. The costs include additional administrative burdens and potential reductions in market participants' willingness to trade close to real time.

3.22. FPNs are submitted at gate closure, and for any individual asset should be consistent with expected physical output (and across a portfolio should, given the underlying incentives of the NETA market design, also reflect the net contracted position). However, FPNs may not reflect actual dispatch on an asset-by-asset basis as market participants only have a financial incentive to be balanced at the portfolio-level.

3.23. Frontier Economics and AFRY additionally propose introducing positive imbalance information charges whenever there is a difference between expected and actual delivery of volume in the BM.[39] This proposal could help to produce higher quality FPNs but would come with a potential cost resulting from double penalisation of imbalances for market participants which would particularly impact harder-to-forecast intermittent renewables and demand. For example, if a market participant is imbalanced at portfolio-level, it implies that at least one asset of that market participant is also imbalanced at unit-level. As such, that market participant would be penalised twice for the same imbalance. In addition, while imbalance prices are typically market-based, information imbalance charges would have to be set administratively. There is little theoretical guidance on what the optimal level of an information imbalance charge should be, creating a significant risk of excessive imbalance charges, that might potentially lead to excessive overinvestment by market participants in forecasting analysis and other tools to avoid charges, relative to the socially efficient level.

3.24. Under locational pricing, the SO would have greater direct visibility as market participants would only be able to benefit from portfolio balancing across assets located in the same zone or node. As such, the positions at portfolio-level and unit-level should be more aligned. Also, contracts between market players would be settled at a specific zone or node, further improving the information available to the SO. Locational pricing would therefore be expected to reduce any potential benefits of non-zero information imbalance charges.

Overall conclusions on first bucket of proposed reforms

3.25. The proposed reforms discussed above have a common feature of not seeking to change the access rights of market participants to the grid. Rather, all participants would continue the practice of trading in the GB wholesale market as if there are never transmission bottlenecks and would continue to receive compensation payments to the extent that transmission constraints actually occur. Hence, the proposed reforms cited above can, in no way, address the volume of constraints that occur because they do not address the underlying the incentives that the current structure of the wholesale market provides (and would continue to provide if any of the above proposals were implemented).

3.26. The reforms might, at the margin, allow the SO to trade out the (same volume of) constraints in a slightly more cost-effective manner in the BM than now. However, even this would come at a very significant costs surfacing elsewhere. As we have explained above, these costs would include either much reduced trading opportunities for market participants near real time; the introduction of new risks of inefficient actions by the SO; very significant risks of gaming (that would be costly to customers); and the risk of arbitrary charges being levied upon market participants.

[39] The original idea behind information imbalance charges was to avoid gaming at portfolio-level.

Bucket 2: giving discretion to the SO to alter or restrict the scheduling in the wholesale market for certain classes of asset would bring significant risks

3.27. Proposals in the second bucket would in practice end firm access rights for certain assets by giving discretion to the SO to alter or restrict these assets being scheduled in the wholesale market without providing a financial compensation. This type of proposal can be implemented either through dynamic scheduling restrictions or dynamic TNUoS charges.

3.28. With dynamic scheduling restrictions, the SO would set limits on how certain assets can be scheduled in the wholesale market based on system conditions. These limits can change from one hour to another. The aim of scheduling restrictions is to reduce constraint volumes, particularly for two-way assets such as interconnectors. Under the current national wholesale market design, interconnectors are at times scheduled 'in the wrong way'. For example, interconnectors between southern England and France are frequently scheduled to export from GB even though local system conditions in England suggest that they should be importing.[40] Under dynamic scheduling restrictions, the SO would prevent the interconnector from exporting or importing if doing so would exacerbate GB system congestion. The same logic can hold for restricting the scheduling of wind generators, batteries or electrolysers.

3.29. Alternatively, assets that are currently 'wrongly scheduled' in the wholesale market could be made subject to dynamic TNUoS charges. Frontier Economics proposes ex-post TNUoS charges for interconnectors based on the difference between the national price and an ex-post estimate of a zonal price. Knowing that the charge would be applied ex-post would lead market participants to take account of the expected value of the charge when purchasing capacity and nominating flows.

3.30. Different to the Frontier proposal, the SO could set the dynamic TNUoS charges ex-ante. Conceivably, there is no *a priori* reason why dynamic TNUoS charges could not also in principle be applied to other assets classes such as batteries and electrolysers that are deemed by policymakers to merit being charged in this manner. Such charges can be more nuanced than scheduling restrictions (which tend to be binary, with plant either allowed to operate or not), and thus could result in scheduling of generation and demand that is close to what would be expected under locational pricing.

3.31. In theory, optimally-determined dynamic TNUoS charges could replicate the effects of nodal pricing by encouraging market participants to take likely network constraints into account when bidding in wholesale markets. However, we note that Ofgem have already stated that TNUoS signals are an inefficient means of influencing real-time operational decisions by market participants.[41]

---

[40] This can lead to very high BM costs to unwind such scheduled trades – see Current-news, 2022. "UK buys power from Belgium at record prices of nearly £10,000/MWh" (link).

[41] Ofgem, 2023, Open letter on strategic transmission charging reform (link).

3.32. The costs and benefits of ending firm access therefore look similar to those of locational pricing – at least for the asset classes impacted. However, applying non-firm access to only a subset of the market participants is quite obviously and unambiguously discriminatory. Whether this discrimination is due or undue, it would need to be considered very carefully – but there seem to us several very significant additional risks of this approach, of which we highlight four: legal concerns, transparency of price formation, risk management and inefficiencies in scheduling.

i. **Legal concerns**. Proponents of dynamic scheduling restrictions or dynamic TNUoS charges argue that such changes may be simpler to implement than locational pricing. However, it is far from obvious that this is the case in practice. Removing firm access rights would have significant negative impacts on the revenues of the relevant assets and likely result in protracted legal challenge.[42] In the specific case of interconnectors, decisions by the GB SO would also affect wholesale prices in interconnected countries and thus consumer costs and producer income in those countries. EU regulations have historically been interpreted as prohibiting dynamic restrictions on cross-border lines, as they are seen as obstructing the free trade of goods across borders.[43] It is unclear how dynamic restrictions or dynamic TNUoS charges on interconnectors would affect current agreements on cross-border trade.

ii. **Transparency of price formation**. Because dynamic scheduling restrictions or TNUoS charges would imply much more discretion of the SO than locational pricing,[44] they are unlikely to have the benefits of transparency or bankability that market pricing can achieve.

iii. **Risk management**. Market participants may find it hard to forecast how the net revenues they will receive will develop over time as scheduling restrictions or dynamic TNUoS charges evolve. Market parties would be unable to hedge these risks in forward markets.

iv. **Inefficiencies in dispatch**. Depending on the precise implementation of dynamic scheduling restrictions or TNUoS charges, there are high risks of inefficiencies in scheduling. Below we highlight three examples.

    a. In case a specific asset-class is subject to dynamic scheduling restrictions or dynamic TNUoS charges, situations could occur where assets from that class are constrained-off where it would be more efficient to constrain off another class of asset that is not subject to such restrictions or TNUoS charges.

---

42 This has been seen in practice with the more limited reforms to charging methodologies that Ofgem has introduced in recent years (see e.g., link).

43 In Art. 16 (8) of Regulation (EU) 2019/943 it is stated that: "*Transmission system operators shall not limit the volume of interconnection capacity to be made available to market participants as a means of solving congestion inside their own bidding zone or as a means of managing flows resulting from transactions internal to bidding zones.*" In 2011, The European Commission's inquiry regarding the Swedish TSO purposely limiting interconnection capacity for cross-border trade led to the voluntary commitment of the Swedish TSO to split the formerly national Swedish price zone in four. The Commission acted on the basis of Art. 102 TFEU, the EU provision prohibiting abuses of dominance. It considered that, by creating obstacles to market integration, the Swedish TSO was violating EU competition law.

44 For example, daily decisions on what restrictions to introduce based on heuristics or untransparent and often updated methodologies for calculating dynamic TNUoS sensitive to in-house often non-public forecasts of system conditions.

b. If all asset classes are subject to dynamic scheduling restrictions, inefficiencies could occur because market participants may be tempted to distort their wholesale market bidding to ensure that they are scheduled (so-called 'disorderly bidding').[45]

c. Dynamic TNUoS charges would inevitably result in inefficiencies due to forecast errors. If they are set ex-ante, errors by the SO in setting optimal charges would lead to inefficient scheduling. If they are set ex-post, errors by market participants in anticipating likely charges would lead to inefficient scheduling.

3.33. Locational wholesale pricing could be expected to produce benefits relative to dynamic scheduling restrictions or TNUoS charges because more optimal scheduling of all assets would result directly from the reflection of transmission constraints in the wholesale market clearing algorithm. Moreover, locational wholesale prices would be determined by transparently agreed upon delineations of bidding zones, and not by potentially untransparent and unpredictable discretionary choices of the SO. Finally, locational wholesale prices can be hedged via long-term contracts possibly in combination with financial transmission rights ("FTRs").

Bucket 3: changing how prices are determined in the BM would often lead to higher consumer costs rather than lower

3.34. The third bucket of proposals contains two ideas for redesigning the balancing mechanism, with the aim of reducing the costs of constraints for a given volume. We first discuss the most radical proposal, before discussing a proposal which has been implemented in other European countries with mixed success.

Proposal 3.A: Bolting a pay-as-cleared nodal market in the BM onto a national wholesale market further exacerbates gaming opportunities

3.35. The current pay-as-bid and proposed pay-as-cleared market structures represent two fundamentally different approaches to the balancing mechanism. The current pay-as-bid arrangement in the BM, in a context where most of the BM actions are taken for system constraints reasons, is in its essence meant to be a cost-recovery mechanism. Generators are dispatched at their bid/offer prices to minimise balancing costs. This is reflective of the stated purpose of the BM, as a mechanism for corrective action by the SO to ensure that dispatch does not violate system constraints.

3.36. Introducing a market-based approach through a nodal pay-as-cleared arrangement in the BM (as suggested by Frontier Economics) implies that the BM becomes a real-time, more locationally granular extension of the current wholesale market. While such an approach might be superficially appealing, it would introduce major distortions that would amplify the already-existing gaming opportunities. Concretely, demand in export-constrained areas and generation in import-constrained areas would have a greater incentive to transact more in the BM rather than the national wholesale market, which can further exacerbate constraint volumes (and ultimately constraint costs).

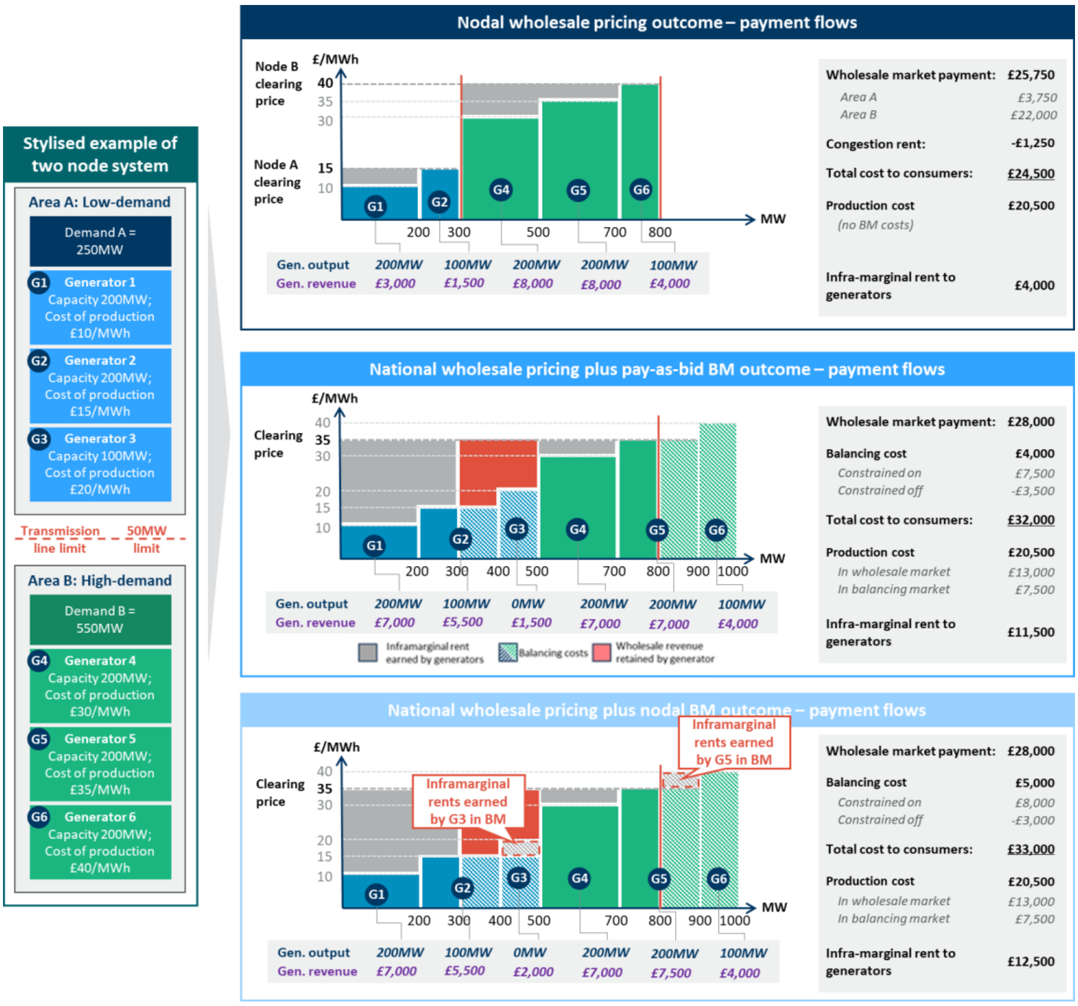
[45] For example, in the case that a generator is expected not to be scheduled due to local constraints (even though the wholesale price is still positive), the generator would be incentivised to bid as low as possible to be sure it is not constrained off. As such, it could be that eventually another generator is constrained off with a lower marginal cost.

3.37. We illustrate these dynamics based on a worked-out example introduced in Figure 6; full details are in the Appendix. The figure shows the market outcomes and payment flows under three market designs: nodal wholesale pricing, national wholesale pricing plus a pay-as-bid BM (status quo), and national wholesale pricing with a nodal, pay-as-cleared BM.[46]

3.38. Under ideal assumptions (see Figure 2 for why these often do not hold in reality) the final dispatch is the same under all three market designs. However, there are significant differences in consumer costs and producer rents in this example:

- Nodal wholesale pricing results in consumer costs of £24,500 and producer rents of £4,000.
- The status quo arrangements result in consumer costs of £32,000 and producer rents of £11,500.
- National wholesale pricing with a nodal (pay-as-cleared) BM results in consumer costs of £33,000 and producer rents of £12,500.

Figure 6: Stylised example of nodal wholesale pricing vs national wholesale pricing plus pay-as-bid BM versus national wholesale pricing plus nodal BM



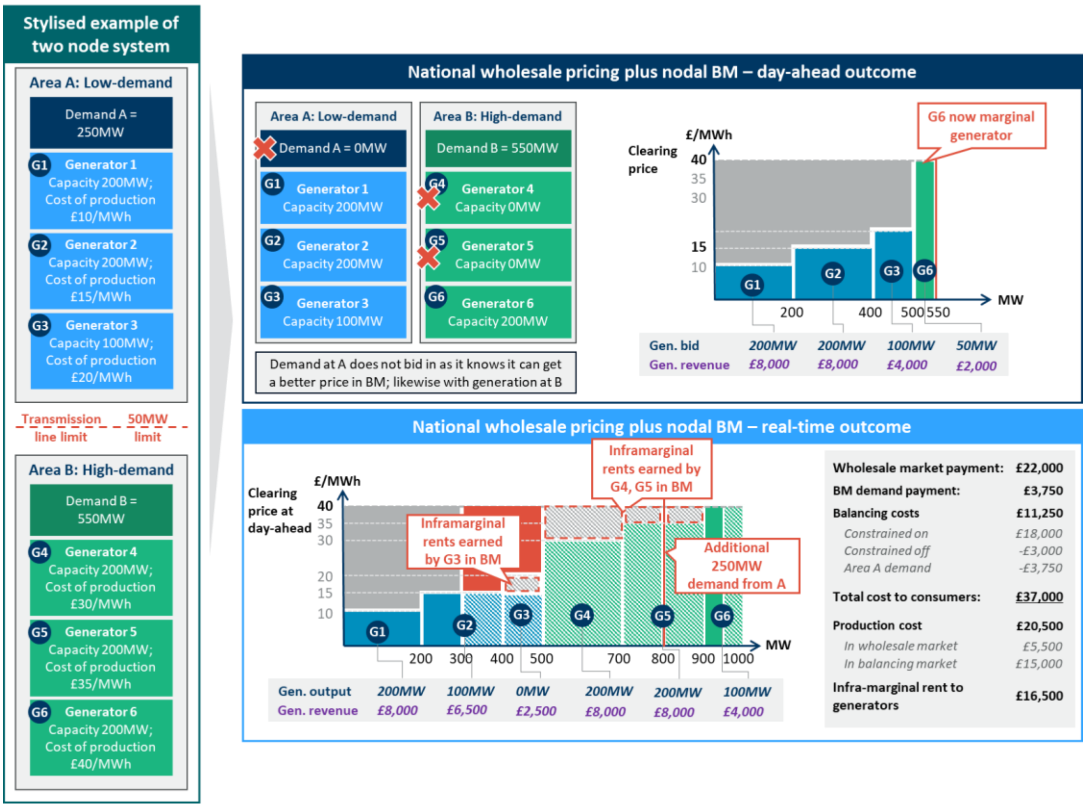
Source: FTI Consulting

46 In this example, we refer to the two zones as 'nodes'. We note that, assuming no intra-zonal constraints, the outcomes would be the same under a nodal or zonal wholesale market.

3.39. The above distribution of rents is predicated on the assumption that all market participants bid their entire supply and/or demand in the day-ahead market. However, under a pay-as-cleared BM, their optimal strategy is different, even without considering any manipulations of bids.[47] Namely, a pay-as-cleared BM price will disincentivise participation in the wholesale markets for those participants who can go into the BM and get paid a more favourable nodal balancing price than the prevailing ex-ante national wholesale market price. This potentially distorts the ex-ante market outcomes, leading to a greater quantity of corrective actions required by the SO in the BM.

3.40. Figure 7 restates the market outcome under national wholesale pricing with a nodal (pay-as-cleared) BM but now assuming this profit-maximising strategy, such that: (i) demand in the export-constrained Zone A waits to contract; and (ii) two generators in the import-constrained Zone B also do not offer their capacity in the wholesale market. Full details of this worked example are also in the Appendix.

Figure 7: Incentives under a national wholesale market with a nodal BM



Source: FTI Consulting

3.41. The final dispatch is the same as in Figure 6, but the quantity and cost of corrective actions taken by the SO are even greater. This leads to:

- a higher ex-ante wholesale price of £40/MWh;
- a consequently higher cost to consumers of £37,000, relative to the national wholesale market plus pay-as-bid BM (assuming no manipulation) cost of £32,000, and nodal market outcome of £24,500; and

[47] Including the possibility to bid higher than marginal costs in the pay-as-bid BM would lead to similar issues.

- higher inframarginal rents for producers of £16,500 at the expense of consumers.

3.42. In short, a pay-as-cleared nodal balancing price allows market participants to capture the locational value of energy, but by doing so can significantly exacerbate BM volumes at the expense of consumers.

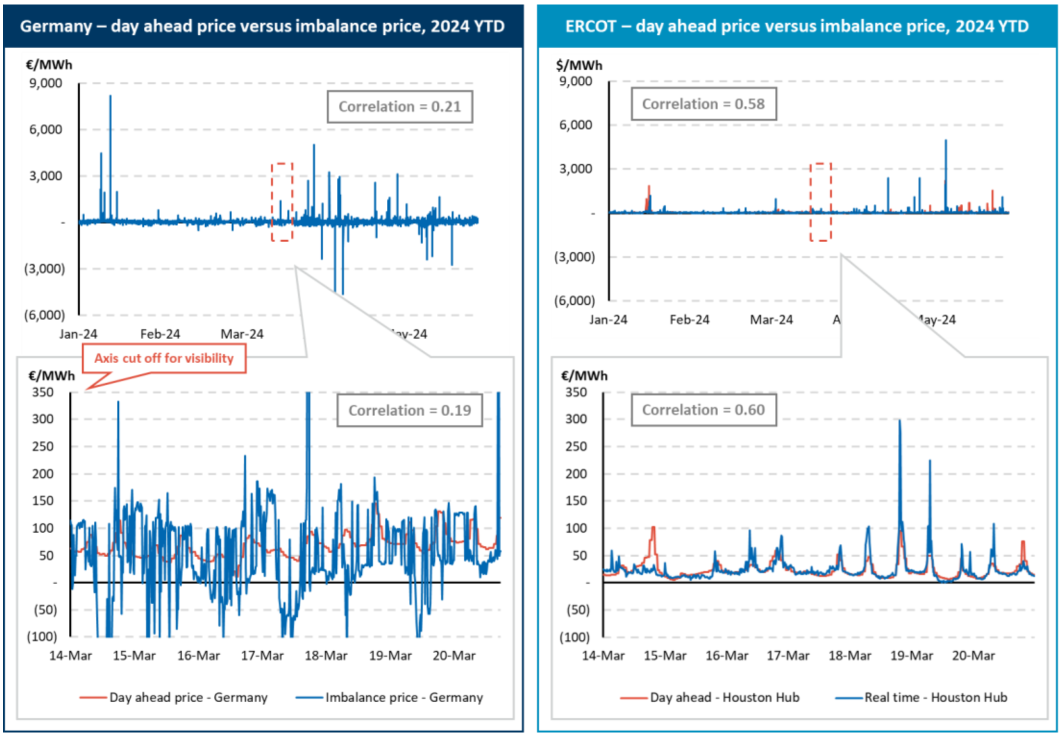
Proposal 3.B: introducing pay-as-cleared pricing for balancing actions triggered to resolve energy imbalances will likely lead to higher BM costs

3.43. Currently, all assets that are activated in the BM are paid their bid price. However, balancing actions that have been called upon to resolve system constraints are flagged – and the marginal bid from all unflagged actions per market settlement period sets the imbalance price that is paid by unbalanced market participants. The idea behind this proposal is to price balancing actions to resolve energy imbalances pay-as-cleared while balancing actions to resolve other constraints would remain under a pay-as-bid arrangement. As such, prices for balancing actions to resolve energy imbalances would align with the way imbalance prices are currently computed. The idea is that such a change would create a 'real-time price' for balancing actions resolving energy imbalances where market parties optimally bid in their marginal cost, which would increase competition and could be used as a basis to settle forward contracts.

3.44. In theory, there are benefits of creating a real-time price for balancing actions resolving energy imbalances, particularly as the share of intermittent renewables increases. However, separating the pricing of BM actions for system constraints and energy imbalances while keeping the current national wholesale market in place would create significant problems. Figure 8 shows the day-ahead wholesale and imbalance prices in Germany (left) and the day-ahead and real-time wholesale prices in the Houston Hub in ERCOT (right). In Germany, SO actions to resolve system constraints are paid their regulated costs, while SO actions to resolve energy imbalances are paid-as-cleared (similar arrangements are in place in many other EU countries included in the Single Electricity Market for the island of Ireland). The shown imbalance price is derived from the latter. In ERCOT, the nodal wholesale market is organised as a two-settlement market (day-ahead and real-time). Imbalances are charged the real-time price.

3.45. In the case of ERCOT, is it immediately apparent that the day-ahead price serves as an expectation of the real-time price. Overall, both price series are highly correlated, i.e., the correlation between day-ahead and real-time wholesale prices is 0.58 for the considered period. In contrast, the German imbalance price resembles a very erroneous signal floating around the day-ahead price, with a correlation of 0.21 for the considered period. The high volatility of the imbalance price makes it unsuitable to expose consumers directly to the imbalance price, or to use the imbalance price to settle long-term contracts. The lesson here is that a lot more adaptations to the current market design would need to be introduced to obtain a valuable real-time price signal other than merely changing the pricing rule for BM actions resolving energy imbalances.

Figure 8: Hourly day-ahead wholesale and 15-minute imbalance prices in Germany (left) and the hourly day-ahead and 15-minute real-time wholesale prices in the Houston Hub in ERCOT (right) for January to June '24 (top) and a two-week snapshot from 14-21 March '24



Source: FTI Consulting

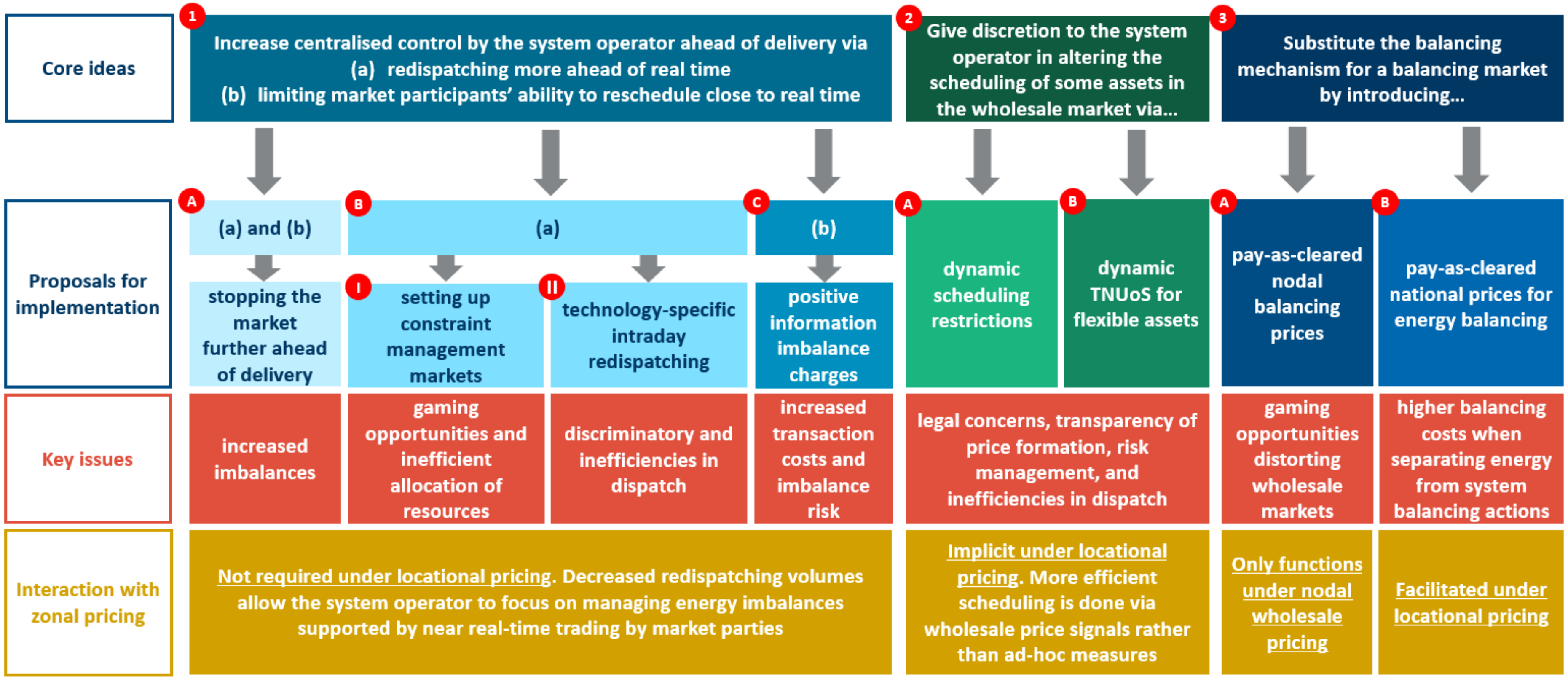
3.46. In addition, accurate flagging of BM actions resolving system constraints is nearly impossible. In practice, BM actions need to be taken with both purposes in mind, as balancing actions aiming at resolving energy imbalances can affect required constraint management actions, and vice versa. In other words, there is no separable merit order for both type of BM actions. The sequence of actions taken by the SO will impact flagging and, as such, could have unpredictable and arbitrary impacts on the pricing of balancing energy. Entirely separating the processes to resolve energy imbalance and system constraints, as in the case of Germany and prescribed by the EU Regulation on electricity balancing, is neither practical nor efficient for either the ESO or for market participants. An additional layer of complexity would be added to what is an already complex market which could act as another barrier to participation.

3.47. Under locational pricing, as the volume of flagged BM actions would reduce, the balancing energy price signal would become less polluted making it more straightforward to price BM actions resolving energy imbalances pay-as-cleared. To obtain a robust real-time price signal, more market design adaptions are likely required to allow the introduction of real-time market substituting a balancing mechanism.

Summary

3.48. Overall, therefore, it is highly doubtful that the proposals we have considered would be effective in addressing the current, recognised, problems with the GB electricity market. Moreover, many of the proposals have the potential to create significant unintended consequences. Figure 9 provides an overview of likely unintended consequences of the different reform proposal.

Figure 9: Taxonomy of BM reform proposals and unintended consequences



Source: FTI Consulting

4. Advocates of the reforms to the status quo fail to consider their longer-term impacts

4.1. Many of the proponents of reforms to the status quo are, under the current market design, the beneficiaries of significant transfers – notably (but not only) generators that benefit from constrained-off payments. This is acknowledged in Frontier's report, which states that "*…there are complex transfers between generators and customers which may mean, in some specific cases, minimising dispatch costs does not minimise customer costs. We do not consider these specific cases as part of this report.*"

4.2. This is a very material oversight. As well as having distributional consequences that are often undesirable, large payments from consumers to generators for not generating significantly reduce the efficiency of the electricity system and the GB economy as a whole. This is for three main reasons:

- First, **generators and large sources of demand have little incentive to site in areas which would benefit GB as a whole** by reducing rather than increasing constraint costs, since they know they will pay or receive the same price in any case (experience suggests that TNUoS charges are unlikely to be strong or granular enough to provide good signals).[48]
- Second, **national wholesale pricing distorts decision-making about transmission build** – with the result that, under the current approach to assessing network investments, network plans are unprecedentedly costly and likely to be politically infeasible.[49]
- Third, following on from the previous two, **the current national pricing regime significantly reduces GB's competitiveness relative to the rest of the world**, as electricity-intensive demand is unable to access low electricity prices in renewable-rich regions. This reduces the attractiveness of GB as a producer of, say, green steel or green hydrogen relative to locations such as the north of Norway or Sweden (both which have zonal pricing). Attracting incremental demand in the right places would lead to a substantial reduction in the current waste of zero carbon electricity via curtailments.

4.3. As a result, reforms to the BM alone are unlikely to allow GB to achieve net zero in line with current targets, as the cost of doing so (both financial and in terms of, for instance, disruption resulting from network build) is likely to be increasingly challenged politically.

---

48 The converse is true for closure decisions too. Namely, generators in export-constrained areas of the system face an unduly strong incentive to remain on the system due to the availability of national prices and compensation payments when it might be preferable for the system that those generators exit the grid (which would free up capacity for newer, likely more efficient, capacity). Equally, generators in import-constrained parts of the country likely face an unduly weak incentive to remain open.

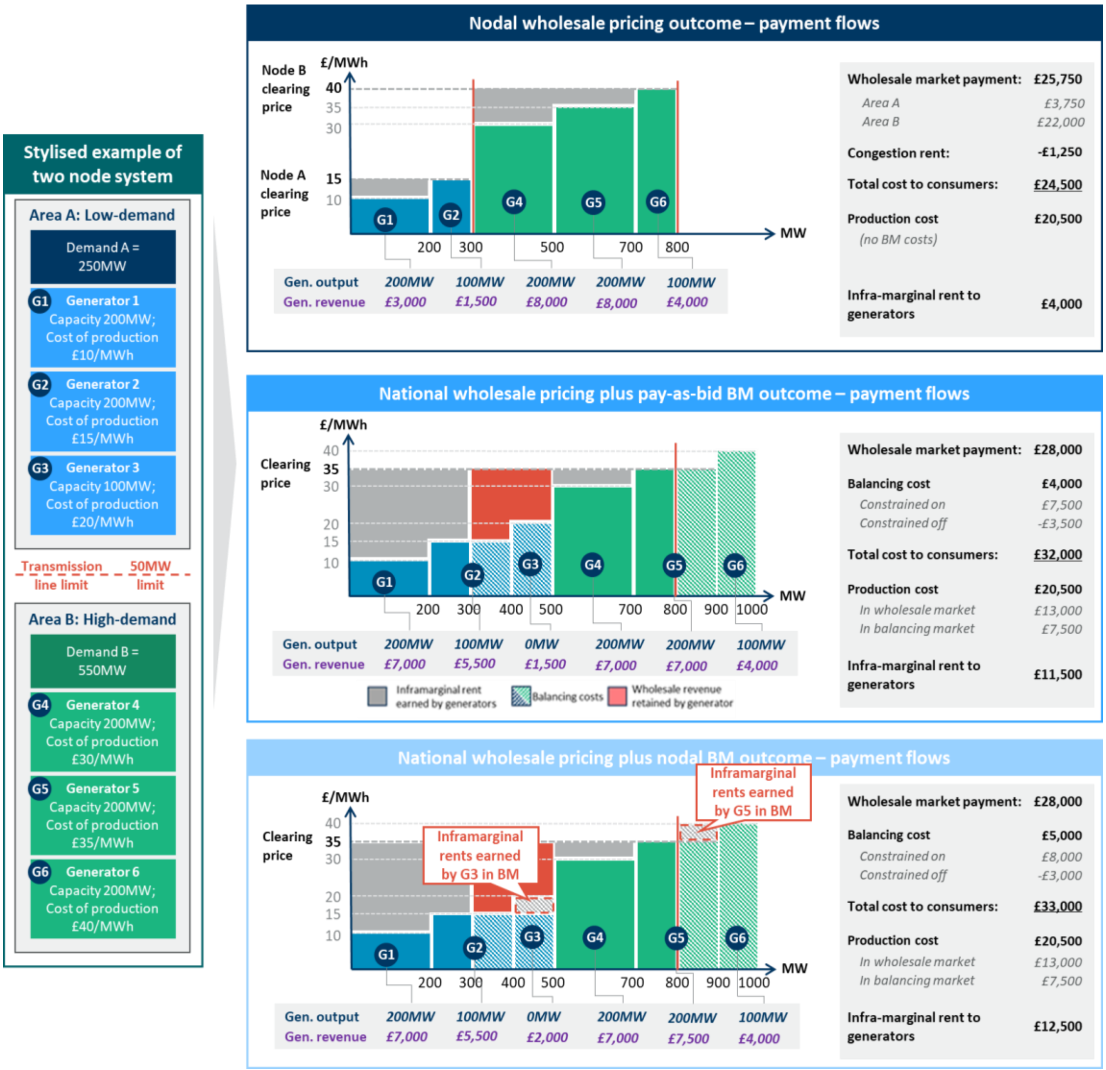
49 Jason Mann, Joe Perkins, and Dan Roberts, 2024. "Transmission investment, flexibility and locational pricing" (link).

Appendix A1: Example of pay-as-cleared BM with a national wholesale market

A1.1 In order to explain the: (i) different distributions of surplus; and (ii) perverse incentives created under a pay-as-cleared nodal BM as suggested in Proposal 3A, we have created a stylised example to illustrate the mechanics at play.

A1.2 The stylised example in Figure A1 consists of a simple two-node system (as set out in the left-hand side of the diagram). We assume for simplicity a one-hour period only and perfect foresight. We start by assuming that market participants bid at their respective marginal costs in the day-ahead market (we relax this assumption in the second part of the example). Demand in this hour is 800MW: 250MW of demand at Node A, and 550MW of demand at Node B. There are three generators at Node A and three at Node B, with limited transmission capacity between the two nodes.

Figure A1: Stylised example of nodal wholesale pricing vs national wholesale pricing plus pay-as-bid BM versus national wholesale pricing plus nodal BM



Source: FTI Consulting

A1.3 The figure shows the market outcomes and payment flows under three market designs: nodal wholesale pricing, national wholesale pricing plus a pay-as-bid BM (status quo), and national wholesale pricing with a nodal, pay-as-cleared BM.[50]

A1.4 Under ideal assumptions (see Figure 2 for why these often do not hold in reality) the final dispatch is the same under all three market designs. However, there are significant differences in consumer costs and producer rents in this example:

- Nodal wholesale pricing results in consumer costs of £24,500 and producer rents of £4,000. Congestion rents of £1,250 are also generated.
- The status quo arrangements result in consumer costs of £32,000 and producer rents of £11,500.
- National wholesale pricing with a nodal (pay-as-cleared) BM results in consumer costs of £33,000 and producer rents of £12,500.

A1.5 The following sections explain how the above figures are calculated.

Market outcome and payment flows under a nodal wholesale market

A1.6 The nodal market respects the transmission network capability in all time scales. In this case, therefore, there is a two-node market and two prices in all timescales. The price is determined by calculating what the cost would be to serve an incremental unit of demand at each node on the system.

A1.7 At the day-ahead stage, market participants bid into the day-ahead market and the SO clears the market. In this case, at Node A, there is 300MW of demand (250MW local demand and 50MW that can be transported to Node B). Hence, G2 is the marginal generator at node A – if demand went to 301MW, the additional unit of consumption could be met by G2. Therefore, the price at Node A is £15. At Node B, the price is £40, as G6 is the marginal generator; an incremental unit of consumption would need to be met by G6.

A1.8 Thus, demand at Node A pays £15/MW for 300MW; G1 receives £15/MW for 200MW and G2 receives £15/MW for 100MW. Demand at Node B pays £40/MW, G4, G5 and G6 all receive £40/MW.

A1.9 This clears the market while respecting the transmission network constraints. There is no need for any further SO actions as we assume there are no energy imbalances.

A1.10 In real time, G1 generates 200MW (and earns £3,000), G2 generates 100MW (and earns £1,500), G4 and G5 generate 200MW (and each earn £8,000) and G5 generates 100 MW (£4,000). In aggregate, generation is paid £24,500.

A1.11 The cost to customers varies depending on where it is sited. Demand at Node A pays £3,750 (£15/MW * 250MW) and Demand at Node B pays £22,000 (£40/MW * 550MW). In aggregate, customers therefore pay £25,750.

[50] In this example, we refer to the two zones as ‘nodes’. We note that, assuming no intra-zonal constraints, the outcomes would be the same under a nodal or zonal wholesale market.

A1.12 A final point: The difference between the amount paid out by customers and the amount received by generators is the congestion surplus / rent. This arises as some of the electricity is sold in the export constrained part of the network at Node A (£15) yet because it is conveyed to Node B on the constrained transmission line, is sold at Node B at £40. Hence (£40 - £15) * 50MW, or £1,250 in congestion rent is generated. This is the genesis of so-called financial transmission rights. Typically, it is passed back to consumers proportional to their load.

Market outcome and payment flows under a national wholesale market

A1.13 The national wholesale market assumes that the transmission network is, in effect, a 'copper plate', and that all sources of generation and demand can freely contract with one another. As above, the price is determined by calculating what the cost would be to serve an incremental unit of demand on the system; but this is determined at a national level. The result is that the SO is required to intervene in the balancing market, such that transmission constraints are not violated in real-time. Under ideal assumptions, the final dispatch (post-BM) is the same as under a nodal market. However, under a pay-as-bid BM and pay-as-cleared BM, the prices paid by the SO (and the underlying incentives) to conduct these corrective actions will differ.

Wholesale market

A1.14 800MW of demand will contract ahead of gate closure with the least cost generators available across both zones, generators G1, G2, G3, G4 and G5. As G5 is the marginal generator, the national day-ahead clearing price is £35/MWh.

SO assessment of system needs

A1.15 At gate closure, generators nominate in line with their respective contracted positions. Hence G1 and G2 nominate 200MW, G3 nominates 100MW, G4 nominates 200MW and G5 nominates 100MW. Demand also nominates its 800MW (250MW at Node A; 550MW at Node B).

A1.16 Having received the nominations, the system operator assesses that the proposed schedule of operation cannot actually meet demand. Even though demand and supply balance in aggregate (800 MW of demand and 800MW of generation scheduled), the transmission constraint means the schedule is, in reality, infeasible. This is because:

- there is too much generation scheduled at Node A (500MW in total) relative to demand at A (of 250MW) and transmission export capacity of 50MW. (i.e., there is 200MW too much generation at Node A); whilst
- the opposite is true at Node B, where demand of 550MW cannot be met by the 300MW scheduled by G4 and G5 and the 50MW that can be imported from Node A (i.e., there is a 200MW shortfall of generation at Node B).

The SO therefore needs to take actions in the BM to resolve the transmission constraint and ensure the schedule is feasible for real time.

SO corrective actions in the BM

A1.17 The SO reduces generation at Node A, by accepting offers to buy from G2 and G3, reflecting their marginal costs of generation. For instance, G3 is indifferent between incurring a cost of £20/MWh to meet its contractual obligation of 100MW (set in the ex-ante market pre gate closure) or paying £20/MWh to absolve itself from its contractual commitment. Hence, the SO sells:

- Under pay-as-bid, 100MW to G3 at £20/MWh, and 100MW to G2 at £15/MWh, reflecting each generator's marginal cost of generation.
- Under a nodal pay-as-cleared BM, both G2 and G3 bid in at £15/MWh and £20/MWh respectively to buy their generation back. Therefore, the clearing price to buy back generation is set by the lowest bid: £15/MWh. The ESO sells 100MW to G2 and 100MW to G3, both at £15/MWh. The lower amount paid by G3 to close out its contractual position, relative to the £20/MWh it received in the wholesale market, leads to inframarginal rents for G3 in the BM.

Therefore, G2 now only has a contractual position, in aggregate, of 100MW, whilst G3 has an aggregate contractual position of 0MW. Total generation in Node A is 300MW (200MW from G1, 100MW from G2), matching the 250MW of demand and 50MW of export capacity.

A1.18 The SO also buys additional electricity at Node B, such that there is sufficient generation to meet the 550MW demand. The ESO buys:

- Under pay-as-bid, 100MW at £35/MWh in the BM from G5 and 100MW at £40/MWh from G6 (as we assume G5 and G6 bid their marginal cost).
- Under a nodal pay-as-cleared BM, 100MW from G5 is bid in at £35/MWh, and 100MW at £40/MWh from G6. The market clearing price is the highest offer: £40/MWh. Hence, despite its lower marginal cost of generation, G5 receives the market-clearing price for additional generation in the BM. As was the case for G3, this additional revenue is inframarginal rent; revenues exceeding the cost of generation.

A1.19 Because of the SO actions in the BM, the planned schedule of generation is now consistent with the transmission network capability. In real time:

- G1 generates 200MW, and earns £7,000, which is total payment from forward sales under both BM market designs;
- G2 generates 100MW, and earns £5,500 under both market designs; £7,000 from forward sales minus £1,500 for buying in the BM;
- G3 generates 0MW, but still earns:
  - Under a pay-as-bid market, £1,500, reflecting £3,500 from forward sales minus £2,000 for buying in the BM; or
  - Under a nodal pay-as-cleared market, £2,000, reflecting £3,500 from forward sales minus £1,500 for buying in the BM.
- G4 generates 200MW, and earns £7,000 under both market designs from forward sales;

- G5 generates 200MW and, like G3, has different revenues under the BM market designs:
  - Under a pay-as-bid market, it earns £7,000, which is total payment from forward sales at £3,500 plus the sale to the SO in the BM for £3,500; or
  - Under a nodal pay-as-cleared market, earns £7,500, which is total payment from forward sales at £3,500 plus the sale to the SO in the BM for £4,000.
- G6 generates 100MW, and earns £4,000 which all comes from sales in the BM.

A1.20 The additional inframarginal rents that accrue to G3 and G5 in the BM is a direct result of introducing a nodal pay-as-cleared BM to replace the current pay-as-bid BM:

- G3 receives an extra £500 ([£20 - £15] * 100), as it has to pay £15/MWh in the BM to not have to generate its contracted 100MW, which would have cost £20/MWh to produce; whilst.
- G5 receives an extra £500 ([£40 - £35] * 100), as it receives £40/MWh in the BM to generate electricity which cost £35/MWh to produce.

Incentives under a national wholesale market combined with a nodal BM

A1.21 In the previous example, we have assumed that all market participants (demand and supply) bid their entire supply and/or demand in the day-ahead market. However, under a pay-as-cleared BM the optimal strategy for the market participants is different, even without considering any manipulations of bids. Namely, a pay-as-cleared BM will disincentivise participation in the wholesale markets for those participants who can go into the BM and pay or receive more favourable BM prices than the prevailing ex-ante national wholesale market price. This potentially distorts wholesale market outcomes. In particular, the demand and generation that helps alleviate constraints is now no longer incentivised to contract ex-ante, but rather wait for the nodal BM to buy or sell electricity, respectively. These incentives are also present under a national wholesale market plus a pay-as-bid BM but would require bids that deviate from their marginal cost.

A1.22 Recall in our previous example the national price was £35. The BM nodal price was either £15/MWh in Area A or £40/MWh in Area B. If we put in place nodal BM prices, then it would follow that the BM price would be £15 or £40 depending on if the participant is located at A or B. Figure A2 summarises the pay-outs in both markets for the different market participants.

Figure A2: Pay-outs in the national wholesale market or nodal pay-as-cleared BM for the different market participants

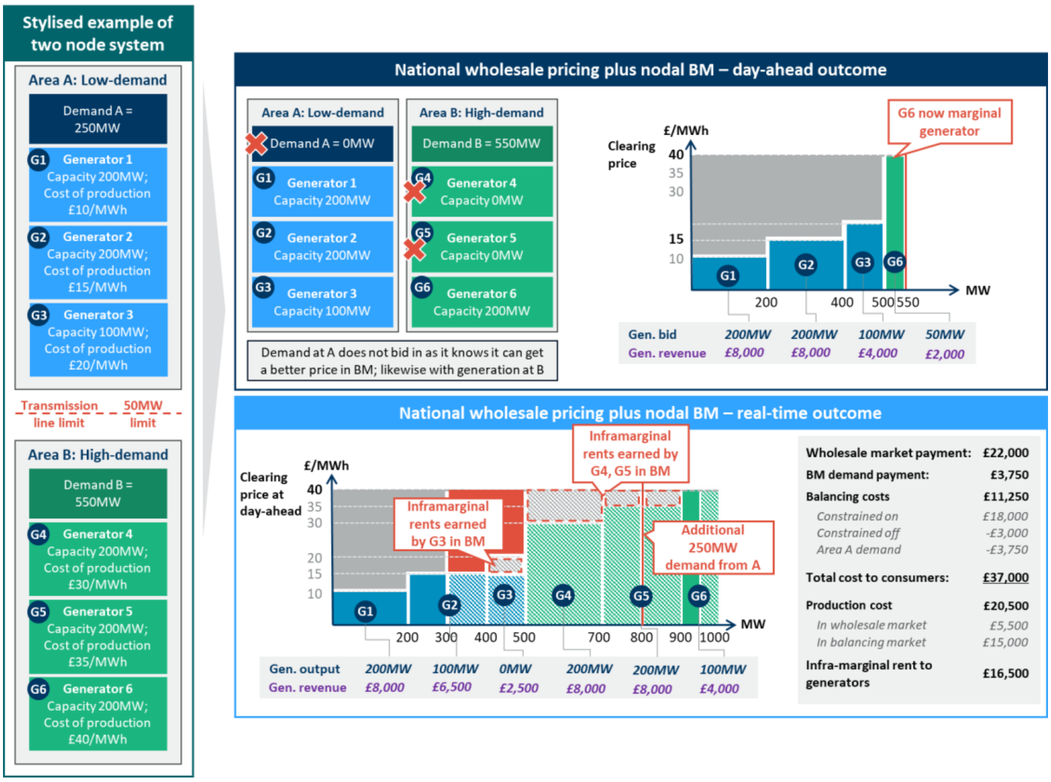
		National wholesale price	Nodal BM price	
Zone A	Demand	£35	£15	Incentive for demand at A to purchase in BM
	Generation	£35	£15	
Zone B	Demand	£35	£40	
	Generation (other than G6)	£35	£40	

Source: FTI Consulting

A1.23 With these incentives, both generation at A and demand at B face a strong incentive to contract ex-ante (rather than to sell at the nodal BM price). But demand at A faces a strong incentive not to contract ex-ante, as it knows if it does not contract at the national wholesale price, it can instead just pay the much lower nodal BM price. Generators 4 and 5, located at B, face an incentive not to contract too. G6 at B is indifferent as it would not be cleared in the national wholesale market and sets the price in the nodal BM market. Taking this to the logical conclusion, a situation appears where: (i) demand at A does not contract ex-ante, and only enters at the BM stage; and (ii) generators 4 and 5 at B also do not contract ex-ante.

A1.24 With this set of distorted incentives in place, Figure A3 updates the pay-as-cleared example in Figure A1, to illustrate the possible market outcome.

Figure A3: Incentives under a national wholesale market with a nodal BM



Source: FTI Consulting

Wholesale market

A1.25 Before gate closure, demand of 550MW at B enters in the national wholesale market while the demand in A is strategic and waits to contract. Total demand is 550MW. Generators 1, 2 and 3 enter the market offering 500MW; this is insufficient to clear the market, and as such, another generator is needed. As G4 and G5 face an incentive not to contract, they hold back. Generator 6 enters the market, and the market clears at £40/MWh. This is not necessarily a good outcome as:

- the national price is higher relative to £35/MWh in the previous example; and
- the thermal constraints are still in place.

Assessment of likely impacts of proposed reforms to the balancing mechanism within a national-price regime

SO assessment of system needs

A1.26 At gate closure, generators nominate in line with their respective contracted positions. Hence G1 and G2 nominate 200MW, G3 nominates 100MW and G6 nominates 50MW. Demand nominates its 550MW (0MW at Node A; 550MW at Node B).

A1.27 Having received the nominations, the system operator assesses that the proposed schedule of operation cannot actually meet demand. Even though demand and supply balance in aggregate (550MW of demand and 550MW of generation scheduled), the transmission constraint means the schedule is, in reality, infeasible. This is because:

- there is too much generation scheduled at Node A (500MW in total) relative to no demand and transmission export capacity of 50MW. (i.e., there is 450MW too much generation at Node A); whilst
- the opposite is true at Node B, where demand of 550MW cannot be met by the 50MW scheduled by G6 and the 50MW that can be imported from Node A (i.e., there is a 450MW shortfall of generation at Node B).

The SO therefore needs to take actions in the BM to resolve the transmission constraint and ensure the schedule is feasible for real time.

SO corrective actions in BM

A1.28 The SO increases demand and reduces generation at Node A. G2 and G3 offer £15/MWh and £20/MWh respectively to buy their generation back. Independent of what the 250MW of demand offers to buy electricity at Node A, if the offer is higher than £15/MWh, the nodal BM clearing price is set at £15/MWh. The ESO sells 250MW to the demand in Node A, buys 100MW from G2 and 100MW from G3. The demand at Node A pays in total £3,750 in the BM.

A1.29 Therefore, the demand in Node A now has a contractual position of 250MW, G2 now only has a contractual position, in aggregate, of 100MW, whilst G3 has an aggregate contractual position of 0MW. Total generation in Node A is 300MW (200MW from G1, 100MW from G2), matching the 250MW of demand and 50MW of export capacity.

A1.30 The SO also buys additional 450MW of generation at Node B, such that there is sufficient generation to meet the 550MW demand. G4 bids in 200MW at £30/MWh, G5 bids in 200MW at £35/MWh, and G6 bids in 150MW at £40/MWh. The market clearing price is the highest offer to clear the total required generation to be bought in the BM of 450MW: £40/MWh.

A1.31 G3 now has a contractual position of 200MW, G4 as well of 200MW and G6, in aggregate, of 100MW.

A1.32 Because of the SO actions in the BM, the planned schedule of generation is now consistent with the transmission network capability. In real time:

- G1 generates 200MW, and earns £8,000, which is total payment from forward sales;
- G2 generates 100MW, and earns £6,500; £8,000 from forward sales minus £1,500 for buying in the BM;

- G3 generates 0MW, but still earns £2,500; £4,000 from forward sales minus £1,500 for buying in the BM;
- G4 generates 200MW, and earns £8,000 from sales in the BM;
- G5 generates 200MW and also earns £8,000 from sales in the BM; and
- G6 generates 100MW, and earns £4,000; £2,000 from forward sales and £2,000 from sales in the BM.

A1.33 The result is a higher cost to consumers, as payments to generators are now £37,000 compared to a national wholesale market plus pay-as-bid BM (assuming no manipulation of bids) total payment of £32,000 and nodal market outcome of £24,500.

A1.34 The total costs for the SO to buy electricity in the BM is £18,000 while only £6,750 is received by the SO from selling electricity. This implies that Balancing Services Use of System ("BSUoS") costs are therefore £11,250, which is proportionally recovered from demand. The overall outcome is that:

- Demand at A pays £0 in the wholesale market and £3,750 in the BM plus its share of BSUoS of £3,516, a total of £7,266; Demand at B pays £22,000 in the wholesale market plus its share of BSUoS of £7,734, a total of £29,734.
- An even greater volume of inframarginal rents (at the expense of consumers) accrue to generators, due to:
  - a higher national price in the wholesale market leading to greater revenues for generators in Area A; and
  - generators 4 and 5 contracting in the BM exclusively, where they receive the full BM price of £40/MWh for their generation, rather than the national price, which would be £35/MWh if they were to enter the wholesale market and contract ex-ante.

DISCLAIMER

This report has been prepared by FTI Consulting LLP ("FTI") for Octopus Energy in connection with a critical review of alternative Balancing Market arrangements under the terms of Octopus Energy's engagement letter with FTI dated 10 June 2024 (the "Contract").

This report has been prepared solely for the benefit of Octopus Energy in connection with a critical review of alternative Balancing Market arrangements and no other party is entitled to rely on it for any purpose whatsoever.

This report is not to be referred to or quoted, in whole or in part, in any registration statement, prospectus, public filing, loan agreement, or other agreement or any other document, or used in any legal, arbitral or regulatory proceedings without the prior written approval of FTI. FTI accepts no liability or duty of care to any person (except to Octopus Energy under the relevant terms of the Contract) for the content of the report. Accordingly, FTI disclaims all responsibility for the consequences of any person acting or refraining to act in reliance on the report or for any decisions made or not made which are based upon such report.

The report contains information obtained or derived from a variety of sources. FTI does not accept any responsibility for verifying or establishing the reliability of those sources or verifying the information so provided.

Nothing in this material constitutes investment, legal, accounting or tax advice, or a representation that any investment or strategy is suitable or appropriate to the recipient's individual circumstances, or otherwise constitutes a personal recommendation.

No representation or warranty of any kind (whether express or implied) is given by FTI to any person as to the accuracy or completeness of the report.

The report is based on information available to FTI at the time of writing of the report and does not take into account any new information which becomes known to us after the date of the report. We accept no responsibility for updating the report or informing any recipient of the report of any such new information.