Report for ESO

Investigation into sequential and co-optimised procurement of energy, reserve and response services

Qualitative assessment

23 July 2024

Disclaimer

This report has been prepared by FTI Consulting LLP (“FTI”) for National Grid ESO as part of a Research and Development Collaboration Agreement entitled “Exploring the economic benefits of co-optimising procurement of energy, response and reserve” under the terms of the Agreement dated 05 October 2023 (the “Contract”).

This report has been prepared solely for the benefit of National Grid ESO in connection with this project and relevant results can be disseminated through the industry (as defined under the relevant terms of the Contract), but 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 National Grid ESO 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, completeness, reliability or adequacy of the report or of the information provided therein.

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.

All copyright and other proprietary rights in the report remain the property of FTI and all rights are reserved.

Copyright Notice

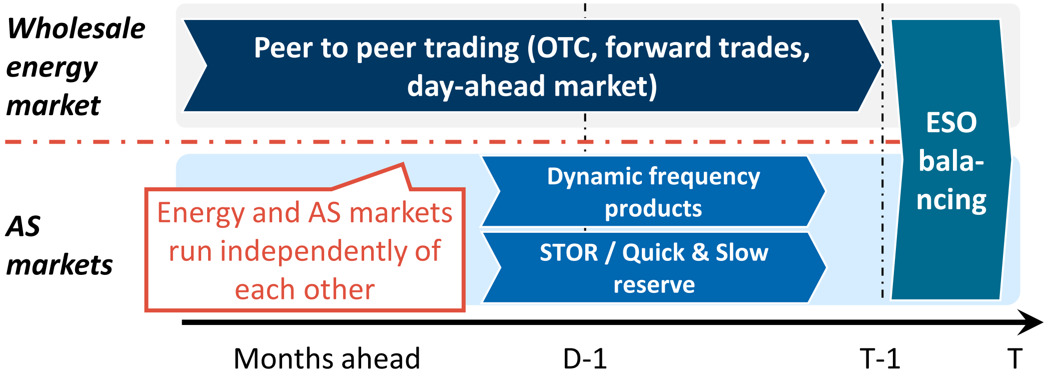
© 2024 FTI Consulting LLP. All rights reserved.

Background and context

The current GB electricity market design features separately operating energy and ancillary services markets, despite inherent links between the two

Wholesale ("WS") and ancillary service ("AS") markets in GB currently operate independently of each other

Simplified current GB design – without co-optimisation



Wholesale energy market

Peer to peer trading (OTC, forward trades, day-ahead market)

ESO balancing

AS markets

Energy and AS markets run independently of each other

Dynamic frequency products

STOR / Quick & Slow reserve

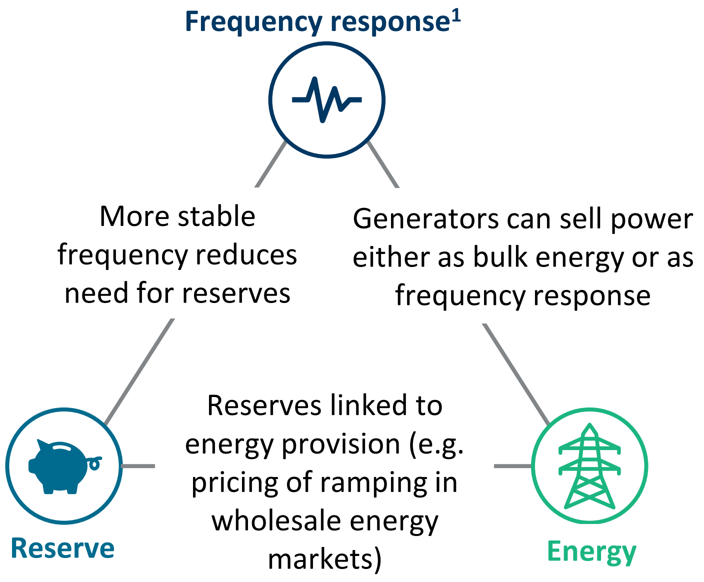
Months ahead

D-1

T-1

T

However, there are close interrelationships between energy, frequency response, and reserve markets



Frequency response[1]

More stable frequency reduces need for reserves

Generators can sell power either as bulk energy or as frequency response

Reserves linked to energy provision (e.g. pricing of ramping in wholesale energy markets)

Reserve

Energy

Markets operate independently and close sequentially, implying that:

1) Resources may need to **choose which markets to participate in**;
2) Resources' bids reflect the **estimated opportunity costs** of participating in the selected market (as by virtue of providing a service in one market assets are not able to provide another service); and
3) The **need to form such expectations** introduces **significant information imperfections**, with potential **adverse impacts on the price formation process** and resulting market outcomes.

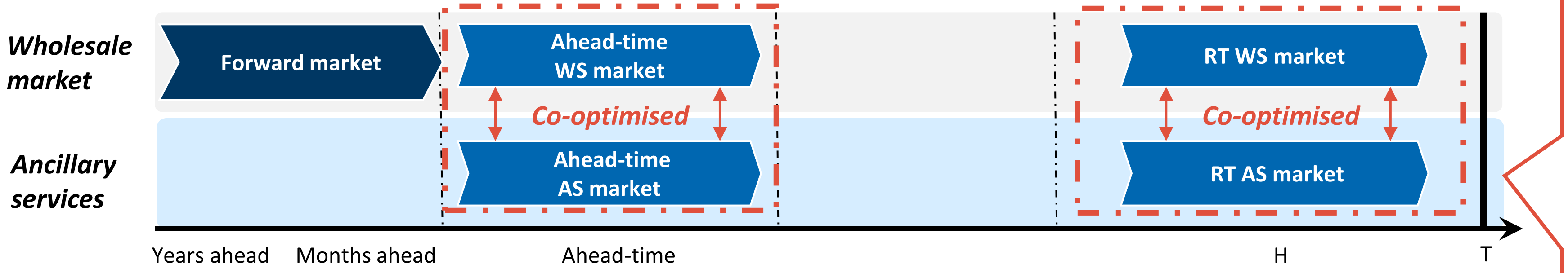
Market participants must first **forecast day-ahead ("DA") WS energy prices** and their resulting production schedule to understand their own **opportunity cost** of providing AS. Forecasts are complicated by **uncertainty** on the final energy demand, renewable production, possible outages etc., as well as the fact that the bidding of other market participants, and hence final market outcomes, will be dependent upon their opportunity cost estimations.

Notes: (1) We use "frequency response" or simply "response" interchangeably in this report.

ESO wishes to investigate whether there is a case for systematically considering these interactions when scheduling energy and procuring AS

- ESO wishes to investigate the potential merits of **co-optimising the scheduling of energy (via WS markets) and the procurement of AS.**[1] This is a process where a **single optimisation process** identifies the optimal outcome across **both WS and AS markets** simultaneously.
- The expectation is that this process would lead to, among other results, a **more efficient allocation of resources** based on the relative needs and conditions of each market and an **efficient formation of prices** across each market. This could be particularly important in future years as new AS markets are launched and AS requirements increase, increasing the potential efficiency gains of co-optimisation.

Simplified schematic of a hypothetical co-optimised market design (with central scheduling[2] and an ahead-time market[3])



This assessment explores the impact of WS and AS markets operating in an integrated, co-optimised manner.

We evaluate the impact of removing the need for resources to choose which markets to participate in and/or perform complex estimations of the **opportunity costs of participating in different markets**.

- In this report we focus on the potential co-optimisation of WS energy and AS markets. Alternative degrees of integration between markets are possible, which we discuss in more depth in the terminology section in the quantitative report.
- Co-optimisation can be performed at forward timeframes and/or real time ("RT"). In this report, we **refer to the forward time frame as 'day-ahead'** because:
  - Currently the reference WS market timeframe in GB is the DA market;
  - Currently US markets co-optimise the WS and AS markets at DA stage.
- However, in the future, co-optimisation could also be performed at the **intraday timeframe** augmenting and/or replacing co-optimisation at the DA timeframe, as the importance of renewable generation forecasting overtakes the long-duration commitment decisions of thermal units.

Notes: (1) ESO has started examining these issues, at a high level, in its recent work on Net Zero Market Reform (link); (2) The merits of moving from GB's current self-scheduled market design to a centrally scheduled one are not examined in this report, for more discussion see slide 64; (3) Here we provide a simplified hypothetical market design, in practice there would be the possibility to introduce WS markets at multiple timeframes e.g., at day-ahead WS markets preceding multiple intraday WS auctions.

The ESO has commissioned FTI to undertake a qualitative and a quantitative assessment of the co-optimisation of energy and AS

- FTI has prepared two reports, published in parallel, setting out our approach, assessment and key findings regarding the sequential and co-optimised procurement of energy, reserve and response services.
- Our assessment comprises the following elements:

Report for ESO

Investigation into sequential and co-optimised procurement of energy, reserve and response services

Qualitative assessment

nationalgridESO

23 July 2024

Report for ESO

Investigation into sequential and co-optimised procurement of energy, reserve and response services

Quantitative assessment

nationalgridESO

23 July 2024

1. Qualitative assessment of the pros and cons of co-optimising energy and AS in GB;

This report focuses on the qualitative assessment. We describe, qualitatively, the benefits (e.g. **more efficient price signals** and **more efficient allocation of resources** across markets) as well as **potential challenges and risks** arising from a potential co-optimisation of energy and AS in the GB context.

2. Quantitative assessment of the historical cost savings of co-optimised day-ahead and/or real time procurement of energy and AS in GB;
3. Quantitative assessment of the potential efficiency savings from the co-optimised procurement of energy and AS in GB from 2025 to 2035;
4. Quantitative assessment of the potential efficiency savings from co-optimisation in a locational market design.

This report focuses on the **quantitative** assessment of potential impacts of co-optimisation of energy and AS (specifically reserve and response) in GB.

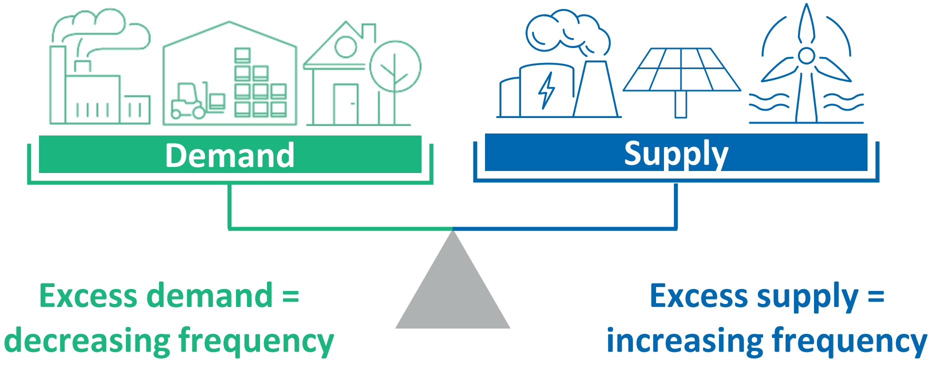
Table of contents

Introduction to GB energy and ancillary service markets

Electricity systems must continuously remain in balance, which can be achieved via two broad type of market design – self or centralised scheduling

Electricity systems must continuously remain in balance

- **All electricity systems must balance supply and demand** on a second-by-second basis at all locations on the grid
- Any **imbalance results in fluctuations in system frequency,** which, if great enough, can exceed system tolerances, resulting in emergency disconnections and, *in extremis*, system-wide blackouts
- As such, **electricity system operators procure a range of services** to manage unexpected imbalances (e.g., following a generator failure or a spike in demand) and to maintain security and quality of supply across the electricity system. These services are known as ancillary services (see slide 10).



Demand

Supply

Excess demand = decreasing frequency

Excess supply = increasing frequency

There are two broad types of market designs to coordinate a balanced electricity system

- Broadly, there are two key market design options to coordinate a balanced electricity system: a **self scheduling market** or a **centralised scheduling market,** both of which are deployed in different jurisdictions around the world.[1]
- Under each design, the roles market participants and the electricity system operator ("ESO") complement each other:

1. Self scheduling

- **Market participants:** Determine their own intended operating schedules, which must be finalised and submitted to the ESO shortly before delivery (e.g. an hour before) – referred to as 'gate closure'
- **ESO:** After gate closure, the ESO takes control of the system, altering operating schedules and dispatching plants (potentially differently to how market participants have self-scheduled) to ensure system constraints and operational limits are respected.

2. Centralised scheduling

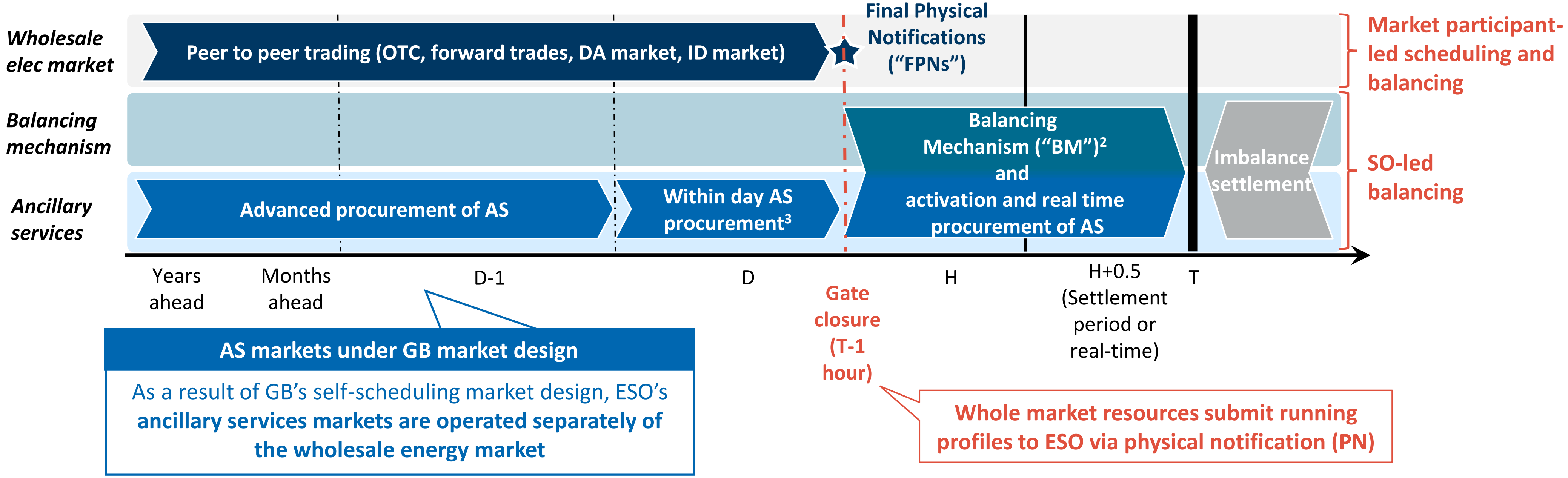
- **Market participants:** Submit bids and offers to provide / consume power at a given price into a centralised market run by the ESO ahead of real time (e.g., a day ahead of delivery), alongside a range of technical parameters (e.g., minimum stable load, ramp capacity)
- **ESO:** Identifies the optimal schedule based on market participant bids, system constraints and operational limits. In real time, can undertake balancing actions and dispatch plants as system conditions require.

Note: (1) In US ISO systems that are centrally dispatched market parties have the option to self-schedule, i.e., they submit an operation schedule to the system operator and become a price-taker. See e.g., PJM (2023) PJM Manual 11: Energy & Ancillary Services Market Operations. Revision: 125. Effective Date: May 18, 2023 (link)

The current GB electricity market is based on self scheduling and was designed to incentivise market participants to balance the system…

- NETA, the basis for GB's current electricity market design, was implemented in 2001 at a time when the system was dominated by a **relatively small number of large fossil fuel generators**
- The current GB design is a self-scheduling market, based on bilateral trading between generators and purchasers of energy with a single national price, designed to **facilitate competition and lead to price discovery**…
- … while also **incentivising market participants to balance the system** over time through the use of imbalance penalties
- ESO was set up to fulfil the role of a **"residual balancer"**, only acting to ensure that transmission constraints and operating limits are respected post gate closure (which occurs one hour before delivery), in line with agreed Security and Quality of Supply Standards (SQSS)

Simplified schematic of the GB electricity market design



AS markets under GB market design

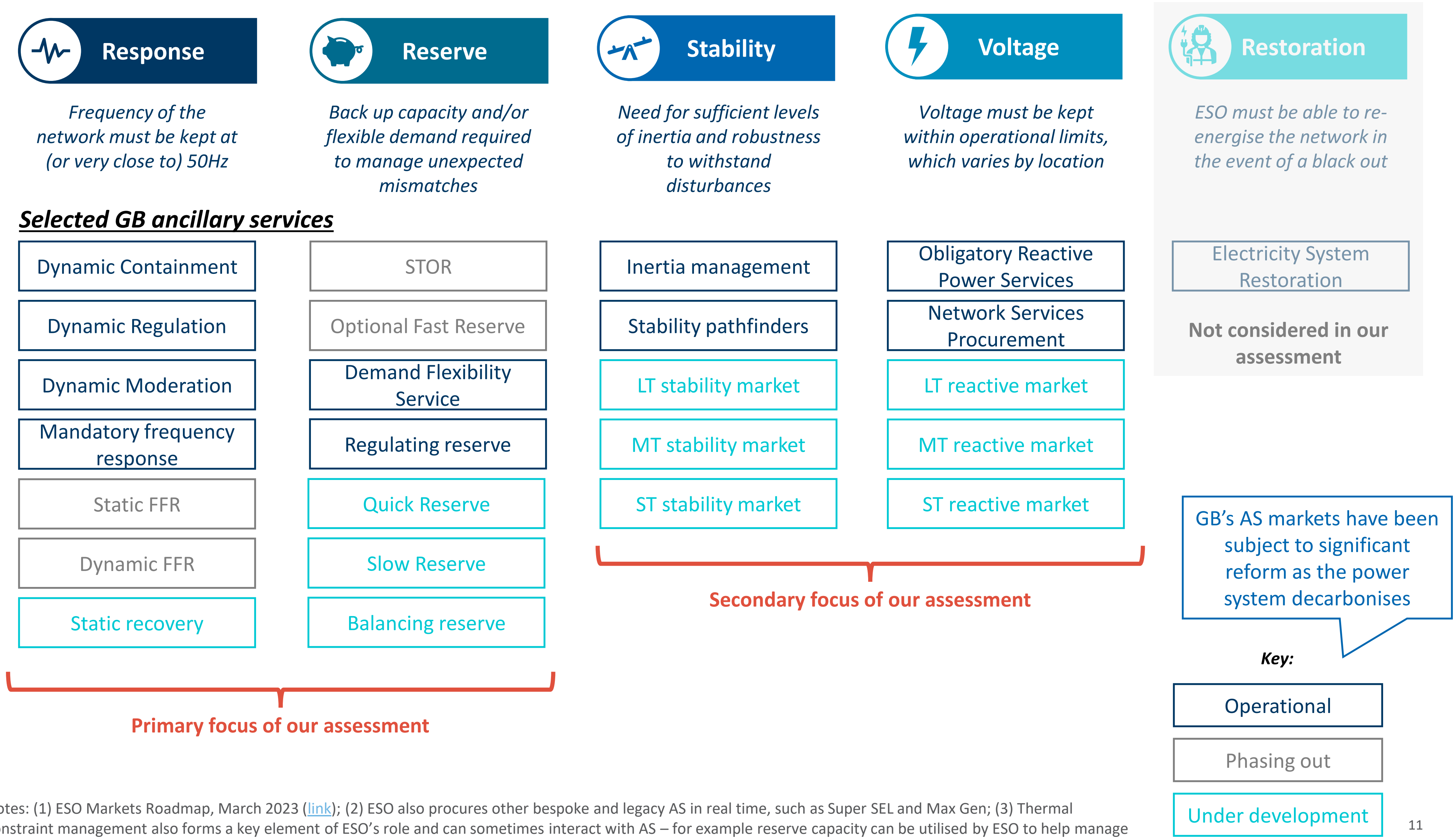
As a result of GB's self-scheduling market design, ESO's **ancillary services markets are operated separately of the wholesale energy market**

Whole market resources submit running profiles to ESO via physical notification (PN)

… with the ESO as the "residual balancer" of the system post gate-closure

Notes: (1) For simplicity, we have omitted discussion around the Capacity Market, and the interplay with other mechanisms such as CfDs. ESO is also able to undertake trades actions prior to gate closure if it considers these actions to be lower cost than via the BM; (2) The BM operates on a "pay-as-bid" principle, where market participants submit bids/offers to ESO reflecting the prices they are willing to accept to modify their electricity production/consumption. The ESO then selects bids and offers in a merit order to ensure that system balanced in real-time while maintaining the stability and security of the system; (3) An example of intraday procurement is Demand Side Flexibility during winter (link).

ESO procures a range of different ancillary services, each meeting a specific system need, independently of the wholesale energy market



Response	Reserve	Stability	Voltage	Restoration
Frequency of the network must be kept at (or very close to) 50Hz	*Back up capacity and/or flexible demand required to manage unexpected mismatches*	*Need for sufficient levels of inertia and robustness to withstand disturbances*	*Voltage must be kept within operational limits, which varies by location*	*ESO must be able to re-energise the network in the event of a black out*

Selected GB ancillary services

Response	Reserve	Stability	Voltage	Restoration
Dynamic Containment	STOR	Inertia management	Obligatory Reactive Power Services	Electricity System Restoration
Dynamic Regulation	Optional Fast Reserve	Stability pathfinders	Network Services Procurement	**Not considered in our assessment**
Dynamic Moderation	Demand Flexibility Service	LT stability market	LT reactive market	
Mandatory frequency response	Regulating reserve	MT stability market	MT reactive market	
Static FFR	Quick Reserve	ST stability market	ST reactive market	
Dynamic FFR	Slow Reserve			
Static recovery	Balancing reserve			

Primary focus of our assessment (Response, Reserve)

Secondary focus of our assessment (Stability, Voltage)

GB's AS markets have been subject to significant reform as the power system decarbonises

Key:

Operational

Phasing out

Under development

Notes: (1) ESO Markets Roadmap, March 2023 (link); (2) ESO also procures other bespoke and legacy AS in real time, such as Super SEL and Max Gen; (3) Thermal constraint management also forms a key element of ESO's role and can sometimes interact with AS – for example reserve capacity can be utilised by ESO to help manage constraints in real time.

AS are procured by ESO over different timeframes and via different mechanisms, including day ahead auctions and the BM

1

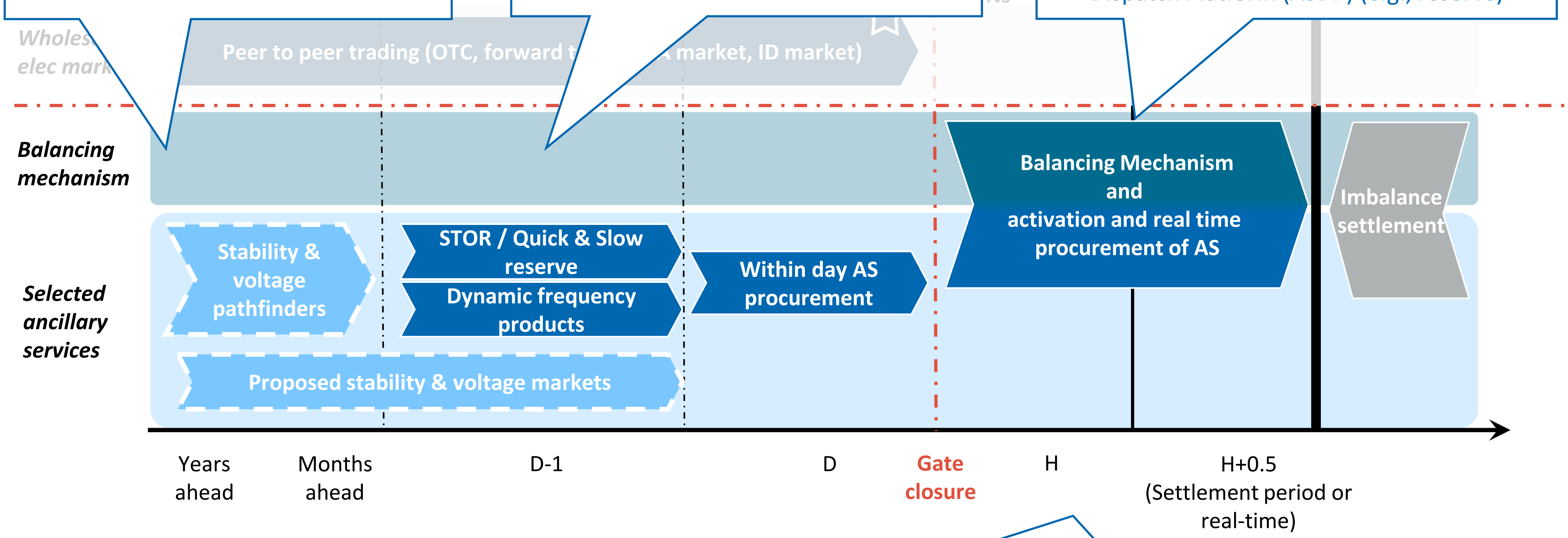
Ancillary service markets have in the past operated up to several years ahead, offering long term contracts to service providers…

2

…but there is a drive to shift procurement to the day ahead and intraday stage, procuring services via competitive auctions

3

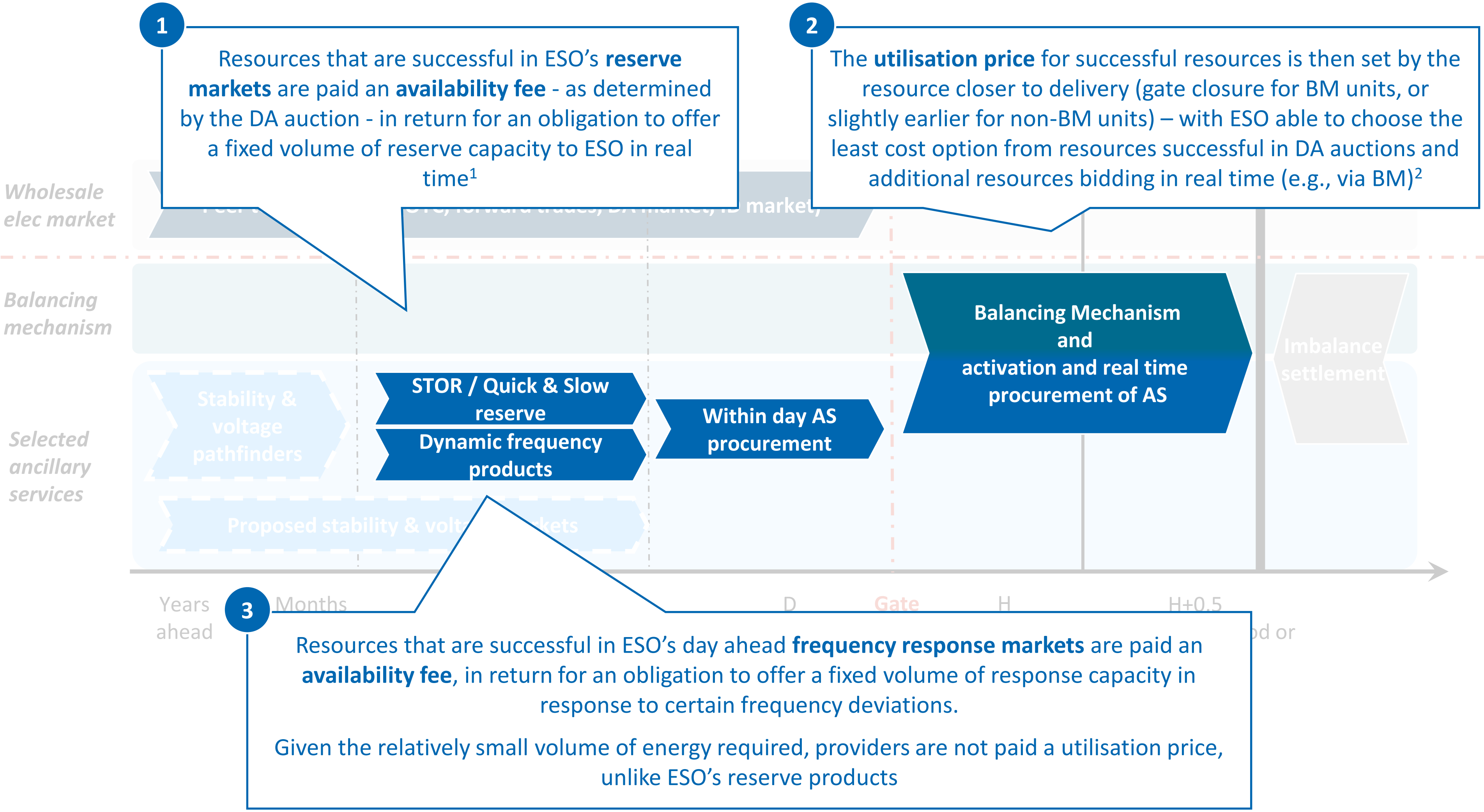
When a system need arises, pre-procured AS are either automatically activated (e.g., frequency response) or manually deployed by the ESO via the BM or the Ancillary Service Dispatch Platform (ASDP) (e.g., reserve)



4

In addition to pre-procured AS, ESO is able to utilise the BM (and trades ahead of gate closure) to re-dispatch resources to provide AS in real time, with the overall objective of utilising the least-cost combination of pre-procured and real time services.
For example, ESO can use the BM to ensure there is sufficient inertia on the system.

For reserve and response, DA markets are used to ensure sufficient provision in real time, however reserve utilisation prices are not fixed until gate closure



1 Resources that are successful in ESO's **reserve markets** are paid an **availability fee** - as determined by the DA auction - in return for an obligation to offer a fixed volume of reserve capacity to ESO in real time[1]

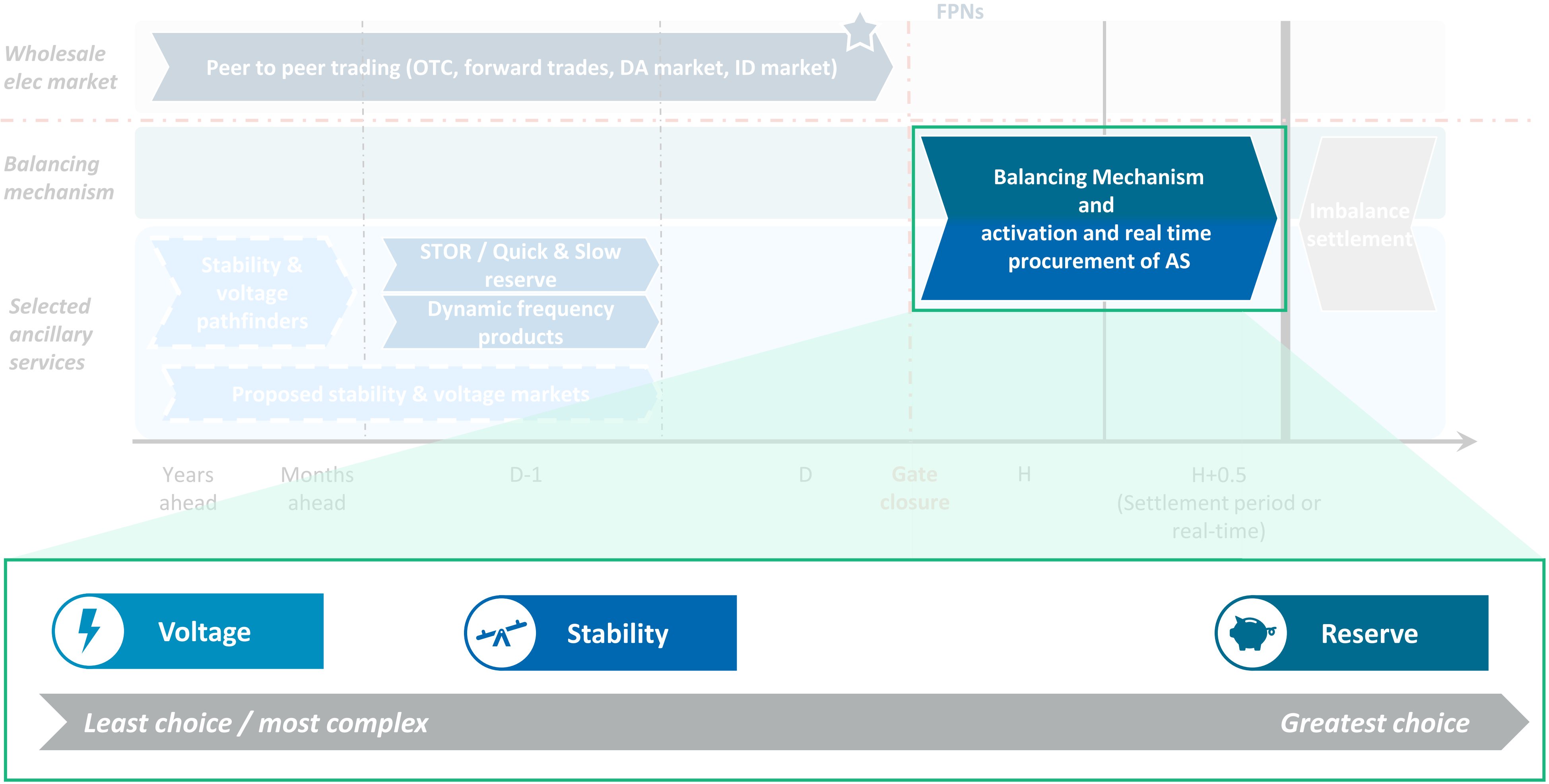
2 The **utilisation price** for successful resources is then set by the resource closer to delivery (gate closure for BM units, or slightly earlier for non-BM units) – with ESO able to choose the least cost option from resources successful in DA auctions and additional resources bidding in real time (e.g., via BM)[2]

3 Resources that are successful in ESO's day ahead **frequency response markets** are paid an **availability fee**, in return for an obligation to offer a fixed volume of response capacity in response to certain frequency deviations.

Given the relatively small volume of energy required, providers are not paid a utilisation price, unlike ESO's reserve products

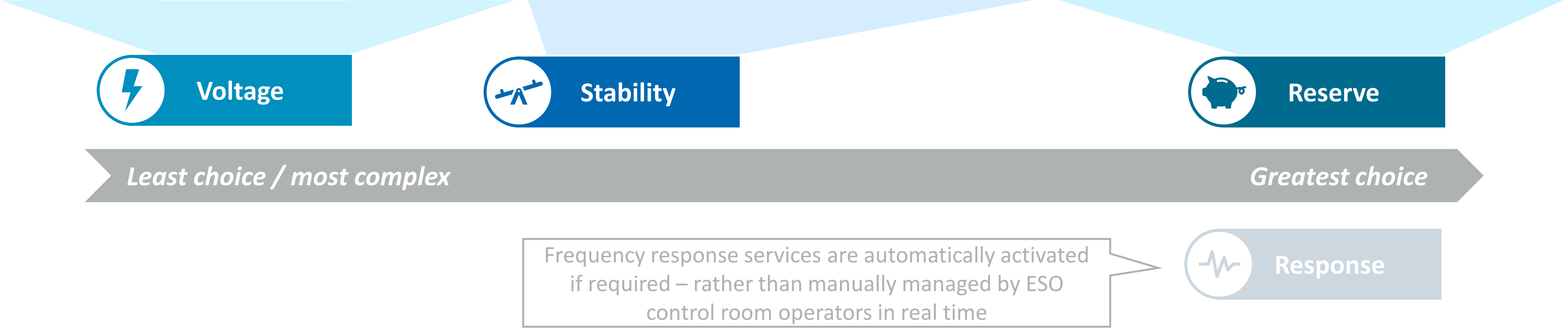
Notes: (1) This report examines, further down, how this commitment to provide reserve impacts (i) participation in other markets and (ii) pricing of different services, particularly in terms of the opportunity costs involved; (2) If additional frequency response services are needed in real time, ESO can call on Mandatory Frequency Response, which is obligatory for certain (typically large) generators.

AS evaluated in this report can be ordered based on the choices available to ESO to find adequate resources to meet the identified system need



Frequency response services are automatically activated if required – rather than manually activated by ESO control room operators in real time

As real time approaches, the level of competition between suppliers and the choice of resources available to ESO varies significantly by system need



- No market-based procurement (outside of recent pathfinders)
  - Real-time, short-term procurement via trades and the BM
  - Provision is obligatory for some resources (e.g., through ORPS)
  - ESO exploring introducing voltage markets in future
- Can only be procured from a limited range of technologies (predominantly thermal synchronous generation and transmission network assets)[1]
- Highly locational and dependent on prevailing system conditions, which further reduces the pool of potential suppliers with the ability to address any shortfalls or surpluses

Voltage

- No market-based procurement (outside of recent pathfinders)
  - Real-time, short-term procurement via trades and the BM only
  - ESO exploring introducing stability markets in future
- Can only be procured from a limited range of technologies (predominantly thermal synchronous generation)[2]
- Inertia is non-locational,[3] meaning the pool of potential providers is not locationally limited (subject to transmission constraints)
- Technically challenging to measure and hard to specify exact needs

Stability

- Market-based procurement via daily auctions is well established and understood (recently reformed, however)
  - Procured at day-ahead stage to provide certainty of real-time supply
- Can be procured from a wide range of technologies with numerous providers
- Non-locational (subject to transmission constraints) – there exists a large pool of suppliers to choose from to address shortfalls.

Reserve

Least choice / most complex

Greatest choice

Frequency response services are automatically activated if required – rather than manually managed by ESO control room operators in real time

Response

Notes: (1) New technology is under development and testing that may allow wind farms, as well as interconnectors, to provide reactive power support. Synchronous condensers can also be used; (2) Resources connected via grid-forming inverters can also provide stability services. (3) Some other aspects of system stability (for example short circuit levels) are more locational, however.

ESO therefore typically ensures system needs are met by starting with those with least choice, before moving to more competitive services

ESO balancing operations approaching real time

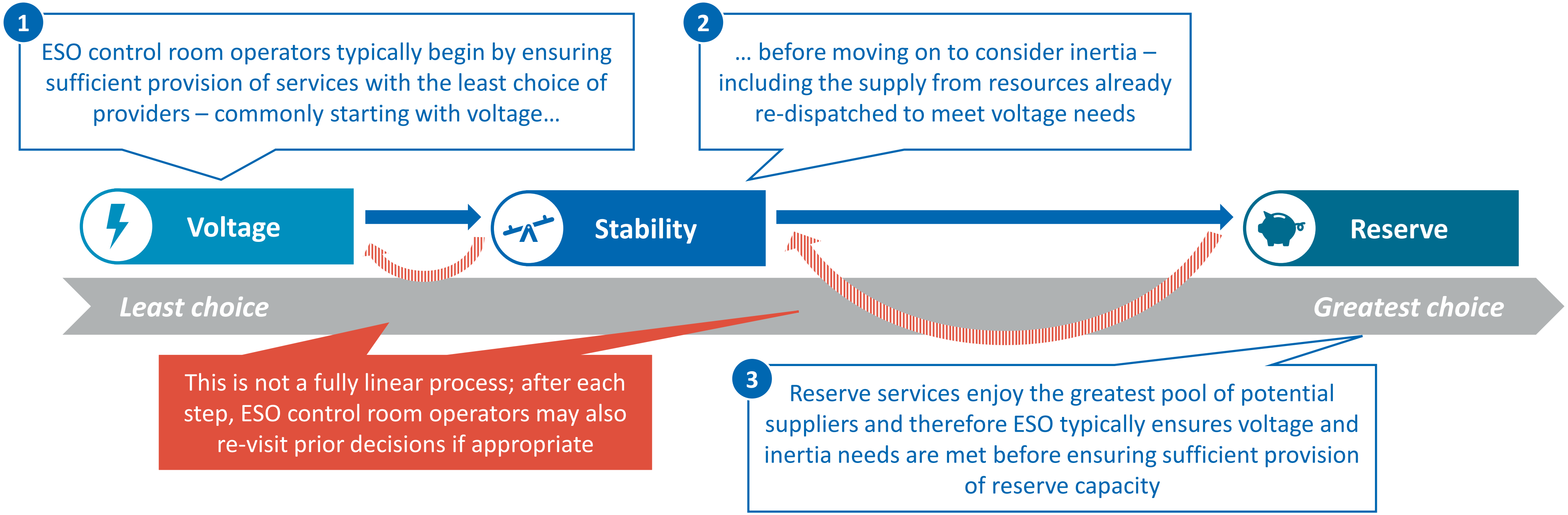
Using the information provided to ESO from the wholesale market (via PNs), the ESO control room develops a System Operating Plan (SOP) and then executes interventions to ensure:

- transmission constraints are adhered to;
- energy supply and demand are balanced; and
- sufficient provision of AS.

Transmission constraint management

- The GB wholesale energy market is a **national market**, operating **without consideration of transmission constraints** across the network.
- **Ensuring these constraints are adhered to in the dispatch** is a critical aspect of ESO's role. Both reserve/response management and transmission constraint management involve ongoing optimisation to balance costs, system security, and efficiency. Actions to manage one aspect can have implications for the other, **requiring an integrated approach**.
- However, constraint management is not typically considered an ancillary service and therefore is not a focus of our report - beyond recognising the need for ESO also to consider constraints when procuring and activating AS.

Stylised timeline of ESO actions to manage AS approaching real time[1]



1. ESO control room operators typically begin by ensuring sufficient provision of services with the least choice of providers – commonly starting with voltage…

2. … before moving on to consider inertia – including the supply from resources already re-dispatched to meet voltage needs

3. Reserve services enjoy the greatest pool of potential suppliers and therefore ESO typically ensures voltage and inertia needs are met before ensuring sufficient provision of reserve capacity

This is not a fully linear process; after each step, ESO control room operators may also re-visit prior decisions if appropriate

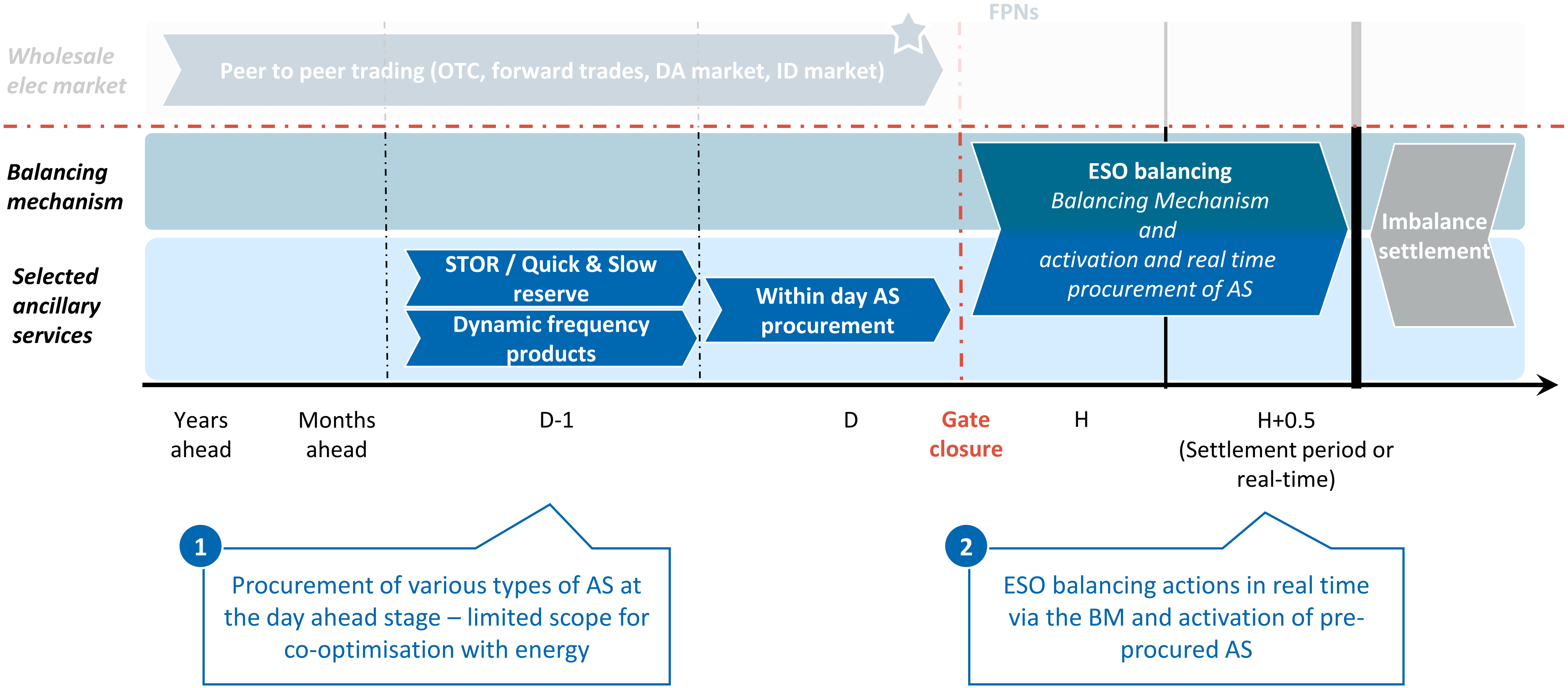
Notes: (1) ESO needs to maintain contingency capacity which is in a state to be capable of synchronising in sufficient time to achieve full output for the relevant peaks/throughs. This requirement usually falls 0MW at 4h ahead of RT as uncertainty regarding system conditions reduces. (2) Pre-gate closure, ESO can also utilise trades to manage the system.

To what extent are GB ancillary services and energy already co-optimised?

The self-schedule nature of GB's current market design limits the potential for full co-optimisation, however some limited scope for optimisation exists

Scope for co-optimisation (or co-ordination) under the current market design

- Due to the separation and sequential nature of the wholesale energy market and the AS markets under GB's self-dispatch market design, there is **limited scope[1] for co-optimisation of energy and ancillary services pre-gate closure**
- However, there are two windows where ESO could – at least partially – co-ordinate the procurement and activation of different services:



Notes: (1) See Appendix 1 for a discussion on why the scope for co-optimisation of energy and AS markets is limited under a self-scheduling design.

At day ahead, ESO is implementing the ability to co-optimise procurement of reserves and response – however this does not extend to real time activation

ESO previously identified significant challenges driven by failure to co-optimise DA procurement of response and reserves…

- Significant co-dependencies between reserve and response products are not currently fully reflected in the market design
- ESO identified in particular that *"Providers blind guess where other parties will tender - risk of oversupplying a single market with others left undersupplied"*…
- …highlighting the risk of imperfect information resulting from the need for market participants to form expectations regarding siloed markets.

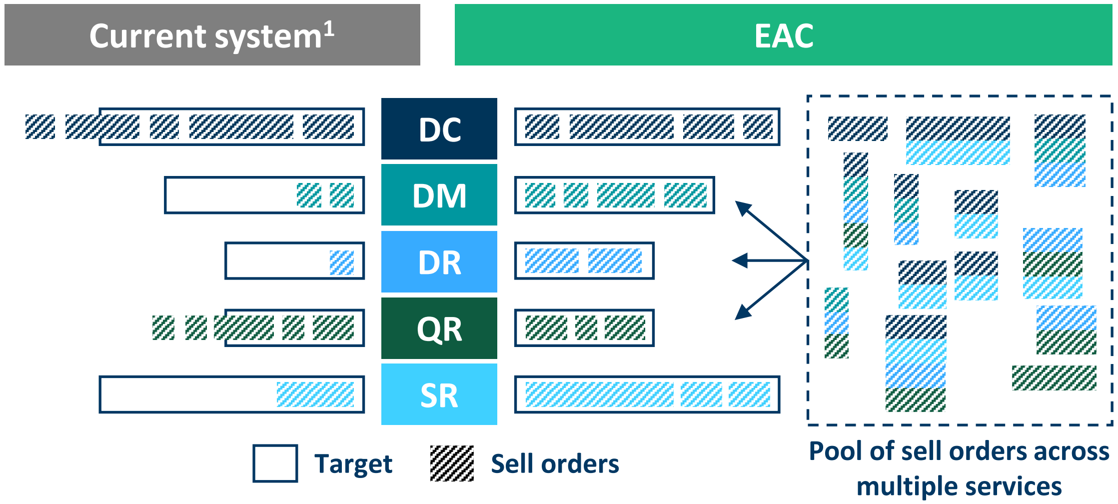
…and this has motivated the introduction of co-optimisation between the two ancillary services

Expected benefits of EAC

- **Reduced risk for market participants**
- **Increase in overall market liquidity**
- **Reduction in procurement costs**

Stylised illustration of current AS market design vs EAC

By requiring participants to bid for a single market, the current system risks inefficient allocations of limited sell orders

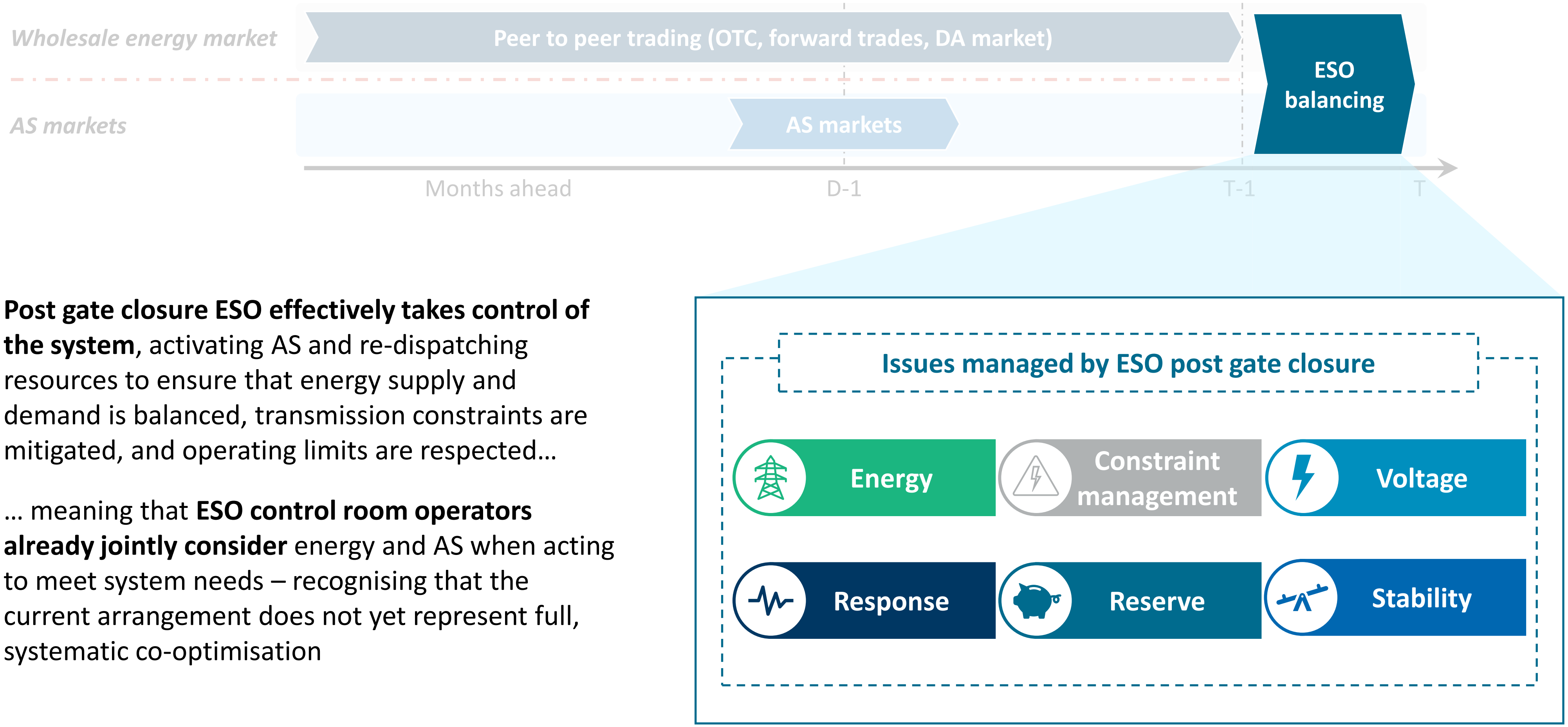


Allowing participants to stack their bids – i.e., submit linked (or mutually exclusive) bids into multiple markets – lets ESO **allocate** them such that markets are cleared **efficiently**

Notes: (1) The introduction of quick and slow reserve are not yet operational (instead STOR is still currently in place).

In real time, ESO can jointly consider some elements of the utilisation of energy and AS via the BM...

Scope for (partial) co-optimisation of energy and AS in real time under the current market design



- **Post gate closure ESO effectively takes control of the system**, activating AS and re-dispatching resources to ensure that energy supply and demand is balanced, transmission constraints are mitigated, and operating limits are respected…
- … meaning that **ESO control room operators already jointly consider** energy and AS when acting to meet system needs – recognising that the current arrangement does not yet represent full, systematic co-optimisation

… however, in practice, co-optimisation of energy and AS via the BM faces several key challenges

Key challenges facing co-optimisation of energy and AS in real time under current GB market design

1

Self-schedule wholesale energy market as starting point

- Under current GB market design, ESO is required to take the schedule as determined via the WS energy market as its starting point for each dispatch interval…
- … then make the lowest cost combination of interventions to this schedule to meet system needs – in line with its residual balancer role
- While these interventions can in theory be optimised, **the resulting dispatch will not necessarily result in a system-wide optimal dispatch** - as would be the case under a centralised dispatch with markets co-optimised at the DA stage – due to the initial inefficiencies in the original schedule determined in the wholesale market
- Furthermore, these **ESO balancing actions to manage transmission constraints will result in balancing costs**

2

Sequential approach to meeting system needs

- Following gate closure, **control room operators typically solve system management issues sequentially**...
- … starting from those services that are most challenging to manage (e.g., where ESO has limited choice of resources to call on)
- **Sequential decision making is inconsistent with co-optimising across different services**, which requires simultaneous consideration of system needs

3

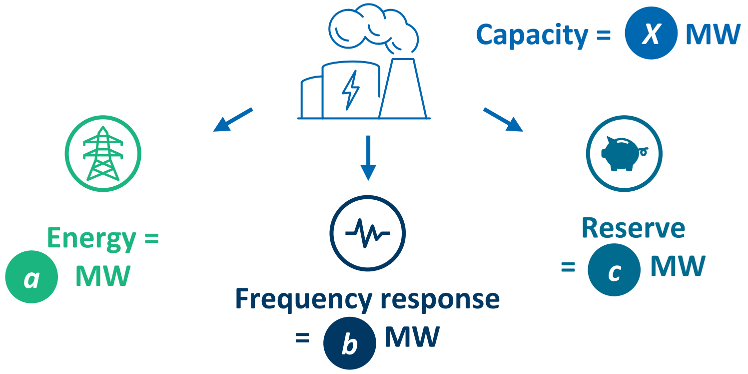
Timeframe for identifying optimal outcome under self-dispatch

- Control room operators have **limited time between gate closure and delivery** to identify optimal actions
- As such, ESO is only able to undertake **relatively simple and more immediately apparent optimisation** choices; e.g., looking to solve an upcoming shortage of reserve and a systematic energy imbalance in a least cost, optimised manner
- As a secondary point, under centrally scheduled markets with co-optimised energy and AS markets, sophisticated software is used to identify the optimal solution, which ESO currently does not deploy (given GB's self-dispatch design)

Challenges with the existing market design

Energy and some AS are partial substitutes - however this is not reflected in the existing market design where the energy and AS markets are independent

Most types of resources that provide AS services can also participate in wholesale and other markets[1]…



Capacity = X MW

Energy = a MW

Frequency response = b MW

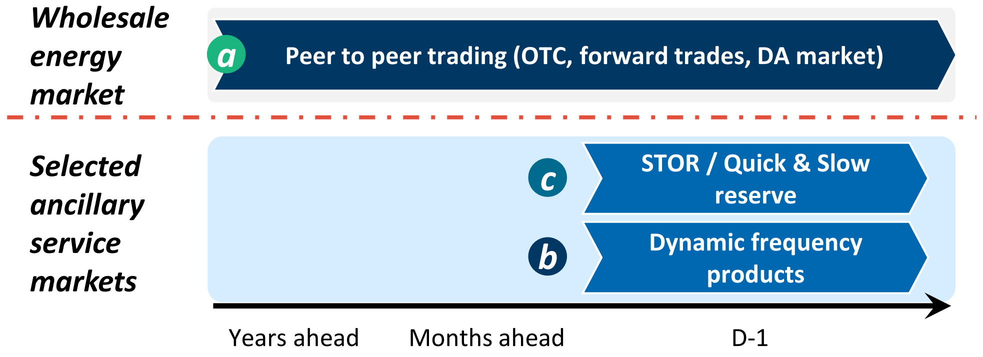
Reserve = c MW

- The provision of energy across different markets / services are **partial substitutes** - some types of resources are able to provide across multiple services…
- … however, the provision of any one of these products **impacts the ability of the generator to provide the other products**
- For example, a resource operating at full capacity in the wholesale energy market can only provide reserve or response by reducing supply into the wholesale market, i.e., the total MW capacity of the resource is the maximum that can be allocated to different services:

$$a + b + c \leq X$$

- As a result, by providing a particular service, the generator **faces an opportunity cost by virtue of not being able to provide another service.** This creates a co-dependency in terms of both volumes and prices of energy and AS being offered in different markets.

… however, GB's energy and AS markets operate independently of one another



Wholesale energy market

a Peer to peer trading (OTC, forward trades, DA market)

Selected ancillary service markets

c STOR / Quick & Slow reserve

b Dynamic frequency products

- Despite this substitutability, **the current market design treats energy and AS as distinct markets, often operated sequentially,** meaning market participants are **required to choose which market, or combination of markets, to participate into** ex ante
- In doing so, market participants have to form expectations as to:
  - ☐ which markets **other participants** will tender in and the **likelihood of being successful** in participating in those markets;
  - ☐ The resulting risk of the resource's **capacity becoming 'sterilised'** (i.e. prevented from participating in other markets, despite potentially being 'in merit' in other markets)
  - ☐ what prices different markets will clear at and therefore what the optimal bids/offers are, as a function of the **opportunity costs** of not participating in other markets.
- The need to form such expectations introduces significant information imperfections, with potential adverse impacts on the price formation process and the resulting market outcomes.

Notes: (1) However, not all resources active in energy wholesale markets have the capability to provide ancillary services and vice versa. Examples of resources active in ancillary services but not in energy wholesale markets are synchronous condensers (providing inertia) and flywheels (providing response). Conversely, large-scale thermal generators typically cannot provide some types of very fast response.

The separation of energy and AS markets can manifest several issues which, in turn, reduce system efficiency and increase costs

A numerical example is provided in the next slide

1

Allocative inefficiency

- Allocative inefficiency occurs when **resources are inefficiently allocated** across different markets, resulting in poorer outcomes for consumers

2

Reduced competition and market depth

- Shallow markets with few participants can result in higher prices, as a result of **reduced competitive pressure** and **low liquidity**

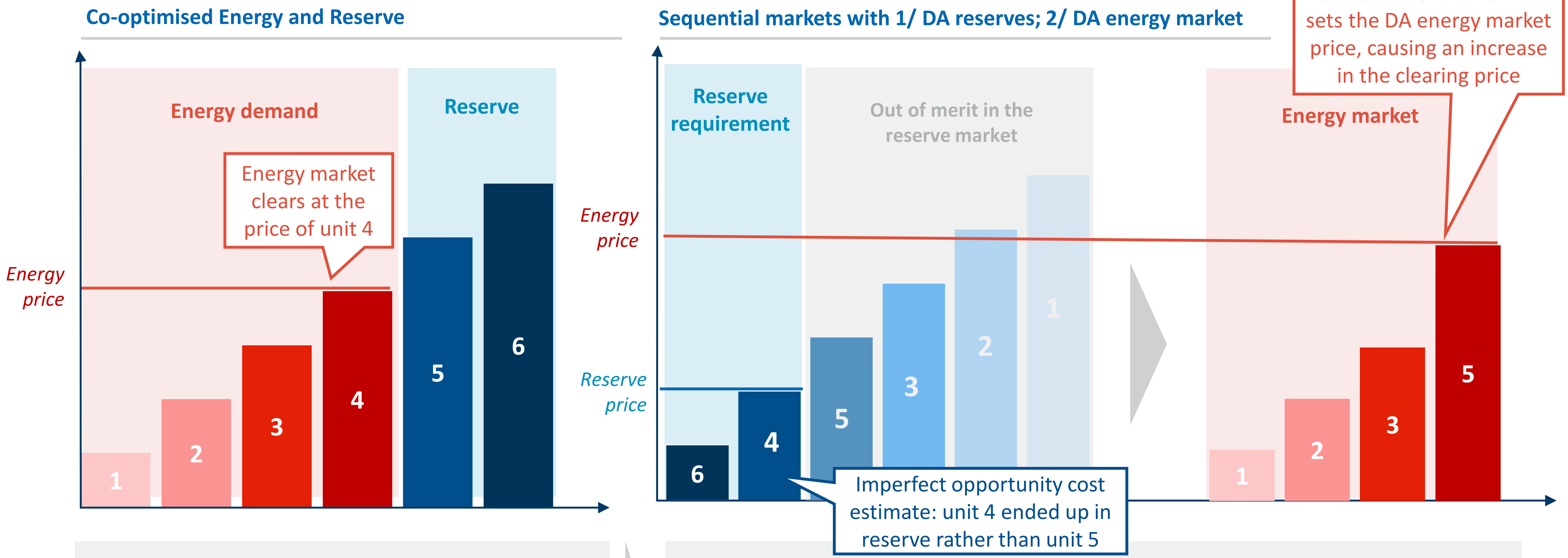
3

Poor price outcomes & dynamic inefficiency

- Prices across markets may **not reflect the true value** of each service (both at an absolute level, and comparatively across markets)
- This, in turn, sends **poor investment signals** to potential investors (or existing assets considering closure), resulting in long run inefficiency

Sequential markets lead to allocative inefficiencies, co-optimisation will impact AS costs and, possibly more important, likely reduce energy costs

- To illustrate the difference between co-optimised and sequential procurement of energy and ancillary services, we assume the market is served by **six generation units**.
- To meet the market's demand**, four units of energy** and **two units of reserve** are needed.



Co-optimised Energy and Reserve

Sequential markets with 1/ DA reserves; 2/ DA energy market

Energy market clears at the price of unit 4

Unit 5 rather than unit 4 sets the DA energy market price, causing an increase in the clearing price

Imperfect opportunity cost estimate: unit 4 ended up in reserve rather than unit 5

The co-optimisation algorithm schedules the units **with the lowest marginal costs in the energy market** while having **the more expensive units in reserves**

… by contrast, in sequential markets, market participants bid in reserve markets based on **estimated opportunity costs** of participating in the energy markets. Imperfectly estimated opportunity costs can lead to **inefficiently committed resources**.

1 Allocative inefficiency:

- Independent AS and energy markets can generate significant allocative inefficiency – i.e., where resources are not allocated optimally across uses - via multiple pathways
- Ultimately, allocative inefficiency will raise prices as the overall merit order of resources is not respected (even though it may be respected within individual markets), increase the risk of system failures and blackouts, and increase costs for consumers

Ex-ante expectations are incorrect

With independent markets, when a resource enters a particular market, it typically considers the opportunity cost of forgoing revenues in an alternative market as part of its bid, based on its ex ante expectations of future market outcomes.

These expectations may be incorrect, resulting in poor bidding behaviour - for example, pricing in an inaccurate opportunity cost estimate into bids - and entering the ‘wrong’ market

Example:
Market participants may expect high prices in the frequency market and low prices in the reserve market – however in reality this may not be true, with instead a shortage in the reserve market arising with oversupply in the frequency market

Resources sterilised from other markets

Market participants who are successful in a given market may, in fact, be more efficiently used from a system perspective by entering a different market. Unsuccessful resources may also be excluded from entering other markets where they could be in merit.

Example:
A flexible resource that has bid all its capacity into the reserve market but is unsuccessful may be unable to enter the frequency response auction, despite it being able to provide the service (and potentially being in merit in that market).

Example:
A relatively low marginal cost resource may be successful in the reserve auction and is therefore sitting idle; more expensive generation is therefore utilised in the wholesale energy market

Suboptimal market entry and dispatch

With independent markets, a resource might choose to enter and be dispatched in one market, when it could provide greater value in an alternative market

Example:
A flexible resource is used for energy provision in the wholesale market when it might be more needed for frequency regulation

Example:
If market participants expect high prices in the BM, they may withdraw from or bid very high prices into the wholesale energy market and DA AS markets, increasing prices in these markets

2 Reduced competition and market depth

- Having independent markets, as opposed to co-optimised markets, can reduce the degree of competition, as well as the depth and liquidity of the market:

Inefficient Resource Allocation

As discussed on the previous slide, resources that could provide value in multiple markets might be, in practice, selected to a particular market (e.g., ancillary services rather than the energy market, or vice versa).

This allocative inefficiency can, in turn, reduce competition in one market while oversaturating another.

Barriers to entry and complexity

Fragmented markets may require participants to navigate multiple sets of rules, bidding processes, and operational requirements.

This complexity can deter smaller participants or those with limited resources from entering certain markets, reducing competition, market depth and liquidity.

Strategic bidding and market power

Larger participants might be able to engage in strategic bidding across multiple markets, or exercise market power in each individual relatively small market.

While market rules would typically be in place to prevent such behaviour, shallow separate markets may be more prone to market abuse. Also, bidding strategies are more complex in non-co-optimised markets (i.e., conditional upon opportunity cost estimations) making it likely more challenging to monitor and prove such market manipulation occurred across the distinct markets.

Example:

A market participant might withhold capacity from one market to drive up prices or to capitalise on anticipated price differentials between the markets. This behaviour distorts competition and price signals.

3 Poor price outcomes & dynamic inefficiency

- Separate markets can also result in poor pricing outcomes, with implications for long run efficiency:

Prices may not reflect true marginal cost of service

In separate markets, the price of energy might not fully reflect the system's marginal cost or value when ancillary services are also being procured, and vice versa. This is driven by the incorrect formation of price expectations which feed into the opportunity cost assessment of participating in a specific market.

Example:
A generator that can also provide AS has a higher energy offer price, the energy market might not dispatch it, even if it's the most cost-effective solution when considering both energy and ancillary service values – requiring potentially expensive additional balancing actions in real time.

Over or under procurement

Without co-optimised scheduling of energy and AS, the system might relatively over-procure one service and relatively under-procure another, leading to inefficiencies and potentially higher costs.

Example:
If prices appear cheap in the reserve market, ESO may procure higher volumes – however this could result in over procurement to the extent that it unintentionally removes capacity from and raises prices in the energy market.

Dynamic inefficiency

If prices in the separate markets don't accurately reflect the true value of services, it provides poor investment signals, discouraging investments in flexible resources that could serve both markets effectively (or, conversely, failing to provide correct signals for assets considering retirement).

Example:
If AS market doesn't adequately compensate for flexibility, traditional power plants may exit prematurely without replacement with other flexible assets. Over time, will result in a shortage of essential flexible resources, compromising grid reliability and efficiency.

Further challenges presented by the energy transition to Net Zero

Historically, the separation of the wholesale electricity market and ancillary services was not a significant issue to the ESO and consumers

Historical ESO procurement of ancillary services

- In the past, the GB power system relied on **large thermal resources** located at transmission voltages to meet energy demand and provide ancillary services
- Ancillary service procurement and scheduling were therefore relatively simple:

A significant proportion of **AS were provided as a by-product** of units already generating; for instance, inertia was provided by large spinning turbines in gas or coal plants

ESO was able to call on a **relatively small number of large thermal generators** to meet its ancillary service needs, reducing the complexity of optimising procurement and scheduling in real time

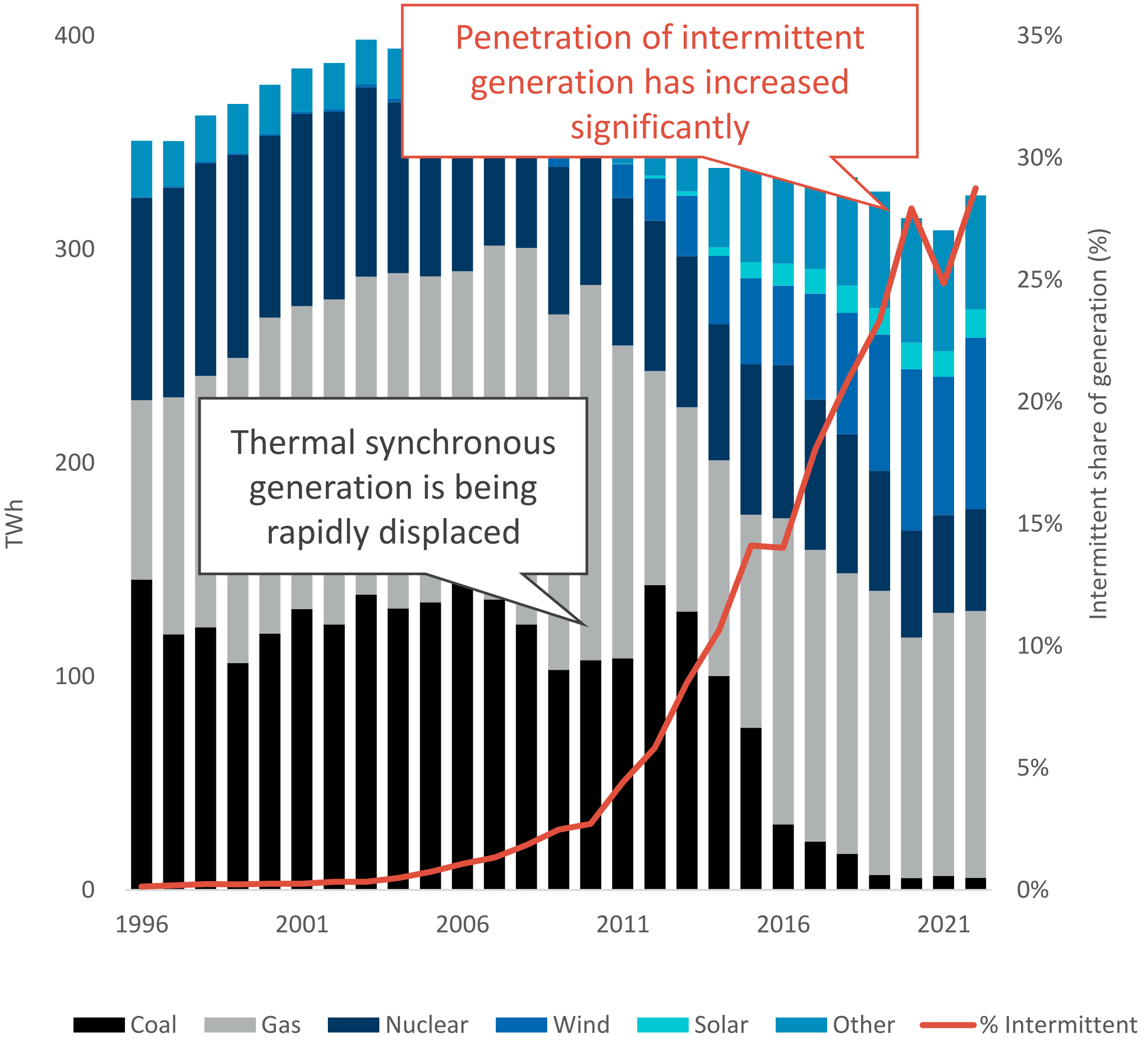
Overall requirements for procuring additional **ancillary services were therefore relatively limited**, resulting in low costs and reducing the impact of any potential procurement inefficiencies

However, the GB energy market is evolving rapidly, with significant volumes of renewable capacity being deployed on the electricity system

The GB energy transition

- The UK, like an increasing number of countries around the world, has legislated for Net Zero 2050, and aims to operate a **fully decarbonised electricity grid by 2035**
- As a result, renewable generation capacity on the system has **increased from 5 GW to 27 GW** between 2002 and 2022.[2] This corresponds to an **increase in intermittent generation** as a share of total generation to almost 30%.
- This change has been facilitated by **technological development, such as batteries and falling costs for wind and solar**, supported by schemes such as Contracts for Difference ("CfDs") and carbon pricing

GB generation mix, 1996-2022 (TWh)[1]


Penetration of intermittent generation has increased significantly

Thermal synchronous generation is being rapidly displaced

Sources: (1) BEIS (link); (2) Dukes 5.7, hydro, wind, solar and biomass

The transition has exposed the flaws in the existing GB market design, requiring significant intervention from ESO and driving up costs

The energy transition is straining GB's market design

- There are two overarching challenges **to ESO's ability to operate the system effectively and efficiently**, each with underlying drivers:

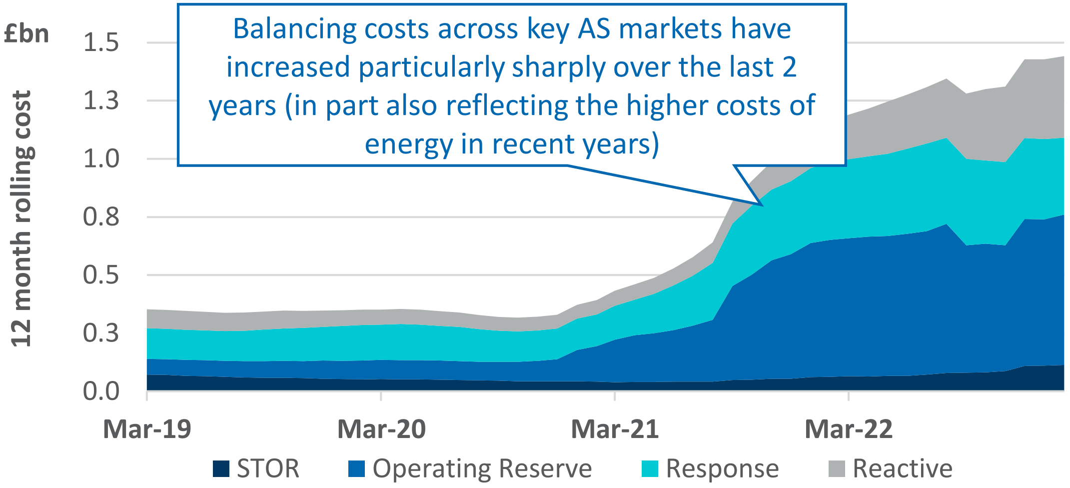
1) Increasing service volume requirements

Increasing system variability	The inherent variability of renewables – in particular wind – and the growth of interconnector capacity increases the need for balancing, redispatch and AS
Reduced AS supply	Inverter-based generation, for examples solar, is less able to provide some AS a by-product of generation (e.g. inertia), and renewables can also face weaker incentives to provide AS[1]

2) Increasing procurement and scheduling complexity

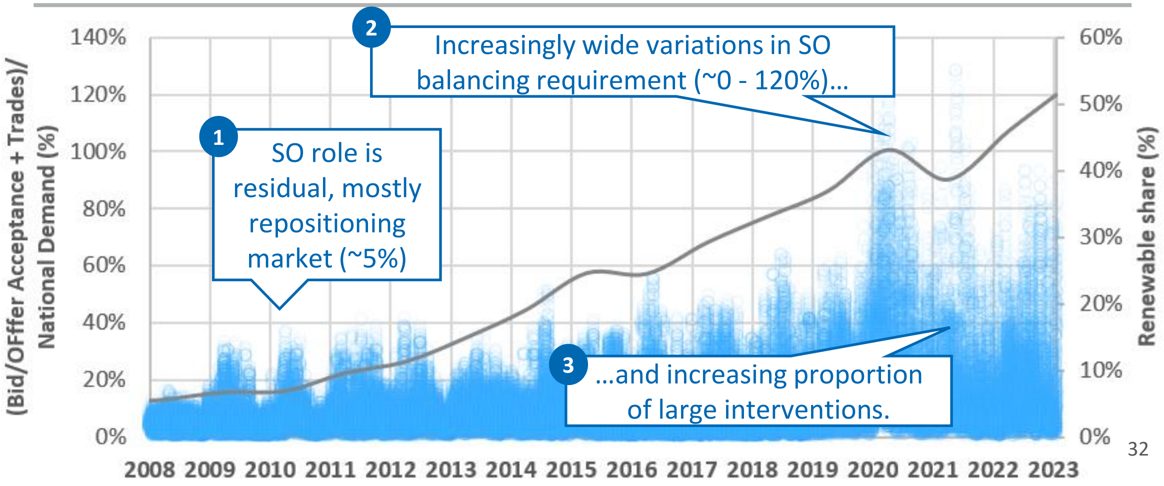
# of market participants	With the rollout of renewables, the number of generators has increased significantly, spread widely across GB
Variety of technologies	Different technologies, for example renewables and interconnectors, have significantly varying characteristics (e.g., dispatchability, duration, scheduling complexity)

Cost of selected ESO balancing services



Balancing costs across key AS markets have increased particularly sharply over the last 2 years (in part also reflecting the higher costs of energy in recent years)

The number of balancing instructions by the ESO has increased



① SO role is residual, mostly repositioning market (~5%)

② Increasingly wide variations in SO balancing requirement (~0 - 120%)…

③ …and increasing proportion of large interventions.

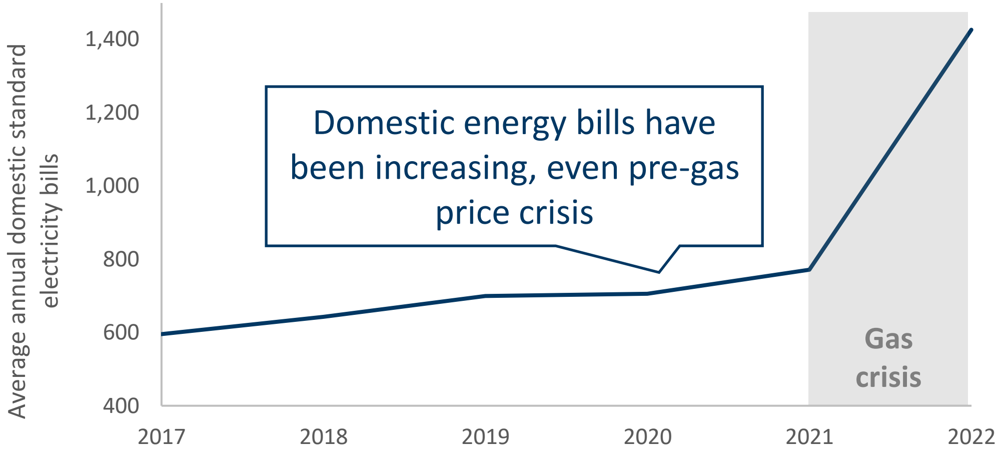
Notes: (1) For example, contracts for difference can reduce the incentive for wind generators to provide frequency responses service

The challenges to the existing market design are expected to grow in future as the system continues to transition towards Net Zero

The challenges to the market design are likely to continue growing in the future

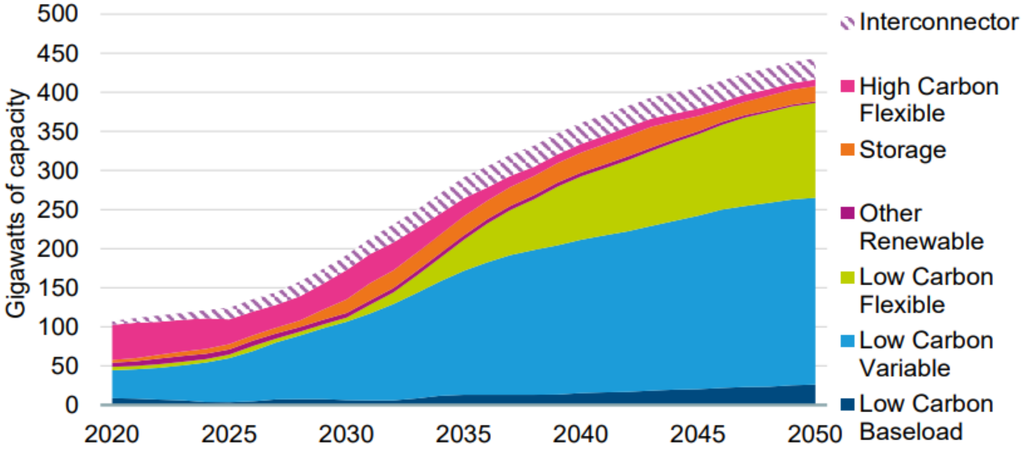
- As the GB generation mix continues to shift towards intermittent resources on the path to Net Zero, the challenges with balancing (see previous slide) are expected to continue to grow
- At the same time, consumer bills have been rising, even pre recent gas crisis…
- …and there is therefore an acute need to ensure consumer demand is met at least cost
- There is therefore a significant pressure on policy makers to deliver a market design that minimises the total cost of meeting Net Zero
- In the context of this report, this means that policy options that may reduce the total energy and AS costs should be explored as a matter of priority

Average annual domestic standard electricity bills (£)[1]



Domestic energy bills have been increasing, even pre-gas price crisis

Forecast Capacity Mix, 2020-2050, (GW)[2]



Sources: (1) DESNZ Annual domestic energy bills (link) – assumes annual consumption of 3,600 KWh/year (2) REMA Consultation 2022, p27 (link).

Co-optimisation in the context of Net Zero

In the context of the energy transition, energy and AS co-optimisation could contribute to a more efficient delivery of Net Zero targets

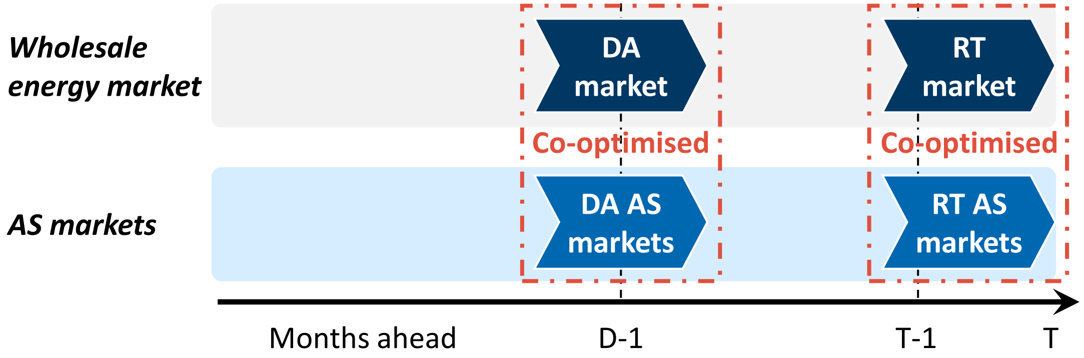
What is co-optimisation of energy and AS?

- Co-optimisation is the process of **simultaneously co-ordinating and producing the optimal schedules and prices** for energy and AS
- This leads to an **efficient allocation of capacity** based on the relative needs and conditions of each market…
- …and the **efficient formation of prices** across each market

What is the opportunity for co-optimisation of energy and AS?

- We take having a **centralised scheduling** in place as a given in this report - the merits of moving from GB's current self-scheduled market design to a centrally scheduled one are not examined in this report, for more discussion see Appendix 1.
- Commonly, **markets that co-optimise operate a day ahead market and real time market** for both energy and AS (as shown on the right)
- This can be implemented both **with and without locational pricing** (to different effects).

Centralised market – with co-optimisation



Energy and AS markets operate in an integrated, co-optimised manner, **removing the need for resources to choose which markets to participate in** and/or perform complex estimations of the **opportunity costs of participating in different markets**.

Notes: (1) ESO currently manages all aspects of ancillary services procurement and scheduling. With a potential move to central scheduling, additional responsibilities would arise in relation to the setting of market rules, structures and price determination ahead of real time and, subsequently, extended into co-optimisation with AS. This role could be undertaken by a Market Operator ("MO") or added to the ESO's current responsibilities. The choice between these two approaches is not within the scope of this report.

Theoretical benefits and risks of co-optimisation

We have considered the theoretical benefits and risks of co-optimising energy and AS markets in the context of GB's transition to Net Zero

- In this subsection, we set out the theoretical benefits and risks (both perceived and actual) of implementing the co-optimisation of energy and ancillary service markets:[1]

Key benefits	Key risks
Allocative efficiency	Mathematical complexity / feasibility
Lower risk and bidding strategy complexity for market participants	Transparency
Greater competition	Reduced revenues / profits
Operational and dispatch efficiency	Operational challenges for ESO
Dynamic efficiency	

We consider impacts on **market outcomes** (e.g., energy and ancillary service prices), as well as broader impacts on **market participants**, **ESO** and GB's **transition to Net Zero**

- For each key benefit and risk area, we summarise our impact assessment from 'significant positive' to 'significant negative':

↑↑ **Significant +ve impact**
↑ **Moderate +ve impact**
↔ **Neutral impact**
↓ **Moderate -ve impact**
↓↓ **Significant -ve impact**

Notes: (1) As explained elsewhere in this report, implementing the co-optimisation of energy and ancillary service markets is commonly implemented under central scheduling. We do not consider the potential benefits and risks of moving to a central scheduling, instead focusing solely on the benefits of co-optimisation of energy and ancillary services.

Co-optimisation can improve allocative efficiency and reduce system costs by committing and utilising resources where they are most valuable

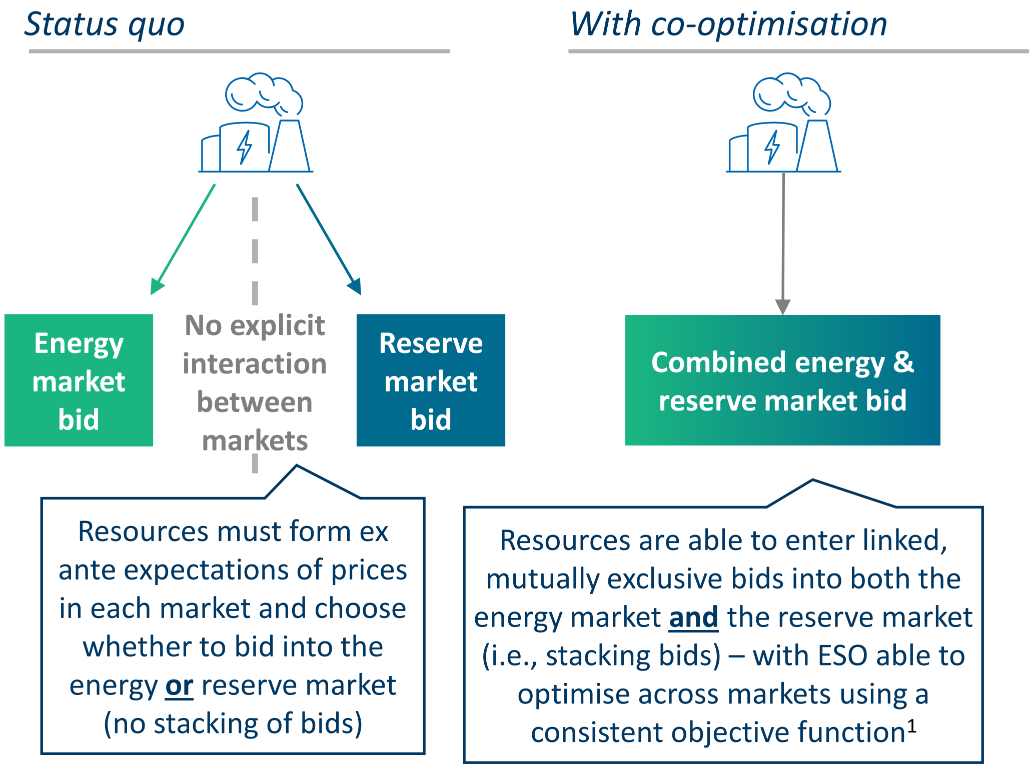
Impact of co-optimisation on allocative efficiency

- A key limitation of GB's existing market design is the risk of allocative inefficiency - **co-optimisation can eliminate this risk:**
  - ESO operates both the energy and AS markets simultaneously, meaning it can allocate resources across system needs in the most efficient way (i.e., resources are deployed where most valuable), using a consistent objective function across markets
  - Market participants can enter linked, mutually exclusive bids into multiple markets (i.e., bid stacking), eliminating the risk of entering the 'wrong' market
- Optimally allocating resources across markets will, in turn, reduce total system costs, generate more efficient price signals across markets and support the decarbonisation of the power system

Supporting evidence (see Appendix 2)

Midcontinent Independent System Operator (MISO): MISO, the system operator for the Midwest USA, estimates that introducing its co-optimised AS markets in 2009 (in place of the existing system of 26 separate balancing authorities across the network) has generated at least $60m per year in benefit (i.e., these are cost reductions in AS procurement, the indirect cost reductions with regards to energy costs have not been evaluated).

Electric Reliability Council of Texas (ERCOT): ERCOT, the system operator for Texas, has recently decided to introduce co-optimisation of energy and reserves in real time (in addition to its existing day ahead co-optimisation), with expected annual savings of $160m in AS markets and $4/MWh reduction in average energy prices saving a further $1.6bn per year



Status quo

Energy market bid

No explicit interaction between markets

Reserve market bid

Resources must form ex ante expectations of prices in each market and choose whether to bid into the energy **or** reserve market (no stacking of bids)

With co-optimisation

Combined energy & reserve market bid

Resources are able to enter linked, mutually exclusive bids into both the energy market **and** the reserve market (i.e., stacking bids) – with ESO able to optimise across markets using a consistent objective function[1]

Evaluation summary

⬆⬆

Co-optimisation of energy and ancillary service markets can eliminate allocative inefficiency, reducing total system costs

---

Note: (1) Introducing central scheduling and dispatch can lead to changes in bidding formats not necessarily requiring linked mutually exclusive bids but rather multi-part bids including an energy offer price and the technical characteristics of the relevant resources

Revenue uncertainty and bidding strategy complexity can also be reduced with co-optimisation, with the need to estimate ex-ante opportunity costs removed

Impact of co-optimisation on revenue risk and complexity

Revenue risk

- With independent markets, prices across markets can be inconsistent and vary significantly
- As such, market participants' revenues can be uncertain and volatile…
- …with high revenues in under supplied markets and low revenues in over supplied markets
- Market participants also face the risk of being unsuccessful in their chosen market and being sterilised from other markets
- Co-optimisation can reduce the volatility and risk of a market participant's revenues: prices are consistent across markets, resources are able to stack bids, and the risk of sterilisation is removed

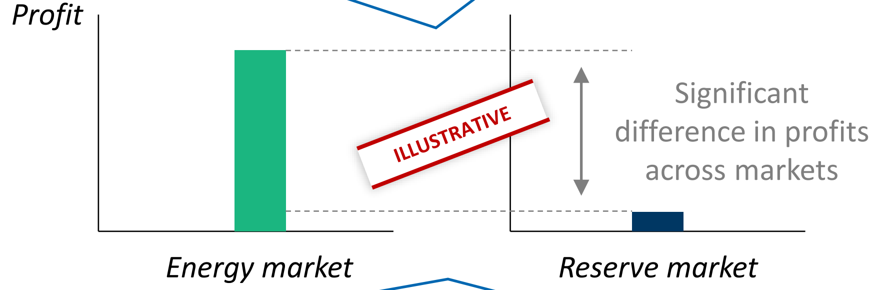
Co-optimisation can improve the consistency of prices across markets, increasing revenue certainty and reducing the need for complex bidding strategies

Bidding strategy complexity

- Uncertainty associated with independent markets increases the need for complex bidding strategies (based on ex ante expectations of market outcomes) in order for resources to maximise revenues
- This, in turn, may benefit large established market participants, who may be able to develop more sophisticated strategies and hedge across markets
- Co-optimisation reduces this issue, with market participants no longer required to form ex ante expectations of market outcomes to determine bid strategy

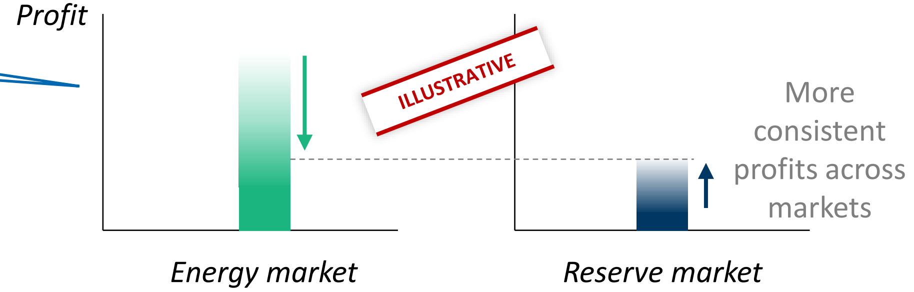
Status quo

Resources have to choose market to enter based on (imperfect) ex ante expectations of these profits…



… which can result in over/under supply across markets - increasing price uncertainty and volatility between markets

With co-optimisation



Supporting evidence

- During ESO's EAC consultation process, stakeholders commented that they "*welcomed addition to the new market clearing algorithm as it reduces the "game theory" aspect of the current market structure and allows us to de-risk many of our optimisation processes*".[1]

Evaluation summary

⬆

Co-optimisation of energy and ancillary service markets improve revenue certainty and de-risk investment for market participants

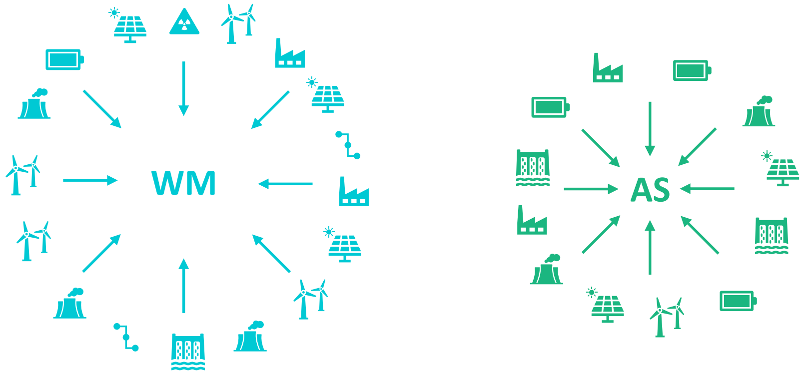
Notes: (1) ESO consultation on EAC, page 1, (link)

Resources can enter multiple markets simultaneously with co-optimisation, increasing competition and liquidity

Impact of co-optimisation on competition and liquidity

- With co-optimised markets, market participants are able to enter multiple markets simultaneously, with stacked bids
- As such, each market can have a greater number of competing resources which, in turn, can:
  - Place downward pressure on market prices;
  - Facilitate more effective price discovery and transparency;
  - Reduce liquidity issues associated with shallow markets, for example high price volatility; and
  - Reduce the likelihood of ESO being unable to meet its optimal levels of AS procurement.
- Greater competition can also reduce the ability of larger market participants to exercise market power, unlike the status quo with independent AS and energy markets

Status quo



With co-optimisation

With co-optimisation, ESO would be able to optimise the dispatch across combined, and therefore more competitive and liquid, energy and AS markets

Supporting evidence

- The potential savings from co-optimisation of energy and ancillary services have been estimated to reach between 30-50% of ancillary service costs for markets in New Zealand, Australia, and Singapore[1]

Evaluation summary ⬆

Co-optimisation of energy and ancillary service markets can increase competition across markets, with resources able to stack bids into multiple markets, reducing system costs

Source: (1) Read, E. Grant (2010), co-optimisation of Energy and Ancillary Service Markets

Co-optimisation can reduce the volume of balancing actions required in real time and the need for manual interventions by ESO in compressed timescales

Impact of co-optimisation on operational efficiency

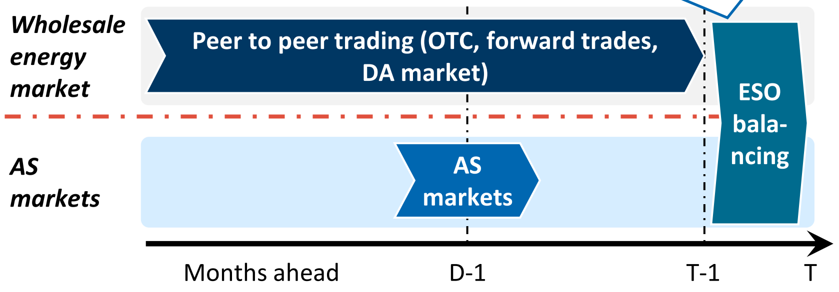
- Implementing co-optimised energy and AS markets (commonly accompanied by the introduction of centralised scheduling) can help to **reduce the operational complexity** of ESO operations
- In particular, ESO would be able to clear the energy and AS markets simultaneously, close to real time, which can **reduce the volume of out of market balancing actions** required by ESO…
- … while also **increasing the likelihood of the optimal allocation** of resources being identified, boosting operational efficiency
- Under co-optimisation the **automated clearing process** is likely to be more **auditable** and reduce any perceptions of control room bias towards certain technologies
- Improved operational efficiency can, in turn, reduce costs for consumers while also increasing system security and supporting the decarbonisation of the power system

The benefit of reducing the need for close-to-delivery balancing actions is partly dependent on the locational design of the market.

The impact of market design on the benefits of co-optimisation is discussed in more detail in slide 51

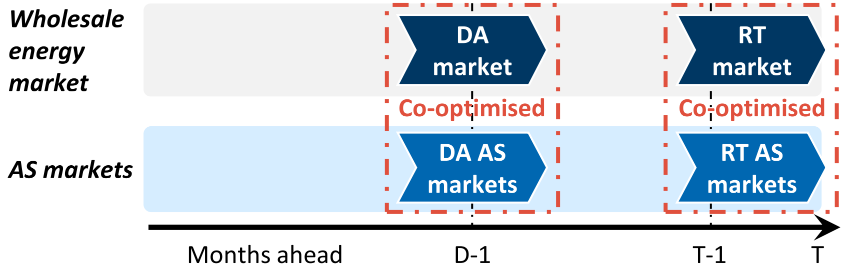
Status quo

Significant manual balancing actions required in compressed timescales



With co-optimisation

Computer algorithm identifies optimal outcome across both energy and AS markets, reducing need for manual balancing actions



Evaluation summary ↑

Co-optimisation of energy and ancillary service markets can increase operational efficiency, by reducing need for significant balancing interventions in compressed timescales, and improve system reliability

Co-optimisation improves system efficiency and price signals, and in the long run incentivises most valuable resources to enter and remain in the market

Impact of co-optimisation on dynamic efficiency

- GB's transition to Net Zero will require significant investment in new energy infrastructure, including clean generation resources
- The benefits described over the preceding slides can improve investment signals and therefore support the decarbonisation of the power system:

Short run impacts	*Long run dynamic efficiency impact*
Allocative efficiency	**Encourages investment** in resources that can **provide more valuable services** (e.g., flexible resources required to operate a decarbonised electricity system)
Lower risk and bidding strategy complexity for market participants	**De-risks investment**, particularly in **clean generation technologies where a high proportion of costs are incurred up front**, and **reduces the complexity required to develop bidding strategies**, **lowering barriers to entry**
Greater competition	Facilitates **price discovery and mitigates market power of large established players**, encouraging investment into value service areas while also prompting the efficient exit of resources by influencing decision to retire inefficient plants early and prolong the operational life of more valuable ones
Operational efficiency	Increases **operational efficiency** and **reduces out of merit dispatch**, which could further **incentivise investments in small and decentralised assets** such as batteries

Evaluation summary ⬆

Co-optimisation of energy and ancillary service markets can improve investment signals and, in turn, dynamic efficiency as GB decarbonises

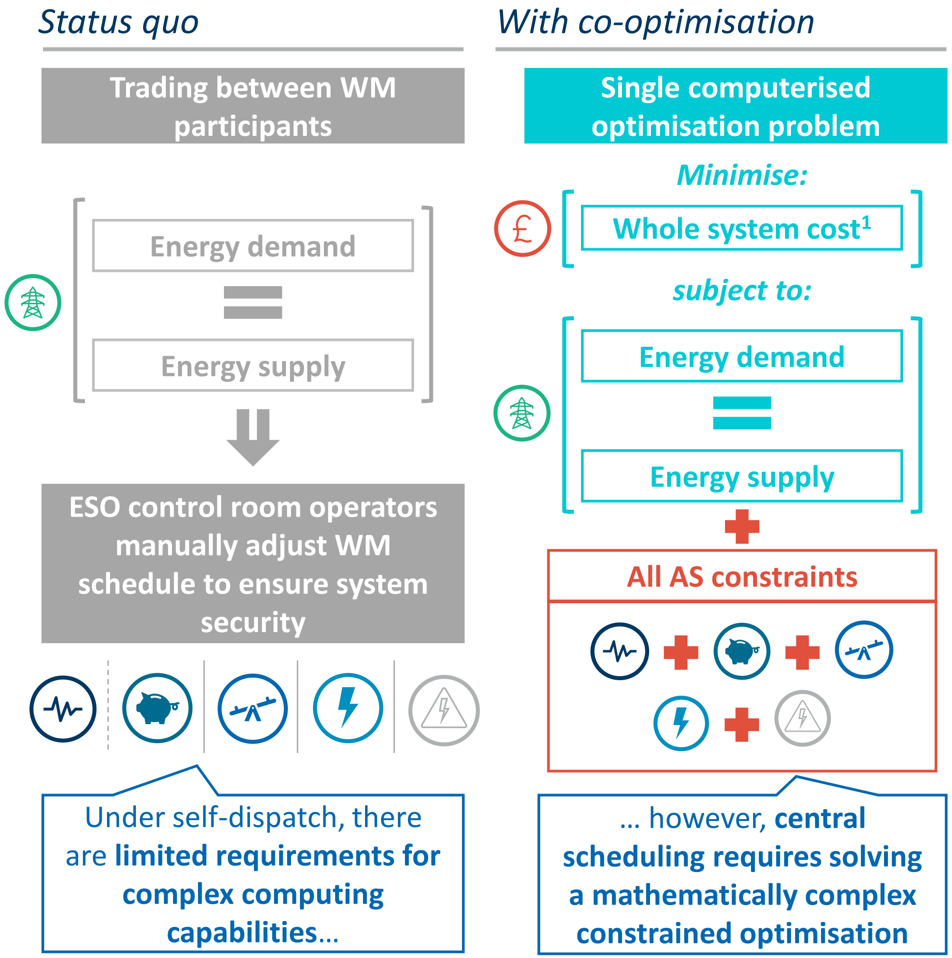
Co-optimisation represents a complex problem requiring powerful computing capabilities; however, it has been successfully implemented in many markets

Impact of co-optimisation on mathematical complexity

- Simultaneously computing the optimal allocation of energy and ancillary services is computationally challenging for several reasons:
  - **Power system operations often involve non-linearities and non-convexities**, leading to multiple local minima; it is therefore challenging to compute the global minimum or maximum of the objective function (this problem also exists in energy market-only optimisation problems, but to a lesser degree)
  - There is **interdependence across markets**, creating further complexity in the optimisation problem
  - Necessity for **real-time decision-making** imposes a significant challenge, demanding algorithms that can swiftly and accurately find solutions amidst the complexities of the problem
- As such**, identifying the optimal dispatch is a computationally intensive**, time sensitive and potentially intractable optimisation problem with solutions that may make it hard for participants to discern clear patterns or rationales

Risk mitigations and wider considerations

- **Centralised scheduling markets that co-optimise ancillary services with energy are commonplace internationally,** demonstrating the technical feasibility of designing and implementing such systems
- Co-optimisation of energy and AS was first successfully introduced in New Zealand in 1996 – significant innovation and technological progress has occurred over the past 30 years
- Changes to bidding formats and the pricing rule (linear vs non-linear pricing with uplifts) can strongly reduce computational complexity (and this may even be possible for more complex bidding strategies such as those of storage assets).



Status quo

Trading between WM participants

Energy demand
=
Energy supply

ESO control room operators manually adjust WM schedule to ensure system security

Under self-dispatch, there are **limited requirements for complex computing capabilities…**

With co-optimisation

Single computerised optimisation problem

Minimise:

Whole system cost[1]

subject to:

Energy demand
=
Energy supply
+
All AS constraints

… however, **central scheduling requires solving a mathematically complex constrained optimisation**

Evaluation summary

Co-optimisation can (non-linearly) increase the computational complexity as more variables are added.

Many of the challenges can be resolved through linear programming, advanced algorithms and improved hardware.

Notes: (1) Alternatively, the objective function could be 'welfare maximisation'. Both options exist, and it is beyond the scope of this report to choose the appropriate objective function for the algorithms

Co-optimisation may be perceived as opaque – however stakeholder understanding can be improved via education and its results are auditable

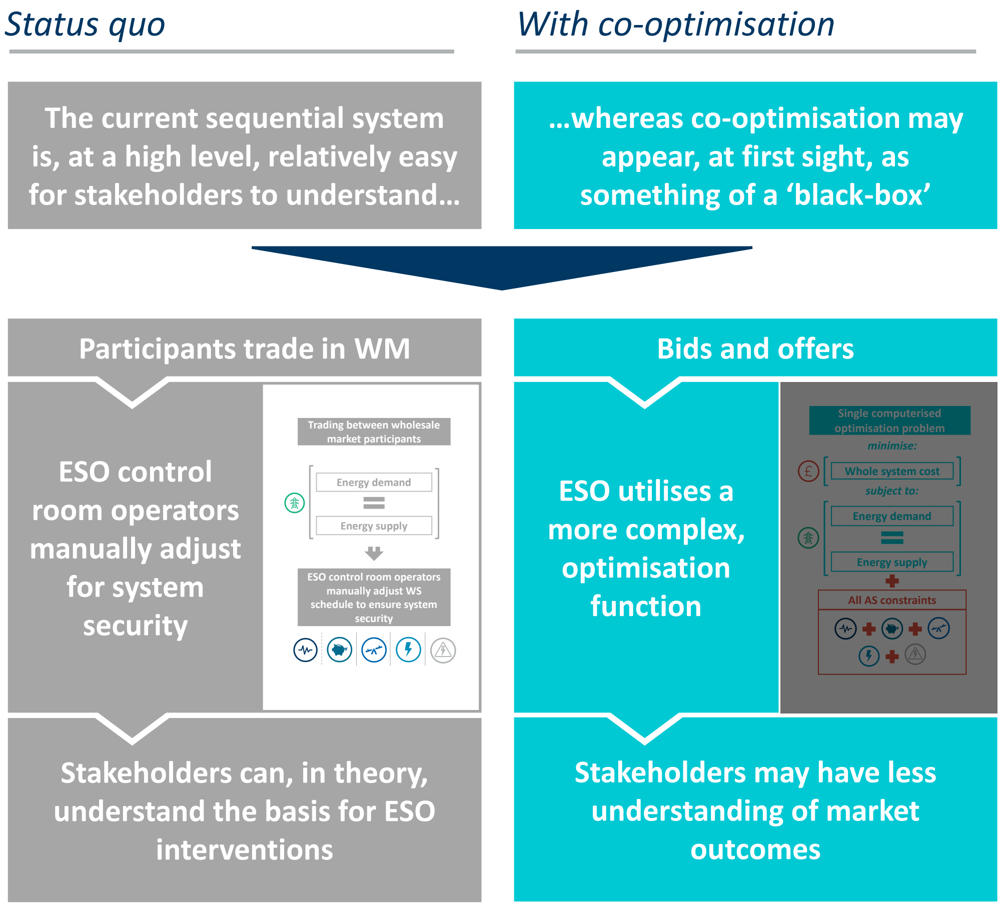
Impact of co-optimisation on transparency

- A move to co-optimisation could risk a **perception of reduced transparency regarding market outcomes**
- In the status quo, the market structure is, at a high level, relatively simple:
  - bilateral trading establishes energy market outcomes
  - post gate closure, ESO makes manual interventions using the BM and pre-procured AS to maintain system security, with ESO actions flagging depending on the driver of the need (e.g., system vs. energy needs)
- In theory, this promotes transparency for market participants
- Under a co-optimised market, multiple markets are being simultaneously solved via a complex computerised dispatch engine…
- … risking the **perception of a 'black box' solution being provided**

Risk mitigations and wider considerations

- Risk can be reduced by providing stakeholders with **training and education** on how co-optimised markets function
- Co-optimisation also brings **benefits in transparency**, in terms of generating increased visibility of prices and better access to market data, due to all services being priced via a single optimisation
- **Outcomes generated via a computerised dispatch are auditable** – this is a key challenge under the current design where ESO actions include a human decision-making element under compressed timescales which can be hard for ESO to communicate (or explain ex-post) and market participants can struggle to understand
- Markets which have implemented co-optimisation have been able to **implement reforms to improve transparency over time**, e.g., PJM [1]

Sources: (1) PJM, Proposed Enhancements to Energy Price Formation, 2017 (link)



Status quo

With co-optimisation

The current sequential system is, at a high level, relatively easy for stakeholders to understand…

…whereas co-optimisation may appear, at first sight, as something of a 'black-box'

Participants trade in WM

ESO control room operators manually adjust for system security

Stakeholders can, in theory, understand the basis for ESO interventions

Bids and offers

ESO utilises a more complex, optimisation function

Stakeholders may have less understanding of market outcomes

Evaluation summary

Initially decreased transparency is a consequence of moving into any new system, which can be mitigated through education. Greater visibility of all prices and 'auditability' may even cause a net increase in transparency over time.

Lower system costs would reduce the revenues of some resources - however this, in turn, would likely benefit end consumers

Impact of co-optimisation on distribution of revenues/profits

- As a result of the improvements in allocative, operational and dynamic efficiency, a move to a **co-optimised system will likely change the distribution of the revenues/profits across different resources**, as well as a reduction in the overall revenues/profits of the market
- Some stakeholders may also **perceive a centralised scheduling to be less efficient** than the existing self-scheduling market (and/or to **stifle innovation** in market participants' bidding strategies)…
- … particularly for assets with complex bidding strategies such as batteries (e.g., if participants are limited by the formats of the bids they are allowed to submit).
- Investments by market participants in developing sophisticated capabilities to determine bidding strategies for the existing market design may become redundant with co-optimisation

Risk mitigations and wider considerations

- A **degree of redistribution is to be expected with any significant market reform** creating winners and losers…
- … and **consumers should benefit from overall lower costs**
- Market participants could benefit from increased revenue certainty and reduced volatility (as previously presented on slide 38)
- Strong stakeholder engagement and testing to define appropriate products/bid formats (noting that many US ISOs are actively exploring options for designing bid formats to address the perceived risk of stifled innovation[1]) can reduce the risk of reduced revenue for some market participants and provide greater clarity to all

Status quo

Relatively higher total system costs and revenues for resources

With co-optimisation

Reduces total system costs and revenues…

… and different distribution of revenues across resources

Evaluation summary

Overall profits are likely to decline with a shift to co-optimisation, but this may be somewhat compensated by a reduction in volatility of revenues and is to the benefit of consumers

Notes: (1) Singhal, N.G.; Ela, E.G. Storage Integration Efforts in the U.S. Wholesale Electricity Markets IESO Energy Storage Design; Project; EPRI Institute: Palo Alto, CA, USA, 2020. (link)

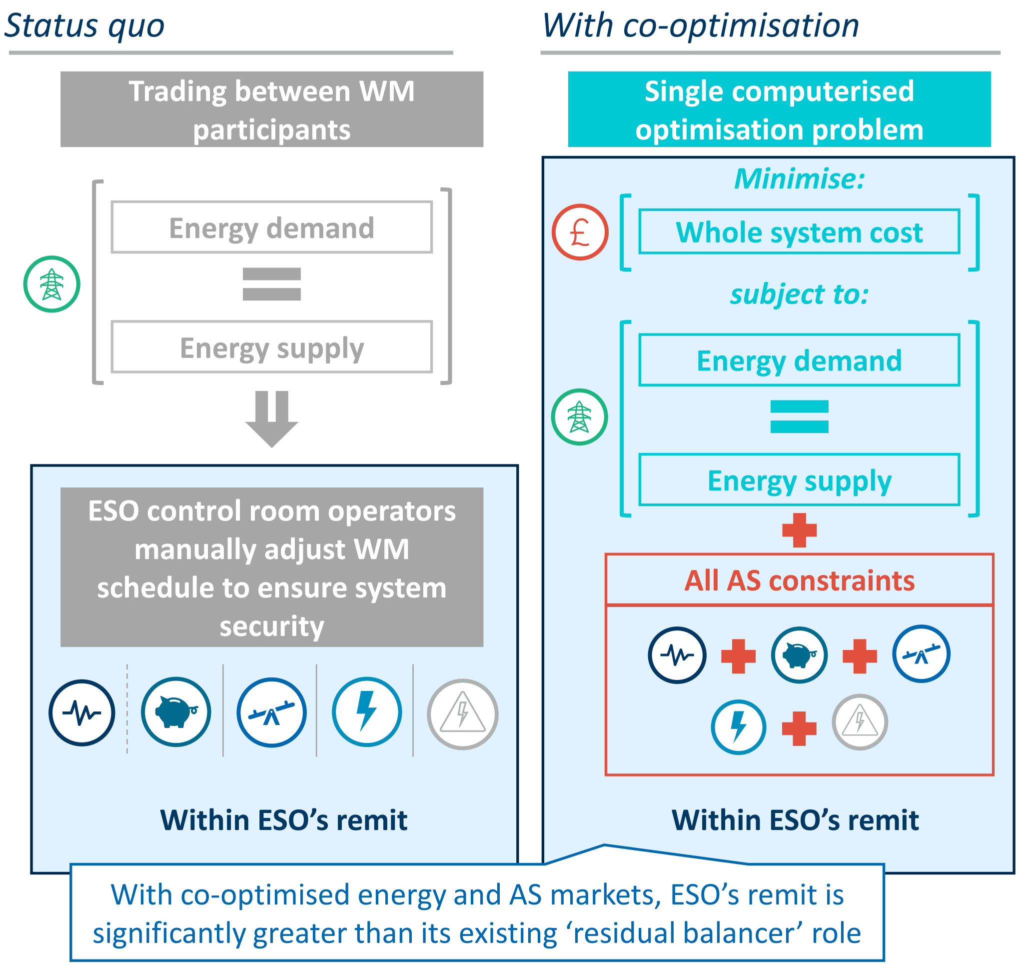
With co-optimisation, ESO's role would be expanded; however, this is primarily a function of transitioning from a residual balancer to a centralised scheduler

Impact of co-optimisation on ESO operations

- The shift to **co-optimisation of energy and AS will require operational changes at ESO** - including the likely introduction of centralised scheduling:
  - — ESO will require more **sophisticated computational infrastructure**…
  - — …. and will need to **integrate datasets**, potentially from various sources
- Additionally, depending on the specific market design implemented, **ESO could be required to manage both DA and RT market clearing**
- As a result, significant **financial investment will be required** in physical hardware, software and human capital
- Increase in ESO's responsibilities, from a 'residual balancer' to market operator for energy and AS markets, could lead to increased complexities, increased likelihood of operational errors, and reputational damage
- In addition, any issues that do occur could have broader implications in a co-optimised market relative to a siloed market
- However, **these risks are predominantly a function of moving from the 'residual balancer' role to a centralised scheduler** and are not incremental to implementing co-optimisation once a central scheduling would be in place

Risk mitigations and wider considerations

- Under current market arrangements, **ESO control room operators are increasingly required to take a significant number of actions** – at times akin to centralised scheduling- without the support of computerised optimisation tools
- Higher computational demands will therefore be partially offset by more straightforward near-real-time balancing requirements
- Skill demands can **be mitigated by training and hands-on experience**
- Some of the costs would be offset by reduced operational complexity for market participants (e.g., bidding strategy)
- There is significant global experience of operating centralised scheduling



Status quo

Trading between WM participants

Energy demand

=

Energy supply

ESO control room operators manually adjust WM schedule to ensure system security

Within ESO's remit

With co-optimisation

Single computerised optimisation problem

Minimise:

£ Whole system cost

subject to:

Energy demand

=

Energy supply

\+

All AS constraints

Within ESO's remit

With co-optimised energy and AS markets, ESO's remit is significantly greater than its existing 'residual balancer' role

Evaluation summary

Co-optimisation of energy and AS markets will extend ESO's role and require significant investment in new computing and other operational capabilities - but will reduce the need for near-real-time balancing actions

Summary of assessment

Summary of benefits of co-optimising energy and AS

Description	Commentary	Evaluation	Supporting evidence
1 **Allocative efficiency**	▪ ESO deploys available resources in **most efficient manner** (with overall **merit order respected**) & **reduces costs** of meeting demand ▪ **Consistent objective function** is optimised across markets ▪ Simultaneous clearing across energy and AS markets lead to **more efficient price signals**	⬆⬆	International evidence of benefits is limited given that co-optimisation is often implemented in parallel with other reforms (e.g., transition to central scheduling)… …and benefits are challenging to disentangle from wider market changes (e.g., RES investments). Nevertheless, we found that: 1. MISO estimated annual savings in AS markets of at least $60m as a result of co-optimisation (see Appendix 2). Academic studies estimated that co-optimisation could reduce AS costs by 30-50% (Read, 2010) 2. In ERCOT's cost benefit analysis for the introduction of real-time co-optimisation the **largest gains are expected in the energy market** ($1.6bn energy cost reduction vs $160 million reduction in AS costs) 3. Studies on the NEM suggest that co-optimisation, combined with other market design elements, results in more efficient long-term signals for investment in flexible resources (link)
2 **Lower risk and bidding strategy complexity**	▪ **Reduced bidding strategy complexity** as no need to choose which markets to enter ex-ante ▪ **No risk of sterilised capacity** ▪ Easier to **stack multiple revenue streams** ▪ **Reduced financial and operational uncertainties** faced by market participants	⬆	
3 **Greater competition and liquidity**	▪ Market participants are able to **enter all co-optimised markets simultaneously**, resulting in deeper markets ▪ **Reduced risk of undersubscribed** markets ▪ Improved **market access and transparency** could lead to **increased competition** in individual markets	⬆	
4 **Operational efficiency**	▪ Co-optimisation can **simplify the dispatch process**: ESO can simultaneously clear energy and AS, **reducing the need for near-real-time balancing** ▪ This is particularly the case if co-optimisation takes into account locational signals (e.g. transmission constraints at DA stage) ▪ Reduced volume of balancing actions required in RT and the automated clearing process is likely to be **more auditable** and reduce any perceptions of control room bias towards certain technologies	⬆	
5 **Dynamic efficiency**	▪ Greater clarity across markets provides **more efficient price signals to support investment** (or avoid unnecessary exits)… ▪ …which is particularly relevant for investment in **flexible resources** needed to achieve Net Zero (avoiding complex bidding strategies across multiple markets) ▪ Facilitates **entry and participation** of smaller / new players	⬆	

Key:

⬆⬆ Significant impact

⬆ Moderate impact

⬌ Minimal impact

⬇ Moderate impact

⬇⬇ Significant impact

Summary of risks of co-optimising energy and AS

	Description	Perceived drawbacks	Reasons for optimism & supporting evidence	Evaluation
①	**Mathematical complexity / feasibility**	▪ Co-optimisation can (non-linearly) increase the **computational complexity** as more variables are added… ▪ …and optimising the objective function becomes more challenging ▪ **Time to solve** the problem and associated **practicability** may be perceived as issues (e.g., lessons from Euphemia) ▪ **Non-linear optimisation** (if needed) may be challenging; interdependence between markets an additional issue	▪ Many of the computational challenges already exist and can be resolved through linear programming, advanced algorithms and improved hardware ▪ Changes to bidding formats and the pricing rule can reduce computational complexity as well	↓
②	**Transparency**	▪ Perception that co-optimised markets are **less transparent** than individual markets as players may not initially understand the optimisation results and allocation	▪ Perceptions of reduced transparency can be alleviated through educational initiatives, stakeholder engagement, ‘shadow’ co-optimisation prior to go-live and comprehensive market reporting… ▪ …these perceptions are therefore likely to be transitory ▪ PJM initially selected pricing method which did not consider AS, and participants could not understand how prices reflect the actual marginal cost, but was later able to remedy the situation[1] ▪ Co-optimisation outcomes via algorithm are auditable	↔
③	**Reduced revenues / profits**	▪ Co-optimisation can **redistribute revenues and costs**, potentially leading to lower profit/revenue outcomes for certain participants ▪ Perception that **ESO may be less efficient than self-scheduling**, particularly for assets with complex bidding strategies (e.g., batteries) – e.g., **innovation can be stifled** if participants are **limited by the formats of the bids** they are allowed to submit ▪ Investments in developing sophisticated bidding strategy capability may become redundant or at least less useful	▪ Stronger competition and improved price signals are expected to lead to lower overall costs to consumers – some of the transfer away from providers of energy/AS is desirable. ▪ Strong stakeholder engagement and testing required to define appropriate products / bid formats - US ISOs are actively exploring options to continue innovating[2] ▪ Overall revenue predictability is likely to improve with co-optimisation, thereby reducing associated market risks.	↔
④	**Operational challenges for ESO**	▪ **Augmented IT and software needs** under co-optimised markets than under current separate markets ▪ **Upskilling and additional staffing** likely required in order to meet additional (and more complex) responsibilities ▪ Potentially increased **reputational risk to ESO** due to the potential for broader implications from operational errors	▪ IT investments in the status quo likely to grow (so incremental IT costs may be limited, particularly in long run as IT costs to manage non-co-optimised system would likely grow, too) ▪ Skills demands mitigated by training and hands-on experience ▪ Some of the costs would be offset by reduced operational challenges for market participants (who save effort on selecting which market(s) to participate in)	↔

Key:

⇈	Significant impact
↑	Moderate impact
↔	Minimal impact
↓	Moderate impact
⇊	Significant impact

Sources: (1) PJM, Proposed Enhancements to Energy Price Formation, 2017 (link); (2) Singhal, N.G.; Ela, E.G. Storage Integration Efforts in the U.S. Wholesale Electricity Markets IESO Energy Storage Design; Project; EPRI Institute: Palo Alto, CA, USA, 2020 (link)

Features of co-optimisation under different market designs

Co-optimisation of energy and AS can be implemented under different specific market designs, for example the extent of locational granularity

Co-optimisation under varying market designs

- Co-optimisation of energy and AS is commonly implemented in a centralised market; and the simultaneous running of markets is a necessary market design elements required for co-optimisation between energy and ancillary services
- However, within these limits, it is possible to implement co-optimisation under a range of designs, for example:

Locational granularity

Co-optimisation has been implemented under a range of different locational designs:

- **National:** Greece had co-optimised the procurement of energy and reserve under a national design
- **Zonal:** The Australian NEM co-optimises the procurement of frequency response and energy under a zonal design
- **Nodal:** Most US ISOs and New Zealand co-optimise between energy, frequency response and reserves under a nodal design[1]

A nodal design is the most common locational market design under which co-optimisation has been implemented.

Timing

Markets which are co-optimised must occur simultaneously, however the timing of the market relative to real time can vary:

- **Day ahead:** In ERCOT, energy and AS are currently procured and co-optimised at the day ahead stage only
- **Real time:** In the Australian NEM and in New Zealand, procurement and co-optimisation between energy and AS occurs in real time only
- **DA & RT:** Most US ISOs, for example PJM, MISO and NYISO, procure and co-optimise both at the day ahead and real time stage

The general trend gravitates towards combining day-ahead and real-time co-optimisation.

Notes: (1) AS service requirements are typically determined per predefined zones (a grouping of nodes) in nodal market designs. See also the NYISO case study in Appendix 2.

Co-optimisation could be implemented with national wholesale pricing or more granular locational pricing, such as a zonal or nodal design

National wholesale price

- Under a wholesale market design, the DA energy and reserve markets **are cleared ignoring transmission constraints**
- In real time (RT), transmission constraints are introduced and the day ahead schedule is **adjusted to reflect the operational constraints of the network**, in addition to any changes in system conditions since day ahead

Wholesale pricing under a locational design

- Under a locational design, transmission constraints are to a **larger extent** (zonal) or **entirely** (nodal) **incorporated** into both the DA and the RT markets…
- … resulting in the potential for different prices per price zone (zonal) or at each node (nodal)
- Under zonal design, balancing actions are **still required** in RT but to **lesser extent than under national wholesale price**
- Under nodal design, the **RT market is solely used to adjust the DA schedule** for changes in system conditions between **DA and RT**

Co-optimisation of energy and AS improves the allocation of resources **between AS and energy markets** under national design…

…and introducing, <u>in addition</u>, locational granularity is expected to **further improve the geographical allocation of resources** in each market (e.g., improved regional distribution of reserve commitments relative to national design)

Impact of electricity market design on co-optimisation of energy and AS

We consider the impact of moving from a **single national price to locational pricing on the benefits of co-optimisation specifically** (**without considering** the wider merits and risks of moving to locational pricing) and **identified four main areas**:

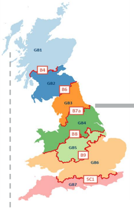
1. Price formation & competition
2. Efficiency of scheduling and dispatch
3. Volume of AS required
4. Differences in risks

The benefits of greater competition and price consistency across energy and AS are likely to be positively impacted by the choice of locational design

Co-optimised national design

Two key benefits of co-optimisation are:

  - **Increased competition within markets:** Resources are able to enter all co-optimised markets simultaneously, resulting in deeper markets and a lower risk of undersubscribed markets
  - **Improved price consistency:** Without co-optimisation, there is a potential disconnect between the prices across energy and AS markets, resulting in inconsistent price signals across markets; co-optimisation reduces this inconsistency across markets
- These benefits will **likely be realised under national design**, given they are inherently driven by the independent and sequential nature of markets operated without co-optimisation



Co-optimised locational design

There may be **additional competition and price consistency benefits** from moving from national to a locational market design.

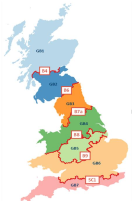
- Price formation becomes more localised and more accurate. Each zone or node will have its price based on more local supply-demand dynamics, grid constraints, and the co-optimised values of both energy and ancillary service
- Competition: Prices in locations with higher variability or renewable integration might place a higher value on ancillary services. Co-optimising energy and ancillary services in these regions can lead to price signals that incentivise resources with flexible operational characteristics.
- On the flip side, specific locations with limited competition or significant constraints might be susceptible to localised market power (although this would have also manifested itself in the national design, through the BM)

① **Price formation & competition** ⬆ **Move to locational pricing would bring more granular price formation with the potential to increase competition. Overall, we believe benefits are likely to be moderately positive.**

Due to the consideration of transmission constraints, co-optimisation under a locational design (particularly DA) will typically result in more efficient dispatch

Co-optimised national design

- Without consideration of transmission constraints at the day ahead stage, as under the national design, the co-optimisation process could commonly generate a schedule that is technically infeasible
- For example, resources are scheduled in the energy market, or for reserves, despite being sited behind a binding constraint
- As such, **when considering the realities of the network, the co-optimisation process under a national design will likely not result in an allocatively efficient outcome** and resources are not scheduled where they are most valuable
- The real time market is then used to incrementally alter the day ahead schedule to ensure a dispatch that adheres to transmission and other operational constraints
- This adjusted day ahead schedule generated in real time will likely result in a dispatch that is less efficient than co-optimising energy and AS with consideration of transmission constraints from the outset at the day ahead stage in addition to significant constraint costs incurred in real time



Co-optimised locational design

- **Consideration given to transmission constraints** when clearing co-optimised markets ensures that resources are allocated based on both the energy and ancillary service requirements of different locations in the grid. This minimises the inefficiencies arising from broader, averaged price signals in a national price system, and it ensures resources are assigned to the market where they deliver the most value.
- **With a locational design, the co-optimisation process at the day ahead stage will result in a schedule that is allocatively more efficient** when considering transmission and other operational constraints.
- **Balancing actions required in real time will also be significant reduced.** In a nodal market, real time actions are only required to adjust for changes in system conditions between day ahead and real time.
- Locational pricing can therefore significantly **increase the efficiency of scheduling and dispatch**

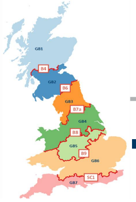
② **Efficiency of scheduling and dispatch**

From a dispatch efficiency perspective, the transition to locational pricing is likely to bring significant positive benefits due to the increased allocative efficiency of dispatch, with reduced need for real-time actions

More granular locational markets can also help to reduce the volume of AS required and increase the strength and precision of price signals

Co-optimised national design

- **Co-optimisation under national design** commonly generating schedules which are, in practice, not technically feasible
- This outcome results in a **potential need for ESO to procure more AS capacity** than it otherwise might need:
  - In real time, ESO is required to adjust the day ahead schedule to reflect constraints in the network, resulting in significant balancing actions
  - The flexible resources used for balancing actions in real time are typically also those that can provide AS capacity (for example, under today's market design ESO may choose to use reserve capacity to mitigate transmission constraints)
  - The **increased need for additional flexible resource in real time may then, in turn, increase the need to procure more AS capacity** in the DA market to ensure sufficient provision
- This will likely increase prices in the AS markets (to meet the higher demand) and, in turn, increase prices in the energy market, as more market participants choose to enter the AS markets and a greater volume of capacity is sterilised from the energy market



Co-optimised locational design

- The co-optimisation of energy and ancillary services in a locational framework ensures that less (zonal) or no (nodal) resources are procured that will not be able to deliver the respective service in real time due to grid constraints. This integrated approach can lead to a more efficient procurement strategy, reducing the risk of procuring unnecessarily high volumes of AS (including in locations where those assets cannot be utilised).
- A further benefit of locational pricing is the **potential for more spatially granular AS procurement and prices**, which better reflect the true value of AS capacity.
- This can be achieved through the introduction of **locational or nested AS procurement targets** - for example, separate AS targets could be set for Scotland and the rest of GB, to ensure sufficient AS capacity is sited below the B6 constraint (which frequently binds during periods of high wind) – see Appendix 2 for an example of NYISO's locational reserve procurement.

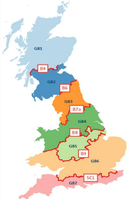
③ **Volume of AS required** ↑

Transition to locational pricing has the potential to have a moderate positive influence of reducing the volume of AS procured

National may appear to offer a simpler co-optimisation process, however the need to take significant balancing actions in real time increases complexity

Co-optimised national design

- Co-optimising under National may appear to be lower risk, as there are **fewer variables to solve** at the day ahead stage (i.e., transmission constraints are not considered in the optimisation process)
- However, this complexity is not avoided under national design – it is only shifted later into the real time market, where balancing actions are required to ensure the dispatch adheres to the constraints of the network
- Indeed, the **tighter timescales** associated with real time may result in a **more challenging optimisation process**…
- … and a higher risk to system security due to the lack of time available to mitigate issues with the optimisation process if they do occur (unlike at day ahead)



Co-optimised locational design

- There is a risk that **locational pricing with co-optimisation may be perceived as excessively complex**…
- … with separate prices for energy and AS at potentially a dozen of zones or thousands of nodes
- However, there is significant **precedent for co-optimised locational markets** internationally (more so for nodal than zonal or national co-optimised designs)
- Similarly, under a locational design, the co-optimised schedule incorporating transmission constraints is calculated at the day ahead stage, reducing the impact of any possible issues with the optimisation process
- However, at the same time, from a market participant perspective standard hedging strategies may be less effective due to the complexities of locational pricing, requiring more complex risk management tools.

④ Differences in risks

↑ **Reduced complexity in real time and greater international precedent**

Co-optimisation can leverage information at DA, RT, or both stages, to manage risks, ensure system reliability, and minimise total system costs

- Co-optimisation timing of wholesale energy and ancillary services markets plays a pivotal role in balancing costs and system reliability, yet it introduces complex challenges that vary depending on the approach taken.
- Different jurisdictions have adopted varied approaches to co-optimisation based on their unique operational characteristics, historical developments, and the evolving needs of market participants. We assessed the strengths and weaknesses of each approach in the figure below:

Timings co-optimisation of wholesale energy and ancillary service markets

	Day Ahead only[1]	Real Time only	Day ahead and Real time
Strengths	**Predictability**: Market participants have a clearer picture of upcoming commitments **Stability**: It offers a certain degree of financial and operational predictability for market participants	**Responsiveness:** Able to handle variations in supply and demand in RT **Optimal Resource Use**: Ensures resources are scheduled and dispatched based on actual, not predicted, conditions	**Balance of Predictability and Responsiveness**: Combines benefits from initial day-ahead planning and flexibility from real-time adjustments. **Better Resource Allocation**: Enables efficient use of resources by considering wider pool of resources.
Weaknesses	**Forecast Errors:** Relying solely on DA forecasts can lead to discrepancies due to forecast errors **Reduced Flexibility**: Managing unexpected changes in demand or supply close to RT might be more challenging **Possibly Higher Costs**: Deviations from conditions expected at DA stage can result in higher system operational costs and potential inefficiencies **Diluted investment incentives:** reduced incentives for investment for flexible capacity	**Operational Challenges:** Managing the grid in real-time without day-ahead commitments can be complex (i.e. compressed timeframe for decision-making) **Limited Participation**: Some participants, with operational constraints (e.g. longer ramp-up rates), might find it challenging to participate **Possibly Higher Costs**: The reduced pool of resources and higher reliance on resources with fast ramping can lead to higher costs	**Communication and Data Needs**: Higher communication and data processing capabilities are needed to manage both time frames effectively, however the increment is likely to be marginal

- While each approach has its merits and challenges, the general trend among modern electricity markets gravitates towards combining day-ahead and real-time co-optimisation to harness the advantages of both while mitigating their individual limitations (e.g. ERCOT –see case study in Appendix 2).

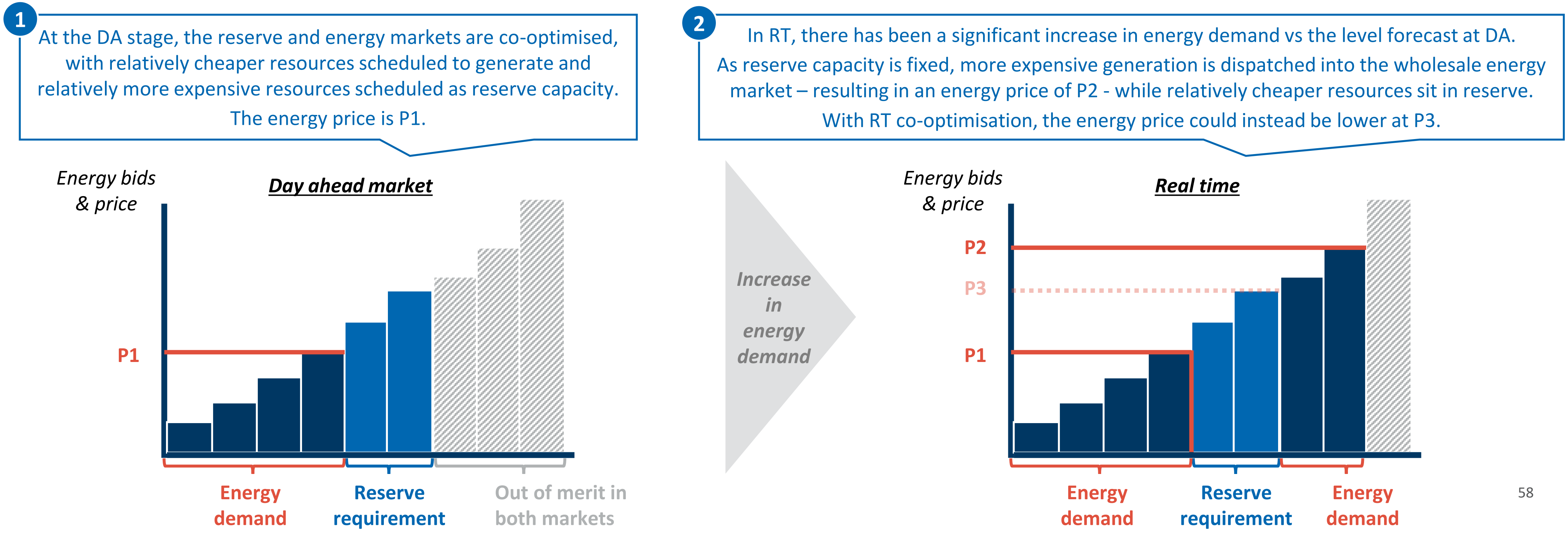
Notes: (1) This report does not examine the optimal frequency and timing of co-optimisation; however, we recognise that there is a tendency towards greater intraday re-scheduling opportunities, see e.g., GE Energy Consulting (2014). "PJM Renewable Integration Study. Tasks 3B & 4. Market Analysis and Mitigation." March 31, 2014. (link)

Based on the experience of ERCOT, co-optimising at the day ahead stage without the ability to re-co-optimise in real time risks creating inefficiency

Co-optimisation in ERCOT

- ERCOT currently co-optimises frequency response and reserve services with energy in the day-ahead market only
- Therefore, in real time, the energy market is optimised taking the DA reserve outcome as a fixed input
- However, this design has been identified as driving inefficiency in dispatch outcomes and driving up energy prices…
- … in particular when system conditions in real time vary materially from those expected at day ahead, which is increasingly the case as the share of variable renewable generation increases
- As such, ERCOT has initiated a project to implement Real Time Co-optimisation (“RTC”) by 2026 (see Appendix 2)

Stylised example – inefficiency driven by lack of real time co-optimisation



① At the DA stage, the reserve and energy markets are co-optimised, with relatively cheaper resources scheduled to generate and relatively more expensive resources scheduled as reserve capacity.

The energy price is P1.

② In RT, there has been a significant increase in energy demand vs the level forecast at DA.

As reserve capacity is fixed, more expensive generation is dispatched into the wholesale energy market – resulting in an energy price of P2 - while relatively cheaper resources sit in reserve.

With RT co-optimisation, the energy price could instead be lower at P3.

Co-optimisation of additional services – inertia and voltage

In addition to reserve and response, there could theoretically be increasing value in co-optimising the procurement of inertia and voltage with energy…

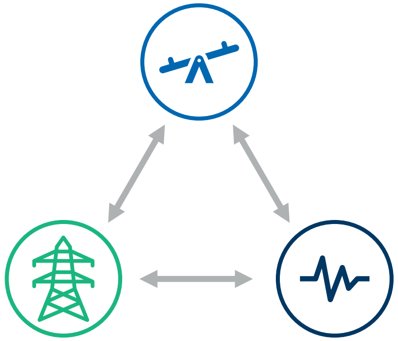
Voltage and inertia are becoming more challenging to manage as the system decarbonises[1,2]

- **Historically, inertia levels were high** in the GB power system, as synchronous generation (such as coal and gas) dominated the generation mix providing the service as a by-product without specific renumeration
- However, inertia **levels have fallen** as renewable generation, which does not provide inertia to the same extent, has grown
- The **ESO is now at times required to intervene to dispatch out of merit units to maintain a minimum inertia level**, with the costs associated with doing so expected to rise
- Likewise, as the prevalence of distributed generation has grown and the share of generation from traditional thermal generator falls, **system voltage levels are increasingly challenging** for ESO to manage…
- … with **high voltages during periods of low load a particular challenge**
- This can manifest itself in how the ESO manages the system in real time: for example, the schedules produced (with or without co-optimised energy, reserve and response) may be upended by meeting inertia or voltage requirements.
- In response to these challenges, **ESO has recently been exploring new options to procure both inertia and voltage support services**, initially through the stability pathfinder projects and more recently through the development of voltage and stability markets

Co-optimisation of inertia and voltage with energy and other AS could theoretically provide benefits

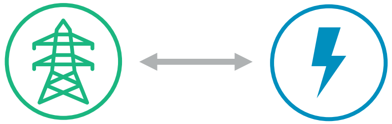
- Given the interrelationships between inertia, voltage energy, and other AS, there could be value in co-optimising its procurement:

Inertia:



- **Inertia** and **frequency response:** the level of inertia on the network determines the RoCoF (the sensitivity of frequency to imbalances in generation and load)
- **Inertia** and **energy:** Thermal synchronous generators provide inertia as a by-product of generating energy

Voltage:



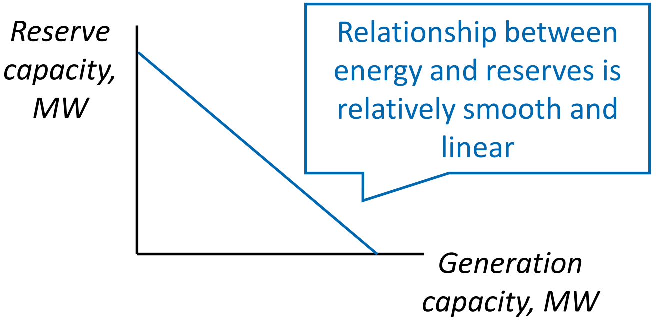
- **Energy** and **voltage:** voltage support services have traditionally been supplied by large synchronous thermal generators via an obligatory service (in addition to transmission network assets), with requirements impacted by load on the transmission network. However, ESO is exploring options to attract new providers of voltage services (e.g. Pathfinders)

Notes: (1) Source: REMA Consultation 2022, p34-36 (link); (2) Source: ESO, Reactive Power Market Design – Innovation Project Summary, 2022 (link)

…however, the technical characteristics of inertia differ from the AS that are commonly co-optimised, which may present challenges

Interrelationship between energy, reserves and frequency response services

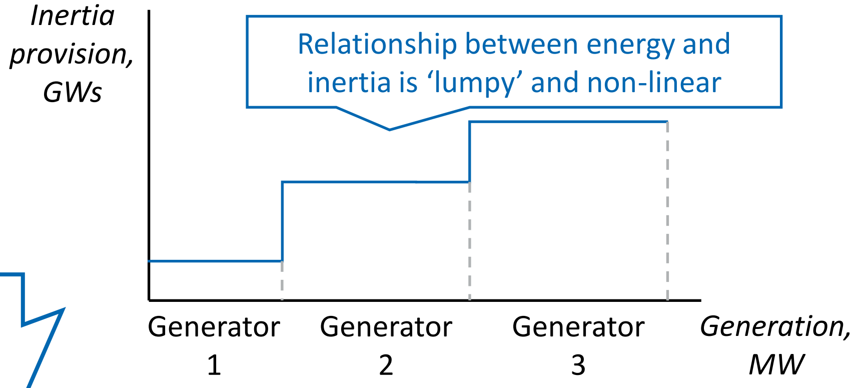
- Broadly, **there is a 1 to 1 negative relationship between providing energy, reserve and frequency response** services…
- … meaning that providing an additional 1 MW of reserve reduces the ability of a generator to provide energy or frequency response by 1 MW
- The relationship between each is broadly linear[1]
- As such, centralised dispatch engines (which commonly rely on linear programming techniques) are able to optimise the allocation of resources across markets by considering this direct trade-off between each



- Policy makers in the Australian NEM have considered the potential to co-optimise inertia with energy and frequency response products…
- … however, the system operator considers this to be technically infeasible using the NEM's existing dispatch engine (see Appendix 2)
- **Implementing full co-optimisation of an inertia spot market with energy and other AS in GB would represent breaking new ground internationally and would likely face significant technical barriers**

Co-optimisation of inertia with energy and other AS

- Unlike energy, reserves and response, **there is not a linear 1 to 1 relationship between inertia and energy**:
  - Inertia provision is typically 'all or nothing' depending on whether the generator is generating/synchronised…
  - … meaning once synchronised, increasing generation by 1 MW has no impact on the level of inertia provision
- As such, the relationship between energy and inertia is non-linear and there is no direct trade-off between supplying one or the other…
- … which means the linear programming methods commonly utilised in dispatch engines may be unsuited to optimising the provision of non-linear services

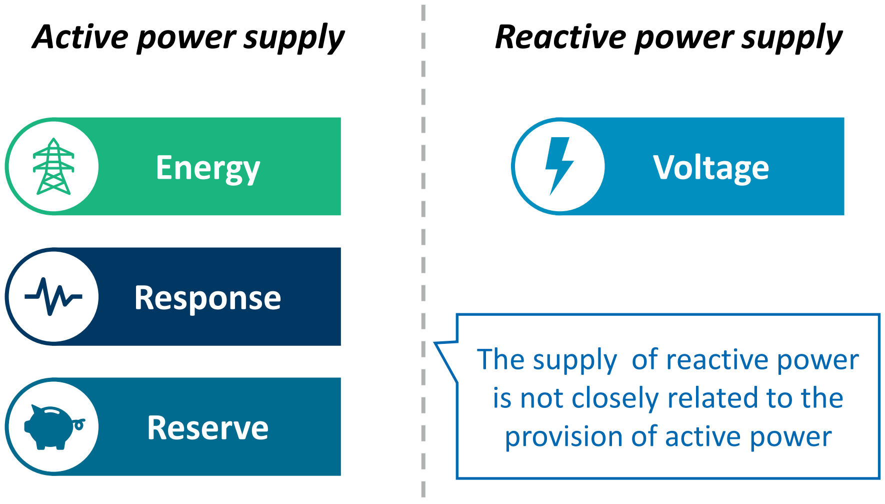


Notes: (1) As discussed on slide 42, a degree of non-linearity exists for the provision of energy, reserve and response. However, the extent of this non-linearity is lower for these service and are currently co-optimised in markets utilising linear dispatch engines.

Likewise, the technical characteristics of voltage vary significantly from existing co-optimised AS, increasing the challenge of co-optimising

Voltage support services

- The provision of wholesale energy, along with the supply of reserve and response ancillary services, concern the delivery of active power
- **Voltage support services, however, require the injection or absorption of reactive power** – a related but separate form of power
- The key challenges related to a potential co-optimisation are as follows:
  - There is **no straightforward (linear) relationship** between the delivery of active power (energy, response and reserve) and the delivery of voltage support services
  - Like inertia, it therefore suggests **dispatch engines that utilise linear programming will be poorly suited to co-optimising the procurement of voltage** with energy and other AS
  - Furthermore, unlike active power and inertia, **reactive power is highly locational** – it does not travel far through the transmission network, meaning balance must be maintained at each location of the transmission network…
  - … with specific requirements and needs difficult for ESO to define in advance
- This complexity is one of the reasons **competitive or market-based procurement of voltage services is not prevalent across power markets globally** relative to other AS (instead obligations are often placed on resources to provide the service as part of connection agreements, for which they are remunerated)
- This is despite a longstanding need for the services (unlike inertia, where procuring additional volumes has only more recently become a challenge commonly faced across power markets)



The supply of reactive power is not closely related to the provision of active power

Given these technical characteristics, it appears a **voltage support market would be highly challenging to co-optimise with energy and other ancillary services** and would be breaking new ground internationally

Appendix 1 – co-optimisation under central or self-scheduling

Central scheduling is commonly in place in systems that co-optimise energy and AS (response and reserves) and expected to facilitate its introduction

The merits of moving from GB's **current self-scheduled market design** to a **centrally scheduled** one are **not examined** in this report. (This potential reform is being considered by the Department for Energy Security and Net Zero ("DESNZ") as part of its Review of Electricity Market Arrangements ("REMA")).

For the purposes of this report, we consider that a potential transition to a **centrally scheduled market design is likely to facilitate the introduction of the co-optimisation of energy and AS markets (response and reserves)** for two reasons:

- **Bidding complexity.** Under self-scheduling, market parties that operate resources that can be active in both the energy and AS markets need to formulate bids for these markets separately. However, the bids need to be interlinked, i.e., for a particular market time unit, if a bid is accepted in the AS market, the volume of electricity that can be sold in the energy market is impacted (and vice versa). Providing such explicit linkages between the energy and AS bids while having the bids reflecting the technical characteristics of the resources is extremely complex.
- **Improved SO visibility.** Under central scheduling, the system operator has significantly improved visibility about the technical characteristics of all resources. This visibility allows for a more efficient allocation of the resources between energy scheduling and AS provision needs.

N-Side's technical assessment discusses in more detail how bidding languages in centralised and self-scheduled markets differ, and the consequent impact on co-optimisation.[1]

It is important to note that **power markets with co-optimisation** of energy and AS in place (at the **DA and RT time frame**) **commonly have a centralised market in place.** Often (but not always), co-optimisation is also accompanied by **nodal pricing:**

DA & RT co-optimisation with nodal pricing: Most US ISOs, for example PJM, MISO and NYISO

Day ahead co-optimisation with nodal pricing: ERCOT

Real time co-optimisation with zonal pricing: Australian NEM

Real time co-optimisation with nodal pricing: New Zealand

Notes: (1) See also the explanatory note from N-side about the SDAC MSD Co-optimization Roadmap Study from 20/10/2020 (link)

Appendix 2 – international case studies

MISO: Implementation of co-optimised AS markets has generated significant cost savings despite the upfront costs

Market background (status 2022)

- Midcontinent Independent System Operator (“MISO”)
- Nodal market
- Centralised scheduling
- 190 GW installed capacity, connecting 15 US states and Canadian province of Manitoba
- Generation mix: 33% coal, 33% gas, 16% wind, 14% nuclear, 2% hydro, and 1% solar.[4]

Implications for GB

- Co-optimised energy and AS markets was implemented as part of a larger reform, hence the substantial implementation costs.
- The MISO experience indicates that the **cost savings** generated by co-optimisation **are significantly larger than the implementation costs**

Throughout the 2000s the MISO market was introduced, with co-optimisation of AS and energy implemented from 2009

2005

- MISO **energy markets are implemented**
- Day ahead and real time markets for energy only
- Balancing and AS needs met by 26 separate balancing authorities across the MISO area

2009

- Separate balancing markets are replaced by consolidated **MISO-wide AS markets**
- Three AS markets covering frequency response and reserves
- Markets are **co-optimised with energy**

Present day

- Markets have been reformed as MISO gained experience and the system evolves
- **5 AS products** are now procured covering **frequency response, reserve and ramp capabilities**[2]

The reform has generated significant cost savings across the market, despite the upfront investment

Implementation cost

- Consolidating MISO’s balancing areas was one of the most significant investments in MISO history
- Cost of c.$75m (not inflation adjusted) – of which a small proportion related to co-optimisation software

Key benefits

- Consolidation helped to reduce reserve requirements significantly[3]
- More efficient scheduling and dispatch as a result of co-optimisation between AS and energy

Annual savings

- MISO estimates annual savings of at least $60m in AS costs alone as a result of the introduction of co-optimisation (indirect energy costs savings are not evaluated)

Notes: (1) Y. Chen (2019) MISO Market Design Overview and Evolution (link); (2) MISO (2022) Business Practice Manuel, Energy and Operating Reserve Markets (link) page 45; (3) MISO (2019) MISO Energy and Ancillary Services co-optimisation ERCOT RTC Lessons Learned engagement (link), (4) MISO fact sheet, (link).

ERCOT: Planned co-optimisation of energy and AS in real-time, in addition to existing day ahead co-optimisation, is forecasted to deliver significant benefits

Market background (status 2022)

- Electric Reliability Council of Texas ("ERCOT")
- Nodal market
- Centralised scheduling[1]
- Circa 98 GW peak capacity
- Generation mix: 42% gas, 29% wind, 11% coal, 11% solar, and others[2]
- ERCOT currently co-optimises frequency response and reserve services with energy in the day-ahead market only
- The current ERCOT market does not have locational reserves

Implications for GB

- Cost Benefit Analyses for ERCOT suggests **significant benefits of real-time co-optimisation**
- These benefits would accrue both to consumers, in the form of lower energy bills, and the SO, through improved system reliability – producer costs also fall
- However, uncertain share of the total benefit would likely arise as a result of introducing a RT reserve market itself, even absent co-optimisation with energy

Introducing real-time co-optimisation (RTC) is forecast to deliver significant benefits across the ERCOT system

- ERCOT currently **co-optimises frequency response and reserve** services with energy in the **day-ahead market** only. It relies on a scarcity price adder mechanism to reflect scarcity of reserve capacity in the real time energy price.[3]
- In **real-time,** adjustments to day-ahead market schedules are currently made via a Supplemental Ancillary Services Market ("SASM") process although this is a thin market often resulting in anomalous prices for replacing reserves.
- ERCOT has been considering the merits of introducing real-time co-optimisation ("RTC") of energy, response and reserve (as requested to do so by the Public Utility Commission of Texas). A study performed by the Independent Market Monitor for ERCOT, estimated that RTC would provide significant benefits (see table below) and, as such, ERCOT has initiated a project to implement RTC by 2026.[5]

Key estimated costs and benefits of implementing RTC in ERCOT

Annual benefits[4]

1. $4/MWh reduction in average energy prices – reducing annual costs by $1.6bn
2. $260m reduction in congestion costs
3. $155m reduction in Ancillary Service costs
4. $12m reduction in production costs
5. Significant improvement in system reliability

Costs

1. Implementation costs of approximately $50m from 2023 to 2026
2. Impact of implementation efforts on staff availability in the context of existing critical workloads

Notes: (1) ERCOT (2023) Item 6: Real-Time co-optimisation (RTC) Update (link); (2) ERCOT Fact sheet, (link); (3) Under potential RT co-optimisation, scarcity pricing would likely remain, but may be implemented in a different way through reserve shortage pricing. (4) Benefits are based on a simulation of the year 2017, see Potomac Economics (2017) Simulation of real-time co-optimisation for 2017 (link). (5) Real-Time Co-optimization (RTC) Update, June 2023 (link).

NYISO: A nested reserve system enables the dispatch process to distribute reserves across the system reflecting expected operational needs

Market background (status 2022)

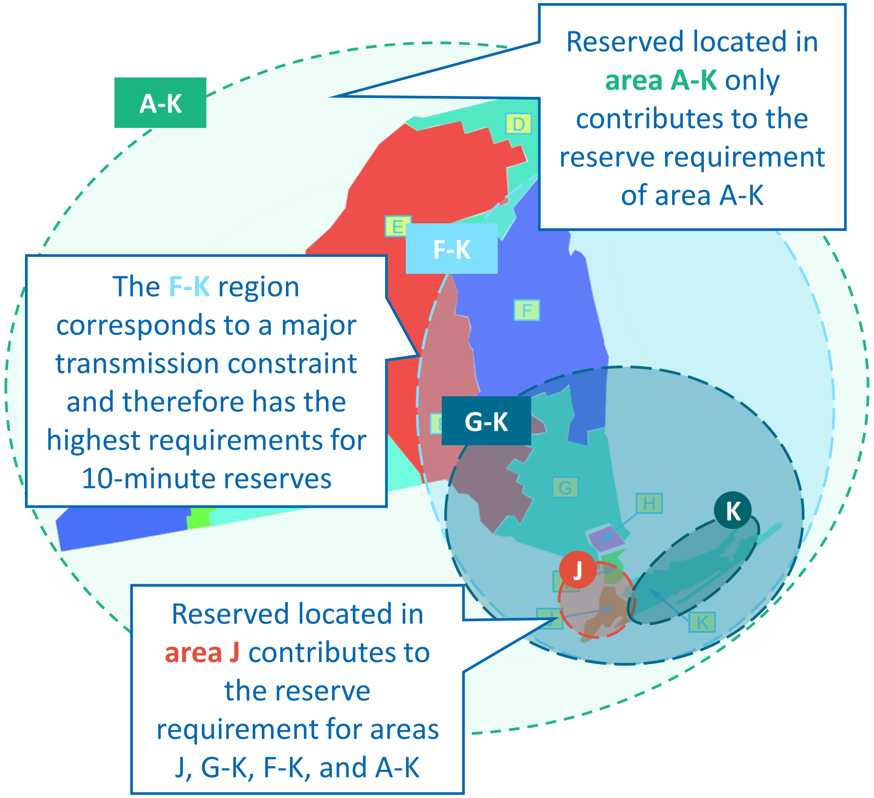
- New York ISO ("NYISO")
- Nodal market
- Centralised scheduling
- 38 GW installed capacity [1]
- c. 70% fossil fuel generation, 10% nuclear, 15% hydro, and 5% wind[1]
- NYISO co-optimises energy, operating reserves, and frequency (termed 'regulation') at DA and in RT
- Co-optimisation is combined with reserve shortage pricing

Implications for GB

- Such a strategy can also be applied to similar areas with high load, but insufficient transmission connectivity
- For example, ESO could define a GB wide reserve procurement target, and a further target for south of the B6 Tx constraint (which frequently binds during periods of high wind) to ensure sufficient reserve capacity is not restricted by this constraint if a large loss occurs south of B6
- This could help to **prevent ESO from having to undertake expensive actions in real time to constrain on additional resources in the south**

Co-optimisation can also be combined with nesting to account for transmission constraints

- NYISO currently employs a 'nested' co-optimisation system, where it:
  - co-optimises energy, reserves and frequency response, and;
  - accounts for contingency based reliability requirements for reserves by setting aside reserve capacity in the areas that would meet the reliability targets.
- Reserve requirements are defined for different areas, some of which are entirely self-contained within other areas…
- … meaning a single resource can contribute towards the reserve target of multiple areas
- The design ensures that reserves are located to meet post contingency reliability requirements, with high penalty prices for not meeting those requirements
- Reserves sited in zones that are most transmission-constricted will typically receive a higher price, reflecting their greater scarcity and, therefore value



Reserved located in **area A-K** only contributes to the reserve requirement of area A-K

The **F-K** region corresponds to a major transmission constraint and therefore has the highest requirements for 10-minute reserves

Reserved located in **area J** contributes to the reserve requirement for areas J, G-K, F-K, and A-K

Illustration of nesting: the letters A-K refer to the zones within the NYISO. The diagram demonstrates that zones J and K (both high load zones relative to the rest) have been allocated reserve capacity. This can be accessed by the rest of zones A-K, but their presence in J and K, respectively, prevents a situation in which those zones are not able to access reserves due to transmission constraints.

Notes: (1) Potomac Economics (2023). 2022 State of the Market Report NYISO (link)

Introducing a co-optimised inertia spot market may not be technically feasible within the NEM’s existing dispatch engine

Background

- National Electricity Market (“NEM”)
- Zonal market - 5 interconnected regions
- Centralised scheduling
- Frequency response products are co-optimised with energy
- Co-optimisation occurs in real time only
- The NEM has no operating reserve products
- There is no day ahead market

Implications for GB

- The **co-optimisation of inertia** via a spot market would be breaking new ground
- The technical characteristics of inertia mean that co-optimisation may not be feasible under the linear programming techniques commonly used
- However, some of the benefits of full co-optimisation of inertia may be achievable via alternative mechanisms

Potential for full co-optimisation of an inertia spot market has been examined in detail in the NEM context, but so far policy makers have favoured alternative approaches

Context

- As with many electricity markets, the NEM is faced with decreasing inertia levels and has been experiencing deteriorating frequency performance.
- Policy makers have expressed an **ambition to move procurement of inertia closer to real time and towards a spot market**, if possible.[1]
- FTI previously examined the concept of using a demand curve to set prices for services such as inertia in a comprehensive report for the Energy Security Board and identified options for operationalising the approach.[2]

System operator concerns

- However, **the system operator is concerned that it may not be technically feasible to create an inertia spot market that is co-optimised** with energy and frequency response:[3,4]
  - The NEM’s dispatch engine is based on linear programming and uses shadow pricing for frequency response services (i.e. price is equal to the marginal cost associated with the frequency response constraint)
  - **Inertia supply is binary by nature** and is **not impacted by varying the level of generation once operating** – unlike frequency response and reserves
  - As such, **the pricing of inertia leads to additional complexities**

Alternative proposal

- Instead, the system operator (“AEMO”) is supportive of the introduction of an Operational Security Mechanism (“OSM”) to provide similar benefits
- This would procure and schedule system security services (including inertia) via a separate mechanism running parallel to the energy and frequency response dispatch – with outcomes across markets optimised in real time.

Notes: (1) AEMC, Efficient provision of inertia - consultation paper March 2023 (link), p12; (2) FTI Consulting: Essential System Services in the National Electricity Market, 2020 (link), ¶7.52 to 7.67. (3) AEMO, Submission to Efficient provision of inertia Consultation, March 2023 (link); (4) AEMO, Submission to Operational security mechanism, August 2020 (link), p14.

Experts with Impact™

FTI™ CONSULTING