Department for
Business, Energy
& Industrial Strategy

Evaluation of the TA – phase 4

Final report

May 2019

Contents

1. Introduction

SUMMARY

This report presents findings from Phase 4 (the final phase) of the evaluation of the Transitional Arrangements (TA) for Demand Side Response (DSR) in the Capacity Market (CM) for electricity. It examines delivery obligations for the second TA scheme, focusing on load turn-down DSR only.

The evaluation was realist and theory-based: during Phase 4, we refined and revised theory developed during Phases 1-3. Phase 4 evidence, comprising National Grid data and a final wave of telephone interviews and email surveys with TA participants, was analysed alongside evidence from participants and non-participants in earlier phases of the evaluation.

This report presents findings from the evaluation of the second auction of the TA for DSR. This realist, theory-based evaluation was undertaken for the Department for Business, Energy and Industrial Strategy (BEIS) by CAG Consultants, in partnership with Winning Moves, Verco and NERA Economic Consulting. Findings from earlier phases of the evaluation are documented in the Phase 1 report[1] (findings about the first TA auction), the Phase 2 report[2] (findings about delivery obligations for the first TA scheme) and the Phase 3 report[3] (findings about the second TA auction). This Phase 4 report covers delivery obligations for the second TA scheme.

Research and policy background

As explained in the earlier reports, the TA formed part of the Capacity Market (CM). Like the CM, the TA aimed to support BEIS's objectives of promoting growth, decarbonisation and energy security, while ensuring affordability of the energy supply. Further details of the scheme can be found in Appendix 1 of this report and Appendix 1 of the Phase 2 report.

In particular the TA aimed to encourage the development of DSR to balance supply and demand in a decarbonised electricity grid[4]. This report uses the CM definition of DSR: the activity of reducing the metered volume of imported electricity of one or more customers below an established baseline, by means other than a permanent reduction in electricity

[1] https://www.gov.uk/government/publications/evaluation-of-the-transitional-arrangements-phase-1
[2] https://www.gov.uk/government/publications/evaluation-of-the-transitional-arrangements-for-demand-side-response-phase-2
[3] https://www.gov.uk/government/publications/evaluation-of-the-transitional-arrangements-for-demand-side-response-phase-3
[4] National Infrastructure Commission (2016) *Smart Power: A National Infrastructure Commission Report.* Available at: https://www.gov.uk/government/publications/smart-power-a-national-infrastructure-commission-report. Accessed 27/7/2016

use. Under this definition, DSR may be achieved through any combination of onsite generation, temporary demand reduction or load-shifting. We use the term 'turn-down' DSR to refer to the last two activities.

The TA scheme involved two auctions for specific types of capacity within the CM, the first for delivery of capacity in the 2016/17 delivery year[5], held in January 2016, and the second for delivery of capacity in 2017/18, held in March 2017. While the first TA scheme was open to all types of DSR and also to small-scale distribution-connected generation in Capacity Market Units (CMUs) of between 2 MW and 50 MW, the second TA scheme was only open to turn-down DSR and had a minimum CMU threshold of 500 kW[6].

The TA auctions were additional to the main CM auctions: the four-year ahead auctions (T-4) and the smaller one-year ahead auctions (T-1) which will deliver capacity from 2018/19 onwards, and the Early Auction (EA) which is delivering capacity in 2017/18. The main CM auctions are open to generation, storage and DSR capacity.

The main steps in the TA process for each 'Capacity Market Unit' (CMU) are outlined in Figure 1.1 below, with drop-out points shown in pink. The main CM auctions follow a very similar process. The grey steps were not observed during 2017/18 as there was no CM Notice or associated 'stress event' between 1st October 2017 and 30th September 2018.

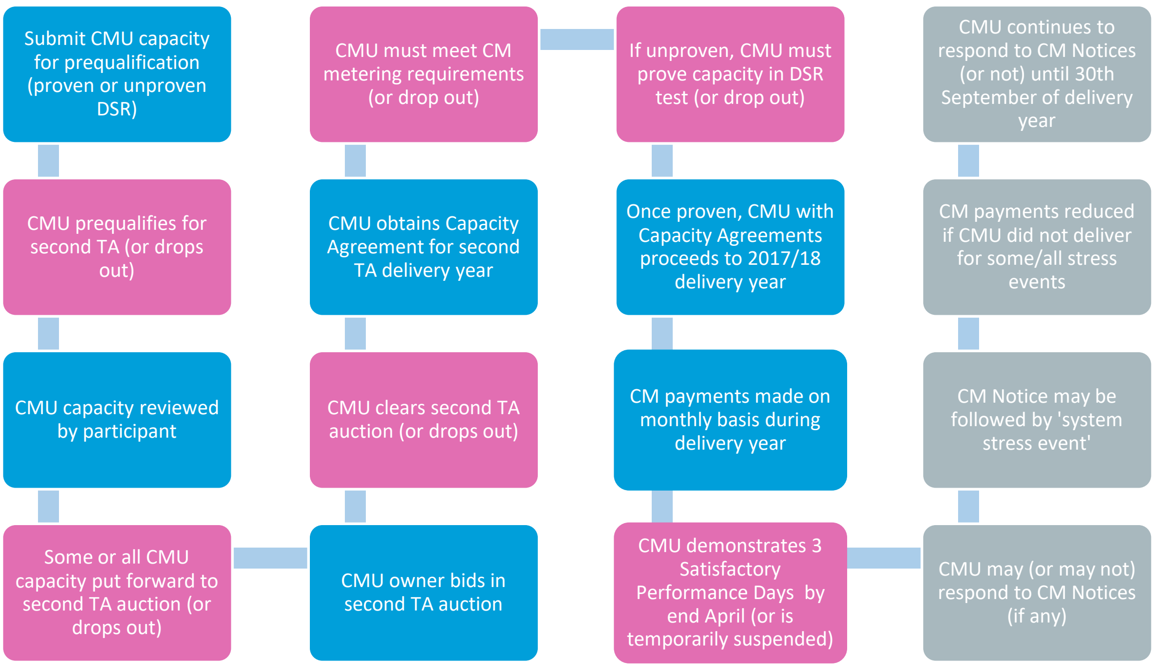

Figure 1.1: Main steps in process for second TA

Further details about the operation of the TA are given in Appendix 1, while a glossary of technical terms is provided in chapter 8 of this report.

[5] The delivery year runs from 1st October of one year through to 30th September of the following year.
[6] In both cases smaller 'components' could be aggregated within one CMU.

Evaluation aims and objectives

The second TA had two main objectives: to encourage turn-down DSR; and to contribute to the development of flexible capacity for the future CM. In contrast to the first TA, BEIS's aims for the second TA did not include a significant contribution to security of supply in the delivery year (2017/18), because short-term system tightness had already been addressed through the introduction of the EA alongside the TA. The objectives of the second TA scheme were therefore:

1. To develop a stock of flexible capacity[7] that could be available for future CM auctions, thereby contributing to competitiveness and liquidity in the CM.
2. To encourage enterprise and develop experience, confidence and understanding so that turn-down DSR will be able to realise its potential and ultimately compete with larger generation assets in the CM.

This evaluation was designed to answer five high-level questions (HLQs) posed by BEIS, in which the desired outcomes stem from the two objectives above. The evaluation also looked for any unanticipated outcomes of the scheme. This report presents findings that are relevant to all of the HLQs below.

- HLQ 1 - What outcomes can be attributed to the second TA and were they as intended by BEIS? What outcomes occurred for whom and under what circumstances?
- HLQ 2 - Through what levers and causal mechanisms has the second TA contributed to these outcomes and the variation by group and circumstance?
- HLQ 3 - Did the second TA represent good value for money to both scheme participants and the consumer?
- HLQ 4 - Which aspects of the second TA's design and implementation account for the findings of HLQ 2 and 3?
- HLQ 5 - What are the implications of the findings for the future contribution of turn-down DSR to the CM?

In addition to these HLQs, BEIS asked the evaluation to research the characteristics of turn-down DSR, including its costs and revenues. This fed into the assessment of HLQ 3 but was also designed to inform future analysis of turn-down DSR by BEIS.

[7] By flexible capacity, we mean electricity demand and generating capacity that is able to increase or decrease in response to signals, to help balance supply and demand of electricity across the GB grid. For the purposes of the TA, flexible capacity does not include electrical storage.

Evaluation design

Our approach to this evaluation was realist and theory-based. A realist approach[8] emphasises the importance of understanding not only whether a policy contributes to outcomes (which may be intended or unintended) but how, for whom and in what circumstances. The realist approach is explained further in the Phase 2 report, chapter 1.

The development of a 'theory' of the TA was central to implementing a realist evaluation as it allowed us to examine rigorously the design and execution of the scheme, and test policy assumptions against available evidence. We developed an initial theoretical framework for Phase 4 of the evaluation, as presented in Appendix 2, which set out the realist hypotheses that we tested against research evidence. The realist hypotheses set out for whom, and in what circumstances (i.e. in what 'contexts'), the policy was expected to lead to particular reasoning and choices being made (i.e. causal 'mechanisms' being activated[9]), leading to desired or undesired policy outcomes. These realist hypotheses are generally known as context-mechanism-outcome configurations or 'CMOs'[10]. Our theory included both 'additional' CMOs (which involved the second TA leading to desired outcomes) and competing 'non-additional' CMOs (in which observed outcomes were caused by other influences).

Realist evaluation uses the idea of generative causality (i.e. a mechanism or reasoning only fires when the contexts are right). In Phase 4, we used realist analysis to test our 'additionality theory' i.e. our CMO hypotheses about whether or not the TA made a contribution to its two objectives (as set out above) and whether this contribution was 'additional' to what would have happened in the absence of the second TA, or whether it happened for other reasons. We tested the strength of this evidence for specific cases using process tracing. Our approach to analysis is explained further in Appendix 4.

Methodology

The evidence that we have gathered during Phase 4, and against which the initial theoretical framework was tested, is set out in Table 1.1. The revised theoretical framework is presented in Appendix 3.

[8] R Pawson, R, and Tilley, N. (1997) *Realistic Evaluation*. London: SAGE Publications Ltd; and Pawson, R. (2006) *Evidence-Based Policy*. London: SAGE Publications Ltd.
[9] In realist terminology, the activation of a causal mechanism is referred to as the mechanism 'firing'.
[10] Definitions for contexts, mechanisms and outcomes are provided in the glossary. Further detail can be found in Pawson and Tilley (1997) (op cit).

Data collection

The evidence summarised in Table 1.1 involved research with direct participants that put their own capacity into the TA, and research with aggregators[11] who put forward capacity on behalf of other organisations.

We also interviewed a sample of eight aggregator clients[12], in addition to the ten aggregator clients interviewed during Phase 3. Across the two phases, we sampled clients purposively to cover the range of DSR asset types identified during Phase 3 analysis. Further detail on sampling is provided in Appendix 4.

The topic guides and email questionnaire for Phase 4 research were agreed in advance with BEIS. The topic guides were designed to test 'additionality theory'[13] in detail, explicitly testing theory hypotheses with interviewees. They also gathered insights into other areas of theory, without explicitly testing those theories with interviewees. This approach was chosen to prevent interviews exceeding one hour, particularly in the light of the number of times that TA participants had already been interviewed during the evaluation. The interviews built on information already available from National Grid data, from the CM registers for the TA and other CM auctions and from any earlier interviews with the same organisations in earlier phases of the evaluation. Tailored topic guides and email surveys were prepared for each interviewee, incorporating this prior information and highlighting priority questions to be probed.

[11] An aggregator is an intermediary organisation that provides a service of collating capacity for flexibility services from a range of other organisations, in return for a share in the revenues generated.
[12] Aggregator clients are the industrial or commercial organisations that put forward capacity via an aggregator.
[13] Additionality theory is explained on page 7: it sets out our hypotheses about whether (and how) the second TA contributed to its objectives in ways that would not have happened without the scheme.

Table 1.1: Summary of evidence gathered in Phase 4 of the evaluation

Evidence source	Phase 4 research tasks	Limitations
In-depth telephone interviews with TA participants (with accompanying email survey)	In April and May 2018 we undertook in-depth interviews with seven of the eight TA aggregators and two of the three direct participants that were awarded Capacity Agreements in the second TA auction. Two of the aggregators subsequently exited as they failed to pass testing requirements. We interviewed one of the drop-out aggregators and the remaining six that passed testing requirements and participated in the delivery year. We also interviewed one non-participant aggregator that submitted capacity via a participant aggregator. All of those going forward to delivery also completed an email survey, during either Phase 3 or Phase 4 of the evaluation.	This was the fourth round of interviews for TA participants, as there has been an interview round in each of the four phases of the evaluation. There were a couple of non-respondents but response rates were high (67-88%). Whilst all of the organisations completed at least part of the email survey, some did not provide all of the information requested. In particular, only seven TA participants (two direct participants and five aggregators) provided cost data at participant-level and CMU-level. The cost data was commercially sensitive and could not be shared with BEIS in disclosive form.
In-depth telephone interviews with aggregator clients (with accompanying email survey)	In April and May 2018, we also undertook in-depth interviews and completed email surveys for eight aggregator clients, from 21 client organisations contacted during Phase 4. These were identified via meter data provided by National Grid. One of the clients had been interviewed before during Phase 2 about the first TA, but there was no overlap with clients interviewed during Phase 3 about the second TA. A total of 18 clients were	The response rate for clients was lower than for TA participants but still reasonable (38%). Across Phases 3 and 4, we interviewed clients from seven out of the eight TA aggregators going forward to delivery. While we did not interview any clients from the eighth aggregator, we did interview a non-participating aggregator that submitted capacity via this aggregator, which gave us insights into some of their capacity.

Evidence source	Phase 4 research tasks	Limitations
	interviewed during Phase 3 and 4, covering all of the DSR asset types identified during Phase 3 analysis.	There were some gaps in the cost data provided by aggregator clients via the email survey and the cost data was commercially sensitive.
Analysis of data from CM register and CM auctions	Updated analysis of performance, revenue and costs (including DSR test analysis to provide capacity breakdown by DSR asset & sector; plus analysis of DSR in TA vs other CM auctions). Our confidence in our characterisation of turn-down DSR in the second TA was significantly higher than during Phase 3. We based our analysis of capacity on DSR test results, which provided a one-off snapshot of capacity for all unproven CMUs. We based analysis of capacity for proven CMUs on the proven capacity of these CMUs.	While we had complete DSR test data for the second TA, our characterisation of capacity by business activity or asset type could only be indicative: Capacity Agreements specify capacity obligations at CMU level, so there can be considerable variation in the contribution of different sites within a CMU between different tests and turn-down events. While DSR test results provided a reasonable indication of the distribution of DSR capacity across different asset types, they probably overstated the total capacity that would be available for a stress event, because clients and aggregators could choose the timing of DSR tests to maximise the chance of passing DSR tests.
Interviews with National Grid, EMRS, ESC, Ofgem	Stakeholder interviews with National Grid, EMRS and ESC (covering TA delivery issues, the influence of TA versus external factors, and the 'value for money' provided by the TA)	Ofgem were unavailable for interview during Phase 4, although they were consulted during the first three phases of the evaluation.
Case study research	Six case studies were undertaken for specific sites in the second TA. These involved further collection and analysis of	To avoid disclosure, we presented excerpts from the case studies rather than presenting the case

Evidence source	Phase 4 research tasks	Limitations
	DSR operating information and cost data. The case studies were selected based on the following criteria: • DSR type – coverage of the main technology groupings identified during Phase 3 (industrial refrigeration and chillers, building Heating, Ventilation and Air Conditioning systems (HVAC), other motors and drives, process heating, and a bespoke industrial process). • Availability of data and agreement to respond to follow up questions.	studies in full. While some of the DSR technologies were relatively generic (e.g. HVAC, industrial refrigeration and chillers), others were bespoke to specific industries and sites (e.g. other motors and drives, process heating and other industrial processes). We used the case studies to explore costs further, including the opportunity costs of turn-down. The cost data gathered was commercially sensitive, so it was used to inform our overall costs estimates but was not directly linked to the case studies.
Analysis of stress event(s)		No stress event or CM Notice occurred during the 2017/18 delivery year to date. This limited the extent to which we could comment on the reliability of capacity provided by the second TA. Instead, we analysed the capacity provided in DSR tests and gathered views on the likely reliability of different types of capacity as part of aggregator client and aggregator interviews.
Validation of findings	Internal workshop with BEIS and peer reviewers.	We undertook an external workshop with delivery bodies and industry representatives about the second TA during Phase 3 but did not think it cost-effective to repeat this during Phase 4.

The email survey gathered information on the assets providing turn-down DSR in each CMU, and on the upfront capital and staff costs of participating in the second TA. Further details about the qualitative research and email survey methodology, including example topic guides, are provided in Appendix 4.

Qualitative analysis

The Phase 4 interviews were recorded and transcribed. We used slightly different analysis approaches for different topics, depending on whether we had formally tested particular elements of theory during the interviews:

- We analysed the evidence on additionality theory against the theory itself – see explanation below - as we had explicitly tested interviewees' agreement with possible additionality hypotheses.

- We undertook more generalised in-depth qualitative analysis of other topics (including reasons for drop-out, DSR testing experiences, reliability and cost of turn-down) because these spanned several areas of theory that were not explicitly tested with interviewees.

Formal testing of additionality theory

For the formal testing of additionality theory, we used spreadsheets to organise and code the Phase 4 and Phase 3 interview responses on additionality against contexts, mechanisms and outcomes (Cs-Ms-Os) in the theoretical framework, and to capture additional or non-additional contexts, mechanisms and outcomes that were supported by the interview evidence but not yet captured by the theory. This coding made reference to evidence from other sources (including interviews in previous phases, email survey data, case study findings, public statements by TA participants, auction data, and CM registers for the TA and other CM auctions) where this was relevant to assessing additionality. We systematically analysed the extent of support for different CMOs in the framework and for potential refined or new CMOs (see chapter 8 for an explanation of CMOs). Additionality findings for aggregators that continued through to delivery were also subject to formal evidence testing using a detailed process tracing approach, as explained in Appendix 4.

Having analysed the degree of support for different elements of the additionality and non-additionality theory in the theoretical framework, we then used realist contribution analysis to refine and revise the additionality CMOs in the theoretical framework. The aim of realist evaluations is to revise the initial theory to reflect the evidence, creating more 'tailored' theory, and then to generalise from this to create 'mid-range theory' which still captures the essence of findings but makes it potentially relevant to other programmes (in this case, the main CM). This theory revision process is illustrated in Figure 1.2 below.

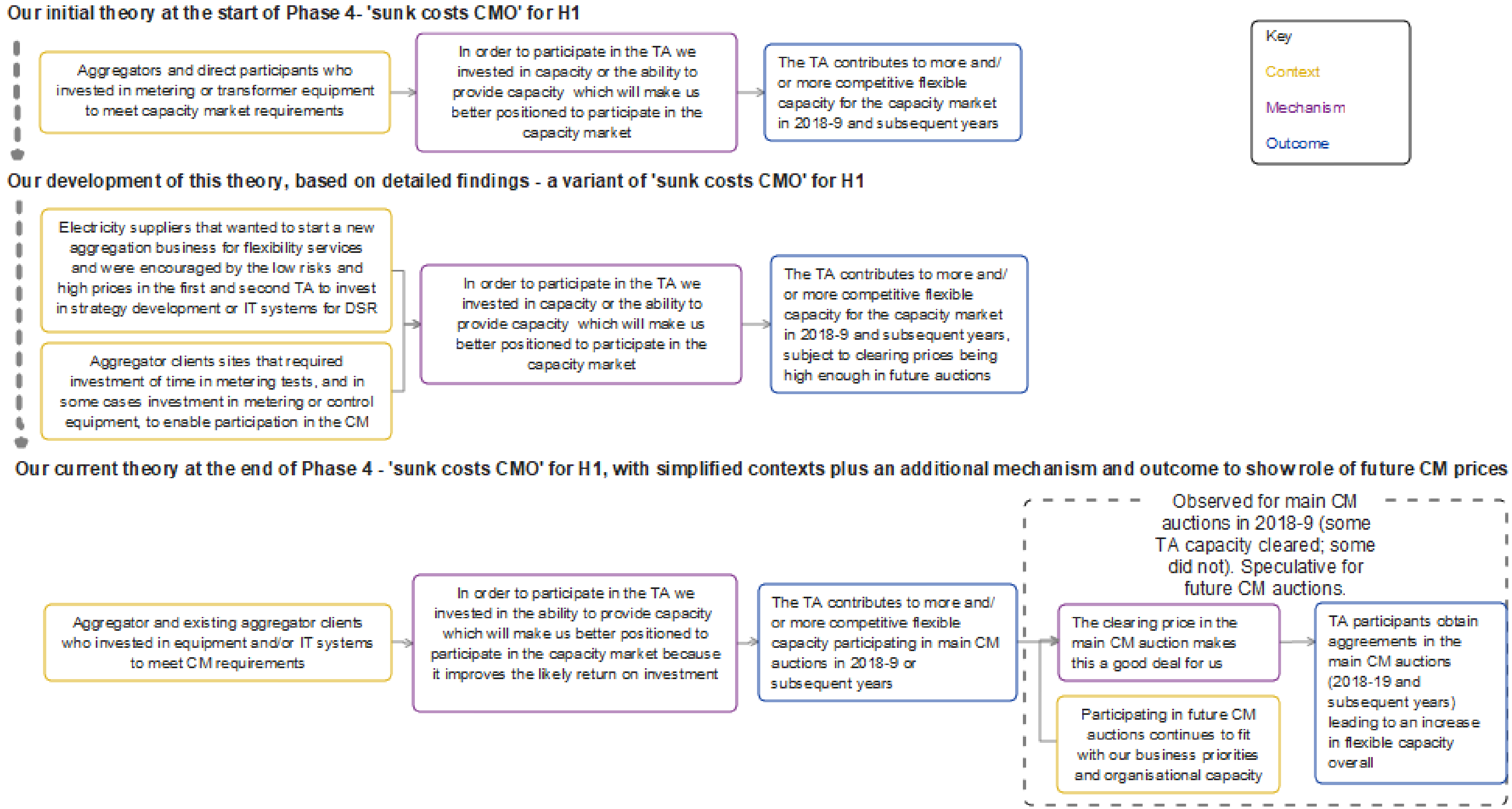

Figure 1.2: Example showing revision of one CMO as a result of findings

Qualitative analysis of other topics

For other topics explored in Phase 4 interviews (i.e. hypotheses or 'theories' about reasons for drop-out, testing outcomes, CMU design for reliability theory and potential Capaicty Market Notice responses), a coding frame was developed prior to analysis, designed around the key topics examined in the interviews. We then used spreadsheets to organise and code findings without explicitly organising these into CMO combinations at this stage. We applied a realist approach to this analysis, analysing what outcomes occurred for whom, in what circumstances and why, but did not attempt to formulate revised CMOs because these topics spanned several elements of theory that were lower priority for BEIS.

As explained above, both these types of qualitative analysis drew on data from other sources where relevant, in addition to the Phase 4 interview responses, including:

- Analysis of TA scheme data (e.g. auction data, CM Register).
- Findings from the email survey (where available).
- Data from previous phases of the evaluation.
- Published information (e.g. website statements, conference presentations).

The coding and analysis were undertaken by two researchers and findings were cross-checked by the wider project team. Members of the wider project team commented on draft findings from the qualitative research and took part in an internal workshop to discuss results emerging from the qualitative research and other workstreams.

Quantitative analysis of performance, costs and revenues

We analysed the overall performance of the second TA auction compared to other CM auctions by comparing and analysing the CM Registers for these auctions. We also characterised and analysed the capacity provided through the second TA by matching meter point data provided by National Grid with commercially available company databases for all CMUs that had passed DSR and metering tests. This enabled us to identify the company and sector of second TA sites with a high level of confidence[14]. We analysed the capacity provided by each site in unproven CMUs using DSR test data for the second TA provided by National Grid. The capacity provided by proven CMUs was based on their proven capacity in DSR tests for the first TA. DSR tests offer a snapshot of how component capacity contributes to CMU capacity, but are likely to overstate the capacity that would be available during a stress event (for reasons explained in Table 1.1

[14] There were a few sites where the address matching was tentative but telephone contact for the aggregator client screening survey subsequently confirmed the identity of the company.

above). We undertook high-level analysis of the initial capital cost and staff time involved in putting forward capacity to the second TA, using data from the email surveys. Analysis of costs for different types of turn-down DSR was not feasible because of the limitations of cost data collected. Finally, we collected some data on typical revenues from other flexibility services that were ‘stacked’ with the second TA. The methodology used for the analysis of costs and revenues is described in more detail in Appendix 6.

2. Findings on HLQ1: what were the outcomes of the second TA scheme?

SUMMARY

293 MW of turn-down DSR went forward to delivery in the second TA scheme, provided by 28 CMUs comprising 333 separate sites or 'components'. Six CMUs were temporarily suspended in May 2018 for failing to demonstrate three SPDs, but only four CMUs (8 MW) were still suspended in September 2018. Capacity in the second TA scheme was dominated by unproven DSR put forward by aggregators, provided by industrial rather than commercial sites. While water and food industries provided a large number of small sites, sites in metal-related, construction-material and manufacturing industries were larger and provided the majority of total capacity in DSR tests.

Costs associated with the second TA scheme appeared to vary widely, depending on the underlying process and the extent to which aggregators needed to recruit new clients. High-level analysis suggest that the average initial capital cost incurred by aggregators and direct participants for the second TA was £150/MW (range £0-580/MW). The average initial staff time input by aggregators and direct participants ranged from minimal to 52 days/MW, averaging 13 days/MW.[15] This is equivalent to an average of £4,800/MW at standard labour rates (range £0-19,300/MW). Average direct participant and aggregator staff time required for ongoing participation in the CM was predicted to be 7 days/MW/year (range 0-21 days/MW/year), equivalent to £2,600/MW/year (range £0-7,800/MW/year). Important caveats about these cost estimates are set out in this chapter and in Appendix 6.

The opportunity costs of turning down in response to a stress event were difficult for participants to assess, because of uncertainty about when a stress event would happen and how long it would last. Participants judged the likelihood of a stress event to be low and no stress event was observed during the delivery year. Case studies suggested that opportunity costs for sites in the second TA were generally low for short turn-down periods (e.g. up to two hours) but that business disruption costs could potentially rise significantly for longer turn-down periods (e.g. four hours or more).

Many of the second TA sites were already participating in cost-avoidance for Triad. The more experienced sites were also active in other National Grid flexibility services.

[15] All participants reported some staff inputs, but the minimum figures rounded down to zero when normalised by proven, contracted capacity.

Volumes of DSR capacity in recent TA and main CM auctions

Figure 2.1 shows that 293 MW passed testing and SPD requirements, out of the 756 MW originally entering pre-qualification for the second TA. This capacity was provided by 28 CMUs put forward by six aggregators and three direct participants, comprising 333 separate sites or ‘components’. The only significant change to TA capacity from the Phase 3 evaluation report was that six CMUs were suspended for failing to demonstrate three SPDs[16]’ before the end of April 2018. Evidence from TA participants, confirmed by National Grid, indicate that SPD problems arose from difficulties with establishing and maintaining the flow of meter data to the settlement body (EMRS), particularly for smaller sites with bespoke metering (see chapter 5 for further examination of these issues). National Grid reported that those TA participants that had suspended capacity were working to resolve these issues. By September 2019, two CMU suspensions had been lifted, and four CMUs (8 MW) were still suspended.

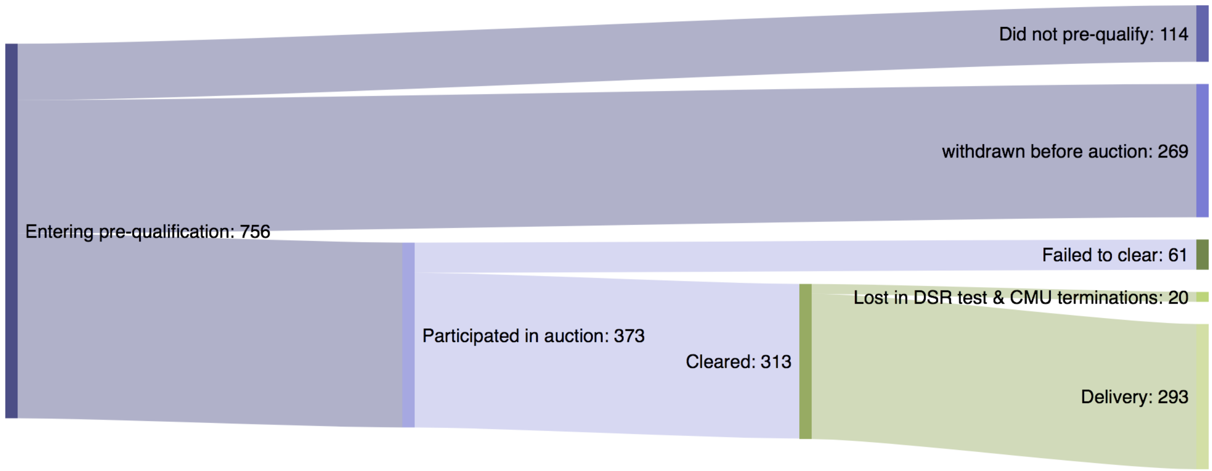

Figure 2.1: MW of capacity participating in the stages of the second TA (source: CM register)

Increasing volumes of DSR (back-up generation as well as load turn-down) have participated in recent CM auctions since the second TA. Just over 313 MW of turn-down DSR cleared in the second TA auction, compared to 475 MW in the first TA auction (which was open to both back-up and turn-down). Significantly higher volumes of DSR cleared in the T-1 auction for delivery in 2018/19 (443 MW) than in the Early Auction for delivery in 2017/18 (209 MW). More strikingly, the volumes clearing in the last two T-4 auctions have been 1.4 GW (for delivery in 2020/21) and 1.2 GW (for delivery in 2021/22). There were restrictions on TA capacity participating in the first and third T-4 auctions, so the increase

[16] CM participants are obliged to provide evidence of three half-hour settlement periods during the winter of a delivery year, on different days, in which they met their full capacity obligation. The CM register shows that were similar SPD problems in the Early Auction (EA) for the 2017/18 delivery year: 28 CMUs were from the EA in May 2018, of which 9 were DSR CMUs.

in DSR capacity in the third T-4 auction cannot be directly attributed to the TA. The contribution of the TA to growth in DSR is discussed further in chapter 3.

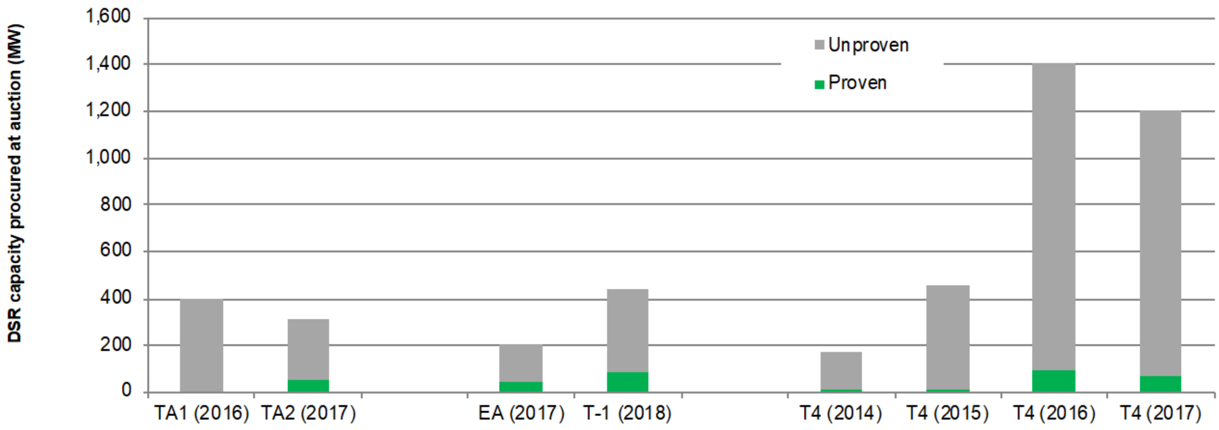

Figure 2.2: Volume of DSR clearing in CM auctions (MW) (source: CM register)

Figure 2.2 also shows that little of the DSR submitted to the main CM auctions was proven DSR. Analysis of the CM register indicates that proven DSR CMUs tended to be single-site CMUs involving one direct participant or a single aggregator client, rather than 'portfolio' CMUs involving multiple aggregator clients. Interview evidence suggests that aggregators with 'portfolio' CMUs submitted them as unproven to allow changes in the composition of CMUs from year to year, as – under current CM rules - CMUs only remain 'proven' while their components remain unchanged. There are proposals from Ofgem to allow some changes to proven DSR CMUs without requiring full DSR testing of the whole CMU, which may encourage more proven DSR CMUs to come forward in future. It is not possible to identify the scale of turn-down DSR procured in the main CM, as no distinction is made between different types of DSR.

Analysis of CM registers indicated that little DSR capacity was put forward by direct participants, with aggregators dominating the capacity that cleared in the TA auctions and in the main CM auctions. This was partly because few industrial and commercial sites had sufficiently large loads to participate directly: the minimum CMU size was 500 kW for the second TA and 2 MW for the other CM auctions. Interview evidence indicated that industrial and commercial organisations also chose to participate via aggregators because of the perceived complexity of participating directly in the CM, and because they saw aggregation as reducing the risks of CM participation. For example, aggregation could reduce the risk of non-compliance with a turn-down request, e.g. where an aggregator put forward capacity from several clients in the same CMU. Aggregators were also perceived as carrying contractual risk, offering specialist knowledge and offering a range of flexibility services to their clients.

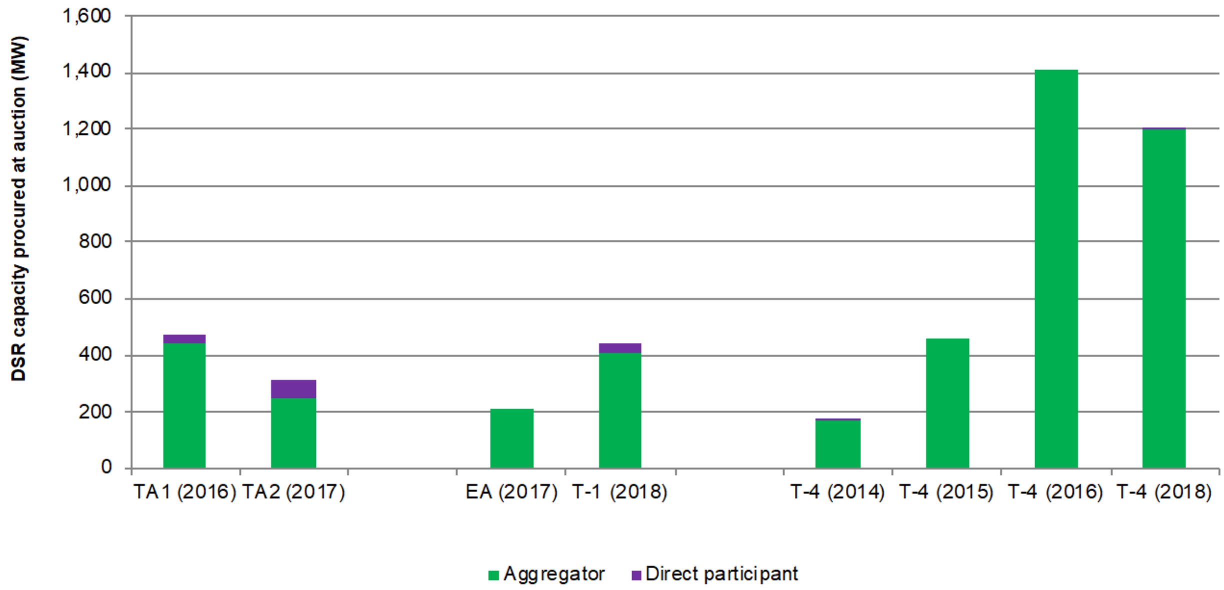

Figure 2.3: Proportion of aggregators and direct participants in recent CM auctions (MW) (source: CM register)

Characteristics of second TA capacity

28 CMUs went forward to delivery in the second TA (including the CMUs that were later temporarily suspended for failing SPDs). Using DSR test data for the 28 CMUs in the second TA, and MPAN data from National Grid, we identified and characterised 333 components with a total proven capacity 493 MW (compared to the proven, contracted capacity of 293 MW)[17]. The proven capacity of these CMUs was – in most cases – higher than their contracted capacity: the degree of 'overstuffing' and reasons for this were discussed in the Phase 3 report.

DSR test data is currently the only available measure of the capacity that individual components can provide[18]. Proven component capacities may (individually and in aggregate) legitimately vary between tests, so long as the contracted CMU capacity is met. While DSR test capacities generally exceed contracted capacities in absolute terms at CMU level, the relative volumes provided by different components in DSR tests provide some indication of the distribution of different types of capacity within CMUs.

[17] The total proven capacity shown in the CM register for the same CMUs is 483MW. We believe the 10MW (2%) difference is due the existence of some meter data including renewable generation which is not discernible within the raw data provided by National Grid.
[18] If there was a stress event, this would provide further data on component capacity, but there has not yet been a stress event within the 2017/18 delivery year.

Type of business activity

Our analysis of the business activities at all second TA sites is shown in Figure 2.4 below. This chart provides an estimated snapshot of the capacity provided by each CMU component, based on capacities estimated from DSR test data.

Figure 2.4 shows that DSR components were spread across a large number of business activities, mostly manufacturing, and that there were few components in the commercial sector. Interview evidence indicates that this was because commercial sites generally involved multiple small loads that were not cost-effective for the CM, particularly if extensive sub-metering was required to comply with CM metering requirements.

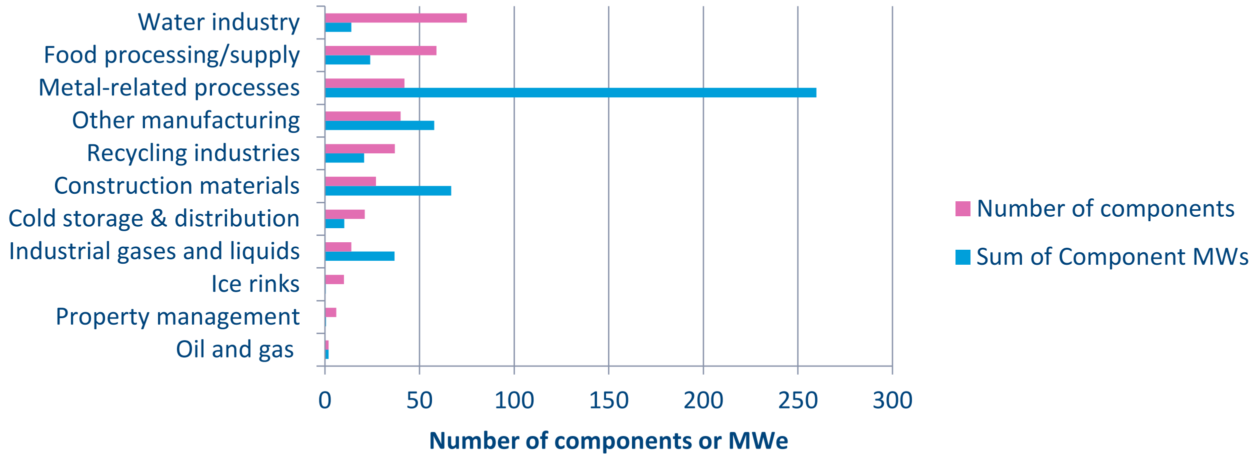

Figure 2.4: Number of components and DSR test capacity (MW) by business activity (source: consultant's analysis of DSR test and MPAN data for sites going forward to delivery)

Figure 2.4 also shows that water industry and food industry sites tended to be numerous but offer relatively low capacities. In contrast, sites with metal-related processes and those involved in the production of industrial gases/liquids and construction materials were less numerous but offered larger capacities per site in their DSR tests.

Type of assets being turned-down

We also analysed the types of electrical assets being turned down for the second TA, using the generic categories set out in Table 2.1. These categories were developed during Phase 3 research, in consultation with industry experts and academic peer reviewers. They have been refined to include lighting, which was identified in email survey responses as a further category of asset not included in the original list used in Phase 3.

Table 2.1: Generic types of DSR assets observed in the second TA (source: consultant's analysis of Phase 3 and 4 data)

DSR asset type	Further description
Building HVAC	Motors attached to air conditioning compressors in a commercial building.
Horticultural lighting	Lighting for horticulture, providing both heat and illumination.
Other motors and drives	Motors to drive conveyor belts, milling machines, crushers. The manufacturing activities associated with this category include: food, animal feed, quarrying and construction products, flour milling metals manufacture and metal recycling.
Process heating	Electrical heating used within an industrial process (e.g. steel manufacture (arc furnace), glass manufacture, insulation manufacture, aluminium smelter, plastic extrusion and baked products)
Pumps	Motors attached to hydraulic or gaseous pumps.
Refrigeration and chillers	Motors attached to industrial refrigerant pumps, compressors and fans. These are found in cold stores, food processing and ice rinks.
Other	Other process-related equipment such as machinery involved in the manufacture of paper, inorganic chemicals and windows.

We categorised the second TA components by asset type, primarily using information from the email surveys, supplemented by interview data, case study information and inference from business type[19]. Figure 2.5 shows the categorisation of all components in the second

[19] In a few cases, where we had no further information from email surveys or published data, we inferred the DSR type from the business activity. For example, in a few cases, we inferred that DSR provided by metal-related processes involved process heating or that other manufacturing involved motors and drives. We attached 'low' confidence to DSR categorisations made on this basis.

TA by asset type, using DSR test capacities for each component. Again, DSR test capacities provide a snapshot of capacity which may legitimately vary between tests. The legend (high, medium, low) refers to the level of confidence in our categorisation.

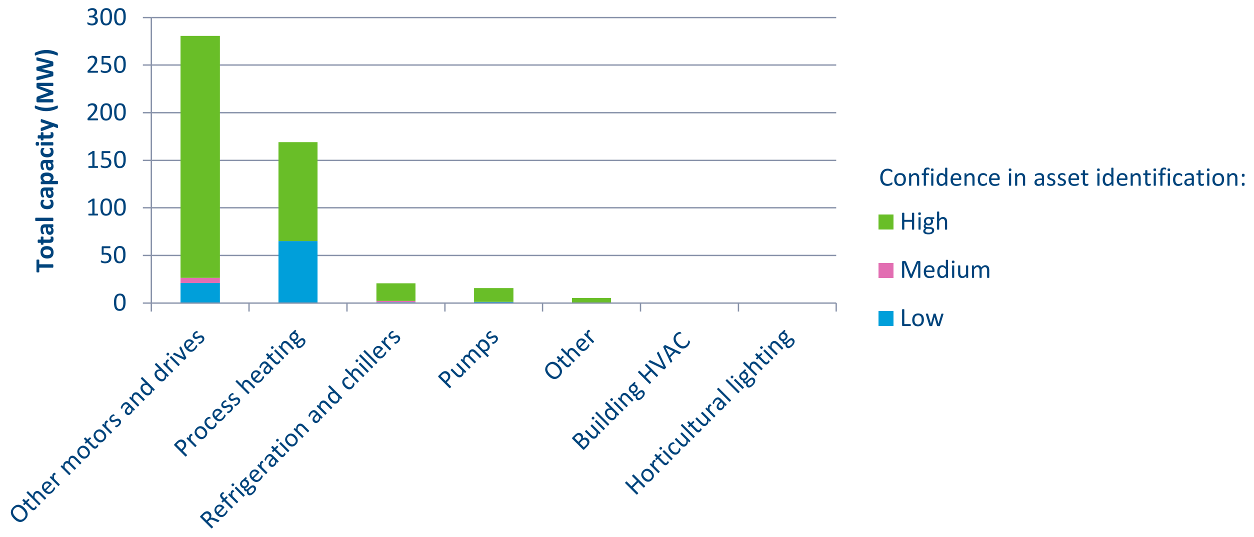

Figure 2.5: Sum of components by asset type, using DSR test capacity (MW) (source: consultant's analysis of Phase 3 and 4 data)

Figure 2.5 shows that the turn-down capacity in the second TA was primarily provided by turn-down of electrical motors and drives and turn-down of electrical heating processes. While the water industry had a large number of sites in the second TA, the volume of capacity provided by water pumps was relatively low compared to these other categories. Very little capacity was provided by Building HVAC and lighting.

This is consistent with the findings in Figure 2.6, which shows the mean and median capacity by component, based on DSR test capacities. Process heating, 'other motors and drives' and 'other bespoke industrial components', had the highest mean capacity per component, while Building HVAC and pumps had the lowest mean capacity per component. The difference between a high 'mean' and a relatively low 'median' for process heating indicated that capacity in this category (and 'other motors and drives') was dominated by a few very large sites.

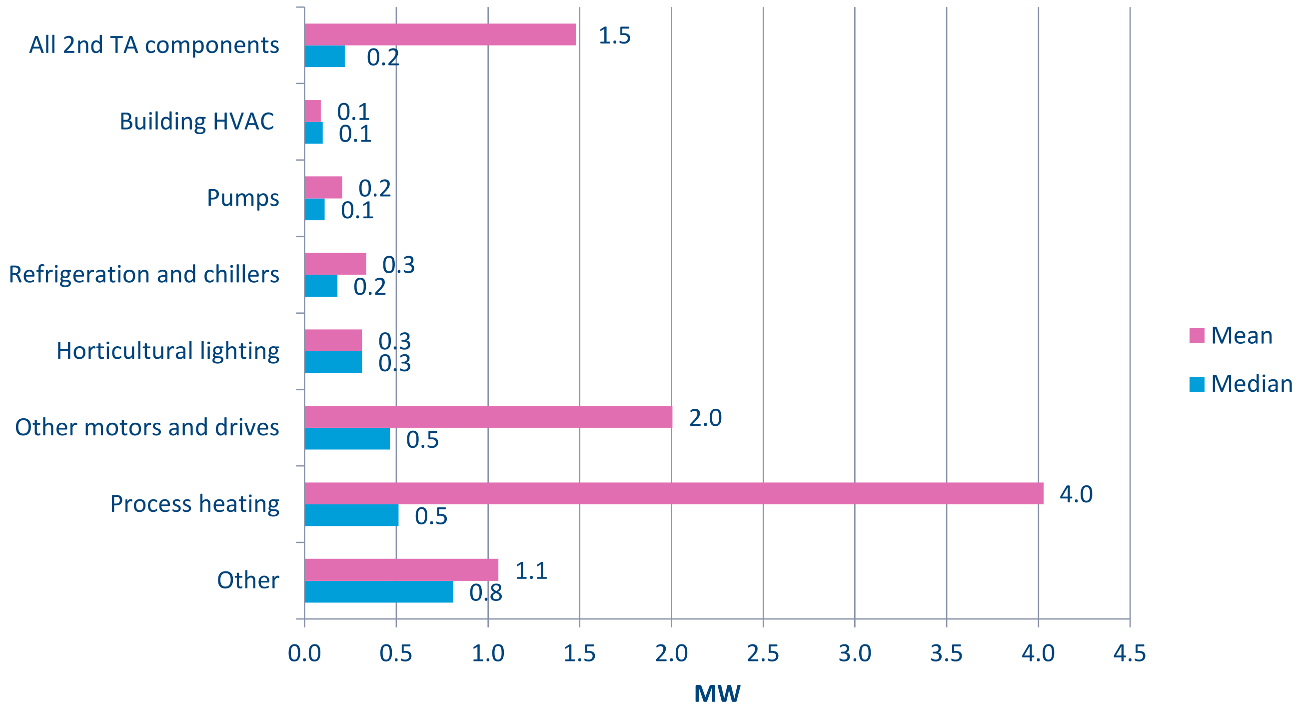

Figure 2.6: Mean and median component capacity by DSR type – from DSR test capacity (MW) (source: consultant's analysis of Phase 3 and 4 data)

The case study research gave us further insights into a few of the components in the second TA, across the main DSR asset types listed above. The capacity provided by these sites is characterised in Table 2.2, but we have insufficient information to assess how far these characteristics can be generalised to other sites in the second TA.

Table 2.2: Characterisation of turn-down provided in case study sites (source: Phase 4 research)

Asset type	Nature of turn-down DSR provided
Case study 1: process heating	• Turn-down is provided by an electricity-intensive heating process using a batch production system • Load shifting: it would be too expensive to turn-down mid-batch, as this would have severe operational consequences, but the plant can provide turn-down by delaying the next batch (provided adequate notice is received) • Production is already scheduled around Triad cost-avoidance, so baseline demand is variable • The site has just started to delivery frequency response services • Production is manually controlled • Metering was installed for the first TA, so the cost of second TA participation was minimal
Case study 2: mixed industrial processes	• Turn-down is provided across a number of sites using heating equipment, cooling equipment and compressors • Load shifting: certain assets can be turned down temporarily and production made up at other times • The sites also participate in Triad cost-avoidance and other flexibility services, so baseline demand is variable • Loads are controlled from a central control room • Turn-down volumes depend on multiple factors, including stock levels and customer demands • The notice period for the CM would allow stock levels to be increased before a turn-down event • Contracted capacity is set at conservative levels to reduce business risks and reduce the risk of non-delivery • No metering or controls were installed for the second TA
Case study 3: pumps	• Turn-down is provided by multiple water pumps at different sites

Asset type	Nature of turn-down DSR provided
	• Load shifting: sites might increase flows temporarily before or after a turn-down event so that water levels can be allowed to drop during the turn-down period • The sites also turn-down for Triad and DUoS cost-avoidance, so baseline demand is variable • Separate back-up generation capacity participates in other flexibility services • The pumps are controlled remotely from operations rooms • No additional control equipment or metering was required to participate in the TA, but investment would be required to bring in additional small sites which currently lack metering or controls
Case study 4: cold storage	• Turn-down is provided by cold storage for food products, comprising frozen and chilled storage • The frozen system accounts for around 80% of load • Combination of load shifting and load shedding: a temporary reduction in cooling system output will lead to storage temperatures rising, resulting in increased cooling demand later to return temperatures to the desired level • Weekends are used as ‘recovery periods’ because temperatures tend to rise during active hours of the week • The site is already providing Triad and DUoS red-zone cost-avoidance, so baseline demand is variable • Turn-down is currently provided using manual controls and no metering was installed for the TA • The company is also conducting frequency response trials for the frozen system, using automatic controls
Case study 5: mixed industrial processes	• Turn-down is provided by shutting down a manufacturing line, comprising a mixture of motors, drives and process heat • Load shifting: there is spare capacity in the manufacturing line so production can generally be halted temporarily without affecting the volume of packaged product leaving the site (providing that a sufficient stock of product is waiting to be packaged) • The site also turns-down for Triad cost-avoidance and operational purposes, so baseline demand is variable

Asset type	Nature of turn-down DSR provided
	• The site involves high-risk equipment involving steam so is not suitable for automatic turn-down • Sub-metering and a manual control system had already been installed so no investment was required to participate in the TA
Case study 6: HVAC	• Turn-down is provided by chillers providing chilled water to HVAC fan coil units within a large office building • Combination of load shifting and load shedding: temperatures within the building will increase as a result of turn-down, but not so much as to reduce comfort or endanger IT equipment reliability • The chillers are controlled by a Building Management System (BMS), supplemented by remote control turn-down for both the TA/CM and frequency response services, with the option of manual override by site engineers • Controls automatically switch on the chillers again if the water in the chiller circuit reaches an agreed maximum temperature • The contracted load is conservatively estimated to ensure they can still deliver in winter, when chilling loads are lower • The site participates in Triad cost-avoidance and other flexibility services using separate back-up generation • Sub-metering and controls had already been installed on the chillers, so no investment was required for TA participation

Cost data

We collected data on the cost of capacity made available to the second TA, using data from email surveys, interviews and case studies during Phase 3 and 4 of the evaluation. To avoid over-burdening respondents, most cost data was collected at participant and CMU level, with only a few data points being available at component-level.

Reasons for analysis approach

Analysis of costs for different types of turn-down DSR proved problematic for a number of reasons:

- Some CMUs had mixed technologies (i.e. they contained multiple components that used different types of turn-down DSR assets). So cost analysis by type of turn-down DSR required analysis of costs at component rather than CMU level.

- There was little component-level cost data and there were low numbers of components in some of the categories of DSR specified above.

- Most of the cost data was provided at TA participant or CMU level, so allocation of costs to component level involved several normalisation assumptions (effectively ‘smearing’ the cost data across individual components). The resulting costs per component were widely distributed and did not provide meaningful results for different technology types.

For these reasons, it was not feasible to develop meaningful costs estimates by type of turn-down DSR. An alternative approach was therefore used. This involved analysis of costs at a higher level, across all types of turn-down DSR. Because component-level cost data was very limited, the data for this high-level cost analysis was sourced from the email survey which provided consistent data across most of the TA participants.

Methodology for high-level cost analysis

High-level cost data was provided by seven out of the nine TA participants, comprising two direct participants and seven aggregator participants. The costs covered 19 CMUs and represented 243 MW of proven, contracted capacity (compared to the total of 28 CMUs and 293 MW of proven, contracted capacity that went forward to delivery in the second TA). The following data was collected from these participants through the email survey. Some of the surveys were completed during Phase 3 and others during Phase 4 of the research:

- Upfront metering and equipment costs (£) – data for 19 CMUs going forward to delivery, collected from TA participants at CMU-level.

- Initial staff time (hrs) – data for seven TA participants[20], provided at organisational level.
- Ongoing staff time – annual (hrs) – data for seven TA participants, as for initial staff time.

The data was normalised by MW to make the figures more useful for BEIS policy purposes. Data was shared with BEIS in normalised and anonymous form to avoid disclosure. Costs were divided by the proven, contracted capacity going forward to delivery in a given CMU, or by a given TA participant. Average costs were derived by dividing the total costs across all seven participants (or all 19 CMUs) by the total proven capacity in these CMUs (i.e. 243 MW). Proven, contracted capacity going forward to delivery was the capacity contracted post-auction for these 19 CMUs, adjusted for any reductions in proven capacity arising from DSR tests and metering tests.

Cost analysis

Cost analysis was undertaken at participant and CMU level. As explained more fully in Appendix 6, analysis of costs by DSR technology type did not provide meaningful results because of the complex assumptions required to allocate participant and CMU-level costs and capacities across multiple components in mixed-technology CMUs.

Data on initial capital costs associated with the second TA was provided at CMU level for 19 out of the 28 CMUs going forward to delivery (covering 243 out of 293 MW). Data on staff time inputs was provided at TA participant level for seven out of the nine organisations going forward to delivery (representing two direct participants and five aggregators). We normalised initial capital costs and staff inputs by the proven, contracted capacity put forward by each participant or CMU, to avoid disclosure and to provide estimates that could be used for BEIS policy purposes.

Limitations of the cost analysis

- The costs presented below do not include the opportunity cost to businesses of turning down demand in response to stress events or tests (see next sub-section).
- As the cost data was collected from TA participants by email (with some responding in autumn 2017 and others in spring 2018), it only includes costs incurred up to the time of the email survey. It does not include costs incurred during the delivery year such as the cost of meeting Satisfactory Performance Day requirements. Interview evidence suggests that there are some costs involved in the coordination of SPDs across multi-component CMUs.

[20] There were eight responses to the email survey, but one of these was for an aggregator that provided capacity for several CMUs to a TA participant on a sub-contracting basis. The cost incurred by the sub-contracting aggregator have been included within the total costs for that TA participant.

- Similarly, the costs do not include any costs associated with metering requirements that arose during the delivery year (e.g. new sub-metering requirements because of renewables being installed on a particular site).

- Not all TA participants responded to the email survey – the costs below are based on staff time data for seven TA participating organisations and capital costs for 19 CMUs. A few additional data points provided by TA participants via interview or case studies are not included because these are not available on a consistent basis.

- The costs and time estimates do not include any capital costs or time incurred by aggregator clients, because of the lack of consistent data across different CMUs. Evidence from our small sample of aggregator clients suggests that the capital costs of metering and control equipment were normally borne by the aggregator and that time inputs by clients were significantly lower than for aggregators. A separate analysis of client costs for a small sample of clients is presented in Appendix 6.

- The data includes a mix of direct participants and aggregators. These are not presented separately to avoid disclosure, because of the small number in each category.

- The email survey asked participants to provide estimated costs specific to the second TA. This means that costs already incurred for participation in the first TA or other flexibility services are not included here.

- The staff time inputs include abortive development time as well as clients/sites that actually went forward to delivery. For example, they include time recruiting clients for CMUs that were not contracted in the second TA auction or that were terminated post auction because of failing to pass testing requirements.

- As the data were provided in 2017/18, they do not capture later developments in the DSR market, such as the progressive trend towards DSR aggregation being offered as part of integrated energy solutions rather than as a standalone service by specialist DSR aggregators.

More detail on the cost methodology can be found in Appendix 6.

Initial capital expenditure by aggregators and direct participants

As shown in Table 2.3 below, initial capital expenditure on metering equipment to enable participation in the second TA varied between £0 and £571 per MW of proven, contracted capacity. The average initial capital expenditure was £144 per MW. This expenditure related solely to metering equipment and its installation. Interview evidence indicated that aggregators chose (wherever possible) to select sites that had simple metering requirements and did not require any up-front capital investment to participate in the second TA. Metering costs were only incurred for relatively large sites that could offer sufficient capacity to justify capital investment for the second TA (with a view to participating in the future CM).

Table 2.3: Initial capital expenditure by aggregators and direct participants for second TA (£/MW)

Item	Average cost £/MW (mean)	Min	Max
a. Capital expenditure on controls	0	0	0
b. Installation of controls	0	0	0
c. Capital expenditure on metering equipment	70	0	290
d. Installation of metering equipment	80	0	290
e. Other capital expenditure	0	0	0
Total initial capital expenditure	**150**	**0**	**580**

Source: email survey data from 19 CMUs, normalised by proven, contracted capacity

Caveat: The indicative costs quoted were calculated by the evaluation contractors using self-reported costs from a small sample of participants in the second Transitional Arrangements auction held in 2017. They may not be representative of all the TA DSR participants or of all the costs faced by DSR market participants more generally. They are likely to be under-estimated.

Initial staff time inputs by aggregators and direct participants

The staff time required to put forward capacity to the second TA varied between zero and 52 days per MW of proven, contracted capacity. As shown in Table 2.4, the average staff time inputs were 13 days per MW up to the beginning of the delivery year, equivalent to £4,800 per MW at standard labour rates. The range of costs was £0-19,300 per MW. Staff time requirements for the delivery year itself were not known at the time of the survey.

Most of the time inputs were associated with aggregators signing up clients and organising testing of client capacity. Some aggregators spent less time on client recruitment because they brought in their existing electricity supply clients or flexibility customers, but significant time inputs were still incurred to meet TA/CM testing requirements. Time inputs by direct participants were much lower, partly because they had fewer sites and partly because they did not need to meet testing requirements, having proven their capacity in the first TA. Although all the TA participants required some time to participate in the second TA, the minimum figures were rounded down to zero when normalised by capacity.

Table 2.4: Initial staff time inputs by aggregators and direct participants for second TA (full time equivalent (fte) days per MW)

Item	Average days/MW (mean)	Min	Max
a. staff time for marketing to clients/sites	1	0	10
b. staff time for signing up clients	7	0	16
c. staff time associated with testing	4	0	23
d. staff time for pre-qualification & participation in auction	1	0	11
e. Other staff time for 2nd TA	0	0	0
Total initial staff time (fte days per MW)	**13**	**0**	**52**
Estimated initial staff cost (£ per MW)[21]	**£4,800**	**£0**	**£19,300**

Source: email survey data from seven TA participants, normalised by proven, contracted capacity.

Caveat: The indicative costs quoted were calculated by the evaluation contractors using self-reported costs from a small sample of participants in the second Transitional Arrangements auction held in 2017. They may not be representative of all the TA DSR participants or of all the costs faced by DSR market participants more generally. They are likely to be under-estimated.

Ongoing staff time inputs by aggregators and direct participants

The predicted time inputs required to participate in the CM on an ongoing basis were lower, because they excluded client recruitment time. As shown in Table 2.5, the average staff inputs by aggregators and direct participants were predicted to be 7 days per MW per

[21] Staff time has been converted into costs using the latest available (2012) Green Book labour rate of £336/day, assuming 5 years of wage inflation at 2%, giving a rate of £371/day. Rounded to nearest £100.

year, varying between zero and 21 days per MW per year. The estimated cost of these inputs, at standard labour rates, averaged £2,600 per MW per year, with a range of £0-7,800 per MW per year. Again, all the TA participants predicted some time requirements, but the minimum figures were rounded down to zero when normalised by capacity. These predicted costs were dominated by ongoing client/site engagement and by the annual auction process. Some testing requirements were still envisaged because of potential changes to the composition of multi-site CMUs, which could necessitate retesting of some sites or CMUs.

Table 2.5: Ongoing staff time inputs for annual CM participation (fte days per MW per year)

Item	Average days/MW (mean)	Min	Max
a. client/internal site engagement	3	0	8
b. adjustments to CMU composition (if required)	0	0	1
c. further testing/other compliance costs	1	0	5
d. annual auction process (e.g. future T-4 or T-1)	3	0	13
e. any other ongoing time inputs (e.g. data flow issues)	0	0	0
Total ongoing staff time (fte days per MW)	**7**	**0**	**21**
Estimated ongoing staff cost per year (£ per MW per year) [22]	**£2,600**	**£0**	**£7,800**

Source: email survey data from seven TA participants, normalised by proven, contracted capacity.

Caveat: The indicative costs quoted were calculated by the evaluation contractors using self-reported costs from a small sample of participants in the second Transitional Arrangements auction held in 2017. They may not be representative of all the TA DSR participants or of all the costs faced by DSR market participants more generally. They are likely to be under estimated.

Opportunity costs of turning down

During Phases 3 and 4 of the evaluation, we researched the opportunity cost to businesses of turning down their loads in response to a particular event (e.g. a CM Notice or test). Where possible, the research team tried to obtain quantitative data on opportunity

[22] Staff time has been converted into costs using the latest available (2012) Green Book labour rate of £336/day, assuming 5 years of wage inflation at 2%, giving a rate of £371/day. Rounded to nearest £100.

costs. This was problematic because of the wide range of business activities present in the second TA and the wide range of operational contexts in which turn-down might be requested. Aggregators usually did not know much detail about the opportunity costs of turn-down provided by their clients, and some aggregator clients and direct participants were uncertain as to what the cost impact would be (depending on the timing and length of turn-down in relation to their operations and other factors) or were unwilling to share details of their internal risk assessments and cost estimates.

We were, however, able to develop some understanding of the types of costs that turn-down might generate for a business, and how these costs would typically vary according to the length of a turn-down event. Our findings are presented in Table 2.6 below. Fuller exploration of the factors affecting businesses' ability to turn down for a specific event are presented in chapter 5.

Table 2.6: Generic stages of opportunity cost by length of turn-down event (source: consultant's analysis of Phase 3 and 4 data)

Stage	Typical duration	DSR service offered	Opportunity cost	Typical cost factors
1. Standard response	< 2 hours	Yes	Negligible - low	Management time
2. Extended / full response	2 to 4 hours	Yes, but most hope it won't happen	Negligible to £13/MWhr* (based on 5 CMUs)	As above + energy and staff overheads, service disruption or minor production loss
3. Long duration	4+ hours	Some can, but for many this would be problematic	High £10+k	As above + temporary production loss/ service interruption
4. Very long duration	8+ hours	No	Significant business impact	As above + business service disruption

The duration of a turn-down event at which a DSR provider will move to the next stage (as shown in Table 2.6) is typically case-specific and depends on multiple variables, such as the weather, time of day, the season, levels of stock and business activity/production orders. These variables affect not only the ability to turn down on the day of the turn-down event but also the site's baseline demand[23]. The notice period for turn-down can also have a significant effect: with adequate warning, several of the case study providers could make provisions to coincide turn-down events with scheduled maintenance or to increase their stock levels to negate or at least mitigate the impact of lost production.

Reliability and opportunity costs are closely related, as DSR providers will cease to provide a turn-down response once cost and risk exceed a certain threshold. Making this judgement is usually the role of a plant operator, although automated thresholds[24] may be used to trigger the end of stages 1 or 2, for example through temperature limits for thermals assets (cold stores, building HVAC, and process heating) and water levels for pumping.

One case study participant was only able to participate in the second TA through an aggregator. Their opportunity costs were too unpredictable to provide capacity with certainty due to uncontrollable issues such as weather and variable baseline demand. For the revenues available through the TA, they could only participate when delivery risks were mitigated through aggregation.

Revenues

Evidence from email surveys, interviews and case studies indicated that many of the second TA sites were already participating in cost-avoidance for Triad. Some were also undertaking cost-avoidance for DUoS red zone costs, while some were putting their second TA capacity into frequency services or STOR. The range of services depended on the operational constraints of their business, their level of experience with flexibility and on the services provided by their aggregator. Interview data indicated that some aggregator clients had recently changed aggregators to access new revenue opportunities for their flexibility, such as the TA/CM, frequency services or wholesale market revenues.

The revenue data collected directly from second TA participants was too patchy and inconsistent to provide reliable estimates of actual revenue from other sources. But the table below compares the level of revenue typically available for different flexibility services.

Table 2.7 shows that Triad cost-avoidance generally offered a slightly greater £/kW value than the CM. Interview evidence indicated that participants found Triad much more accessible (e.g. no contracts or metering requirements) but that Triad required more frequent responses (circa 20-30 each winter season) to ensure that the maximum cost-

[23] Turn-down delivery is measured against baseline demand at specific times over the six weeks prior to a turn-down event. The baseline methodology is explained further in Appendix 1.
[24] Interview evidence indicated that, wherer DSR assets were suitable for automatic control by the aggregator, and where there were important operational limits for the aggregator client's business, controls could be used to offer DSR within certain thresholds. For example controls could be used to turn an HVAC system back on again automatically when the building temperature rose above a certain level.

avoidance was achieved. However, Ofgem's review of embedded benefits may significantly reduce Triad benefits in future, as discussed further in chapter 6.

Table 2.7: £/kW value of common DSR revenue streams

	£/kW	£/kW	Comment and source
	Low	High	
Capacity Market	6	45	Low case is for the 2017/18 T-1 auction and high is 2nd TA
Triad cost-avoidance	26	54	2017 rates as published by National Grid; varies by region.
Dynamic FFR	40	80	Indicative values only; based on National Grid availability and utilisation payments

Dynamic Firm Frequency Response (FFR) was the most lucrative service on a £/kW basis. It was, however, reported to be a demanding service to deliver, requiring the installation of specialist equipment and acceptance of a degree of automated / third party asset control.

3. Findings on HLQ2: how and why the second TA contributed to its objectives

SUMMARY

We examined how the second TA contributed to its objectives of (a) contributing to flexible capacity in the future CM and (b) encouraging turn-down DSR more widely. The 293 MW of turn-down DSR procured in the second TA was significantly higher than the estimated 60-90 MW of turn-down procured in the first TA, which was open to both back-up and turn-down DSR. All the participants in the second TA went on to participate in the main CM auctions in 2018, although not all their capacity cleared at the lower prices in these auctions. We found evidence of positive outcomes attributable to the second TA scheme, rather than other factors, for:

- **Aggregators and their clients that were less experienced with the CM (as the soft conditions in the second TA provided a safe learning ground).**

- **Aggregators that were less experienced with turn-down DSR (as the second TA incentivised them to learn about and recruit turn-down DSR).**

- **Aggregators and clients that invested in controls, metering equipment or systems for the second TA (reducing the costs of future CM participation).**

- **New and existing aggregators that recruited some clients new to flexibility services, attracted by the high price and low credit cover in the second TA scheme (as this expanded their customer base of turn-down DSR for the future CM and other flexibility services).**

We did not find evidence of outcomes attributable to the second TA for:

- **Direct participants and those aggregator clients that were highly experienced providers of flexibility services or were very confident in energy management (as their capacity was already available to National Grid via other services).**

- **Non-participant aggregators that had mixed back-up and turn-down DSR capacity on small sites that would require investment in metering to participate in the second TA (as this investment was not required for the main CM and – for small sites - was not justifiable for the second TA alone).**

The extent to which turn-down capacity from the second TA will secure agreements in future CM auctions will depend on future clearing prices, on participant’s ability to stack CM revenues and on the details of CM rules.

This chapter presents our findings about the second TA's contribution to its objectives and how far these were 'additional' compared to what would have happened in the absence of the second TA.

We agreed with BEIS that the evaluation should focus on the availability of electrical capacity, rather than on how this capacity is used during the delivery year. This is consistent with the concept of the CM for electricity, both in the UK and other countries. The factors affecting the reliability of this capacity for delivery during a stress event are discussed in chapter 5 below.

We tested two additionality hypotheses, based on the two objectives of the second TA:

- H1 – second TA contributes to more (competitive) flexible capacity for the CM in 2018/19 and subsequent years.
- H2 – the second TA contributes to wider encouragement of turn-down DSR.

These theories overlap but are distinct: H1 focuses on the second TA's contribution to flexible capacity in the main CM (including both back-up generation and turn-down DSR). Conversely, H2 focuses on encouragement of turn-down DSR in any services contracted by the National Grid, not just the CM.

Approach to assessing additionality

At the start of Phase 4 of the evaluation, we developed our 'candidate' theory for the additionality of the second TA, as shown in Appendix 2. We tested this theory against evidence about the second TA, collected during Phases 3 and 4 of the evaluation. The evidence includes publicly available evidence about organisations' participation in the two TA auctions and other CM auctions as well as interview evidence and data gathered through email survey responses. We also used process tracing to test the strength of the evidence we had gathered about the additionality of aggregator behaviour, and how much was attributable to the TA compared to other influences. This is explained in Appendix 5.

Our assessment of additionality is based on evidence from aggregators, direct participants and aggregator clients that were involved in delivery for the second TA. The aggregator sample includes one non-participating aggregator that submitted capacity via a participating aggregator.

Competing explanations

There were a number of driving forces for increased flexibility and increased turn-down at the time of the second TA. These included the Medium Combustion Plant Directive (MCPD) which constrained CM participation by older diesel plant; the Power Responsive Campaign, which raised awareness of DSR opportunities; the falling cost of battery technology and the high level of network charges (e.g. Triad and DUoS) for electricity consumers with half-hourly meters. While these influences were pushing in the same direction as the second TA, we attempted to assess whether the second TA itself made a difference to the volume and competitiveness of turn-down DSR that came forward to the CM and other flexibility services.

To some degree, the second TA was building on the experiences of the first, not so much in terms of proven capacity (because of the second TA being restricted to turn-down DSR only, which excluded much of the capacity in the first TA) but in terms of learning for TA participants.

Contribution to H1 - the second TA contributes to more (competitive) flexible capacity for the CM in 2018/19 and subsequent years

The 293 MW of turn-down DSR procured in the second TA was significantly higher than our estimate of 60-90 MW of turn-down procured in the first TA, which was open to both back-up and turn-down DSR. All of the participants in the second TA went on to participate in the main CM auctions in 2018, although not all of their capacity cleared at the lower prices in these auctions. There was evidence of increased turn-down capacity coming in to the future CM from aggregator clients that participated in the TA and that were willing to clear in the 2018 T-1 auction.

> *One of the customers ….started with 3 sites worth 3MW in total, at £45,000 [per] megawatt for this year. They're bringing 4 new sites [into the 2018 T-1], at £6. Even though the price is much lower, they are bringing more capacity into the market. [Is that turn-down as well?] Yes.(Aggregator, Phase 4)*

Interview evidence indicated that - with some reservations, as discussed below - participants in the second TA planned to continue participating in the CM. The volume of flexible capacity that they will actually clear in future CM auctions will depend on the clearing price in those auctions.

While we found evidence of the second TA's influence on DSR volumes in the CM, and evidence of 'CM entry costs' being covered by the second TA (see below), it was challenging to find direct evidence of TA participants bidding at a lower level in the main CM than they did in the second TA. As we did not have access to bid prices (i.e. exit prices) for the main CM auctions, we were reliant either on interview respondents reporting the prices at which they would (in theory) be willing to clear capacity in the main CM or those involved in the second TA clearing the same (proven) DSR capacity in another auction at a known clearing price (e.g. £6/kW for the 2018 T-1 auction).

Interview data suggested that TA participation might have had some influence on the 'reserve price' agreed between aggregators and their clients, where these clients had gained confidence that the CM involved low impact on their business:

> *I suspect the experience in the TAs has given them [clients] confidence that yes, their assumption is right, the number of activations is quite limited, the impact on their business is quite limited, so it allows them perhaps to refine down their assumptions as to what revenues they need. So that might lead to a slightly lower reserve price in that sense. (Aggregator, Phase 4)*

We examined evidence about the extent to which the second TA influenced this ongoing participation in the main CM, through higher volumes, new clients or sunk costs. The contexts and reasoning in which the second TA was 'additional' (i.e. did contribute towards a greater volume of capacity, or more competitive capacity[25]) are outlined below. Some second TA participants exhibited more than one of these reasonings so, for them, the second TA was additional in more than one sense.

Additional contribution to H1

Reasoning: your experience of participating in the TA means the CM seems less risky
Certain aggregators and aggregator clients reported that they gained learning from the second TA that made the main CM seem less risky. The low credit cover in the second TA helped to make it a 'safe place' to learn, as it was only 10% of credit cover in the main CM. This reasoning was observed in the following contexts.

- Aggregators that had not previously completed a TA cycle were able to learn about the CM process and assess client risks and costs more accurately.

> *I think it's just having done a cycle of it you know what the requirements are. As I say, we've adapted our contracts to reflect the risks and costs for the customer. We know better to try to keep up with the timetable and all the pitfalls of the process. It's those learnings. (Aggregator, Phase 4)*

- Aggregators developing a flexibility business were enabled to take risks by the low credit cover, and this accelerated their learning.

> *So I think going forward, we don't want to be losing any more credit cover given that it's £5,000 going forward instead of £500. I think that was the real benefit of the TA, was the fact that the credit cover was a tenth of what it actually is normally. So it was a good learning curve for us in terms of, we need to be more conservative in our analysis of their flexibility. (Aggregator, Phase 4)*

- Existing aggregator clients that were concerned about the risk of delivery were enabled to learn and develop their approach to DSR in the CM. For example, this

[25] We did not attempt to separate outcomes involving greater volumes of capacity from those involving more competitive capacity, as evidence indicated that these outcomes were closely linked. For example, a lower perception of risk could affect both participation volumes and bidding prices in a future CM auction.

applied to clients that were concerned about the variability of their baseline demand; or that had assets/operations which needed a significant notice period or manual control; or those that were concerned that multiple stress events would occur during the delivery year.

What we do, I think it is true, we do see that, over time, as the clients get familiar with the process, it becomes less scary to them. They start to think, "Okay, yes, I can do this for less than I thought I could." …We do see that customers that have been with us for a year, you know, their price does come down compared to what it was when we were first trying to recruit them. (Aggregator, Phase 4)

Reasoning: in order to participate in the second TA, we invested in the ability to provide capacity, which will make us better positioned to participate in the main CM

Firstly, certain aggregators and aggregator clients reported that they invested in controls, metering equipment or IT systems to meet CM requirements for the second TA. Aggregators anticipated that some of the costs associated with getting customers set up to participate in the second TA would not be repeated in the main CM.

The sunk costs, we've already made for the existing customers, are already in place. It's only upside from here on. It's going to be much easier, and we're probably going to spend a lot less time, to bring our current portfolio into next year. (Aggregator, Phase 4)

Secondly, some new aggregators invested in systems and approaches that they would apply to a range of DSR services, including the main CM.

The second TA was our first real go at [DSR aggregation]…So we've got a new platform which is called [name redacted]. (Aggregator, Phase 4)

Reasoning: we have been able to build our customer base for turn-down DSR because the TA offered attractive terms.

All aggregators participating in the second TA reported that they had been able to recruit some clients that were new to flexibility using the combined attraction of the high price and low credit cover. This had enabled new aggregators to establish a customer base for turn-down DSR in Great Britain (GB).

I think it helped us establish a strong foothold. You know, we've developed a reasonably sized turndown portfolio, because we had a good opportunity in the TA. (Aggregator, Phase 4)

While existing aggregators with turn-down clients had seen the second TA as an opportunity to obtain extra revenues for existing clients that could meet CM requirements, it also enabled them to recruit new clients and expand their turn-down DSR portfolios.

So it helps to have a higher price to recruit them as a new customer and get over some of their initial hesitation and concerns, especially if they're not that familiar with … DSR, never done it before. So having a good price, like we did in the second TA, is really great for recruiting a lot of new capacity. ...You know, there is a *premium just to get customers to get over that initial hesitation. (Aggregator, Phase 4)*

For existing aggregators that were not particularly experienced with turn-down DSR, the second TA generated learning about how to explain to clients about turn-down DSR in the CM.

Yes, certainly we've learned an awful lot about it [turn-down], and we're geared up for it now, so we can do it going forward. [Did you do it at all before the second TA?] I think only a little bit really, but not in the Capacity Market in any meaningful way…. And also explaining to clients and talking to them about what the requirements are and what's involved. (Aggregator, Phase 4)

Reasoning: 'the high price in the second TA caught our attention and motivated us to contract with an aggregator'

The aggregator clients that were new to flexibility and were attracted by the high price in the second TA were already turning down for Triad and/or DUoS red zone cost-avoidance, if they possibly could.[26] Triad cost-avoidance was the main driver for their adoption of turn-down DSR and they saw the TA (and future CM) as an easy and low-risk add-on to turn-down for Triad.

To be honest, going back, the TRIAD avoidance proved that we could shut down and reduce energy [..] So, when we saw this as an additional opportunity, we saw it as low risk because the projections are this could be once or twice a year, so it's quite a low risk for us to go into that. (Aggregator client, Phase 3)

However, irrespective of whether they increased their volume and frequency of turn-down beyond what they already did for Triad, there was additionality in that they contracted with an aggregator for an external flexibility service that could be requested by National Grid. Aggregators expected clients to stay on board with them for the main CM, irrespective of whether they had signed multi-year contracts.

All of our customers that made it are wanting to go in again. If they've had their metering statement and if nothing's changed on their metering, we can submit that again. Once they're in, it's a lot easier to keep them in unless they have a business change. (Aggregator, Phase 4)

No additional contribution to H1

There were TA participants for whom we observed contexts and reasoning that led to the second TA making no contribution to its objective of developing additional flexible capacity for the future CM.

[26] The only exceptions were clients that could not turn-down frequently enough to avoid Triads.

Reasoning: we have always intended to participate in the CM and the second TA did not make any difference to that

Direct participants and aggregator clients who were either highly experienced providers of flexibility services or very confident in energy management reported that they would have participated in the CM anyway, going straight into the T-1 or T-4 auctions. The high price in the second TA was mainly just a windfall to them, although it may have made initial conversations easier (both internally and with their aggregator).

> *We signed a relatively long-term arrangement anyway, we knew the [TA] would have finished, yes we took advantage of the initial period of it because the price was very high, we did some analysis, we got some prices, so I think the [TA] wasn't really a big impact; we knew we were going in for a four or five-year contract, so I wouldn't have said the [TA] reshaped that in particular. (Aggregator client, Phase 4)*

Less experienced aggregator clients were not always aware of the difference between the conditions of the CM and TA, because they left the details to their aggregator.

> *I'll be honest, I probably wasn't even aware initially there was a difference between transitional, and T1/T4's, [the second TA] didn't shape our thought process, definitively no.(Aggregator client, Phase 4)*

Some of the organisations who argued that the TA was not additional did not actually clear in the main CM in 2018. Whether a particular organisation cleared in the main CM auctions depended on the clearing price relative to their cost base, the other flexibility revenues available to them and the risks they perceived in relation to CM participation.

Reasoning: the main CM is more attractive to us than the second TA, either because it suits our capacity better or because it offers a steady stream of revenue over several years

The second TA was not attractive to aggregators, aggregator clients and potential direct participants that had mixed DSR capacity (i.e. turn-down loads and non-renewable back-up generation on the same site). This could not easily be submitted to the second TA without investment in metering that would not be required for the main CM, unless the organisation submitted a declaration to National Grid that they would not use the back-up generation while turning down. As reported in Phase 3, few organisations made such a declaration because of the time and regulatory risks involved in getting such declarations accepted.

A further deterrent, mentioned by both participating and non-participating aggregators, was that capacity participating in the second TA was excluded from participating in the third T-4 auction (for delivery in 2020/21). This was a particular concern to those aggregators that expected low prices, and potentially low volumes, in T-1 auctions.

> *The TA was a bit of a joke, and it's not something we really wanted to participate in because of the way that CMUs in it were restricted from other activities [i.e. from some future T-4 auctions], meaning that the actual ability to learn from it was*

reduced almost to zero, so we didn't want to participate in it. (Non-participating aggregator, Phase 4)

A participating aggregator that generated additional outcomes through the TA (exhibiting several of the 'additional' reasonings) also thought that - with hindsight - they might have preferred to go straight into the main CM. It was difficult to assess, from the interview and auction evidence, whether they would have achieved the same degree of additionality through this route because the aggregator would have had a lower price and higher risk product with which to attract new DSR clients.

Other non-additional outcomes for H1

We found no evidence of participants in the second TA being deterred from participating in the main CM by their experience of the second TA. However, direct participants found the administrative burden of the CM heavy compared to other sources of flexibility revenue. Aggregator clients were largely sheltered from this burden because it was carried by their aggregator.

The CM is the one which I hold my nose and get on with. It's not facilitated for extra volume, it's difficult with the base lining, it's administratively difficult, you know, there are lots of rules, it's 247 pages of rules, of which I'm not as familiar with as maybe I ought to be....The practical consideration is that, you know, we just stumble through, effectively. (Direct participant, Phase 4)

Since all second TA participants went on to participate in the main CM, we found no outright evidence of second TA capacity being completely unavailable to the future CM because of the conditions of the main CM. However, there was evidence that certain aspects of the main CM could result in participants choosing to stop participating in the CM in future. These were future risks rather than observed outcomes, as follows:

- The higher level of credit cover required for unproven DSR in the main CM was a significant issue for smaller aggregators facing financial constraints.

At the moment, we put that [credit cover] up and we're a start-up, so it's a reasonable amount of cash to have to put not just at risk but the cash flow point of holding it for eight or nine months or something. (Aggregator, Phase 4)

- Lower clearing prices may make the main CM unattractive, despite the attraction of the TA. This was evidenced by the fact that some second TA participants submitted capacity to the 2018 T-1 and T-4 auctions but failed to clear at the prices of £6.00/kW and £8.40/kW, respectively. Relevant factors were their cost base and their access to flexibility revenues from other sources. Those without other sources of revenue were less likely to accept a low clearing price.

Contribution to H2: the second TA contributes to wider encouragement of turn-down DSR

Our main analysis of the additionality of the second TA is presented under H1 above. Within the analysis of H1, we have considered whether the second TA contributed to more turn-down DSR coming forward for the future CM.

There is one aspect of H2 that does not overlap with H1, and that is the question of whether the second TA brought forward turn-down DSR that will participate in other flexiblitiy services, as well as or instead of the CM. This is explored below,

Additional contribution to H2

Reasoning: we will seek to maximise other flexiblity revenues for the turn-down DSR that we have recruited for the second TA

This reasoning was observed for aggregators that recruited clients new to flexibility services (other than Triad cost-avoidance) for the second TA. Where clients new to flexibility were recruited, attracted by the high price in the second TA, aggregators then sought other revenue streams for these clients.

> *The idea of that is that we'd be able to offer various flexibility services on top of [the CM]. (Aggregator, Phase 4)*

In addition to evidence of clients contracting for flexibility services for the first time, there was evidence of clients switching aggregators in order to access a wider range of flexibility revenues. For example, one client had entered a contract with a proactive aggregator that offered CM participation, frequency reponse trials and energy management opportunities.

> *It's just encouraged me more, they're quite proactive on it. It's the same old thing these days, if you don't go and find the information yourself, you're never going to know. Whereas [aggregator] because they're obviously going to earn something out of it through ourselves, they're quite proactive in informing of these sort of things. (Aggregator client, Phase 4)*

As for H1, there was evidence from those aggregators less experienced with turn-down DSR that the second TA incentivised the recruitment of pure turn-down DSR, in a way that was different from other flexibility services.

> *I mean, the majority of demand-side response comes from, let's be honest, back-up diesel generators. So the focus is on those. They're easy. They're not environmentally friendly, but they're easy. So I think a lot of aggregators were focusing on that. So I would say that, yes, the TA has certainly incentivised a good focus on pure turn-down.(Aggregator, Phase 4)*

The second TA encouraged aggregators to look specifically for turn-down rather than just being agnostic as to whether DSR was provided by turn-down, back-up generation or energy storage. While the technology-specific focus on turn-down may not outlive the second TA, there was evidence that the turn-down capacity and knowledge gained would be carried forward.

> *If we hadn't [participated in the second TA], then it probably would be much harder to reach the same levels of pure turn-down capacity that we have now. It would take us, probably, multiple years. It's only going to be just because a customer is doing turn-down, not by design but by accident… I think that would be how additional turn-down would appear in our portfolio if that was not incentivised – last year – by the TA.(Aggregator, Phase 4)*

No additional contribution to H2

Reasoning: we are confident that we cannot access any additional revenue for our turn-down DSR capacity in the second TA, beyond the CM

This reasoning was observed for aggregators that were already active in the GB market with turn-down DSR, insofar as they were working with their existing clients or clients that were already experienced with a range of flexibility services. For example, one aggregator was already active in STOR while another offered frequency response services as their primary products. Nevertheless, these experienced aggregators were enabled to bring in some additional clients new to flexiblity, who were attracted by the high price of the second TA.

Summary of additionality of the second TA

The evidence presented above indicates that the high price and low credit cover offered in the second TA encouraged both new and existing aggregators to develop capacity and market their services to clients previously only turning down for Triads. The attraction of the second TA also encouraged clients with suitable turn-down loads to contract with aggregators to provide turn-down DSR, not just for the second TA but for the future CM and for other flexibility services. While some of this contribution would have been achieved through participants going straight into the main CM in the absence of the TA, the evidence suggests that the short-term growth in turn-down DSR would have been lower without the high price, low credit cover and low volume thresholds offered by the second TA.

4. Findings on HLQ3: did the second TA represent good value for money?

SUMMARY

While we present a commentary on 'value for money' below, a full assessment of the costs and benefits of the second TA (compared to alternative means of achieving its objectives) was not included in the design of this evaluation. Our limited assessment suggests that the second TA auction appears to have been:

• Expensive by comparison with recent CM auctions in GB, albeit with different auction objectives.

• Slightly more expensive than Demand-Side Balancing Reserve (DSBR), National Grid's interim balancing service for winter 2015/16, but cheaper than frequency services in GB, although these services differ significantly in their requirements.

• Comparable to prices paid for DSR in international capacity auctions.

• Possibly more expensive than it would have been if the target volume in the auction had been reduced by BEIS before the auction, because the supply curve appears to have been steep around the clearing price. However, it is uncertain whether this would have prompted further withdrawals of capacity pre-auction.

Introduction

Our assessment of value for money in the second TA scheme was limited, as in other phases of the evaluation. While we present a commentary on value for money below, a full assessment of the costs and benefits of the second TA (compared to alternative means of achieving its objectives) was not included in the design of this evaluation.

The ways in which we reviewed value for money, within the scope of this evaluation, were:

- By considering how far the second TA scheme contributed to its objectives (i.e. achieved the benefits intended).
- By comparing the clearing price paid to participants in the second TA scheme relative to the anticipated costs and risks of providing capacity.

- By comparing the clearing price in the second TA auction to the clearing price in other CM auctions (subject to the caveat that they procured other types of capacity, not just turn-down DSR).
- By comparing the price in the second TA with prices paid for turn-down DSR in other markets (both GB) flexibility markets and international CMs).
- By considering whether the second TA auction might have cleared at a lower price if BEIS had adjusted the auction parameters before the auction.

These approaches are explored in turn below.

Contribution to second TA objectives

Chapter 3 presents evidence about how, why and in what circumstances the second TA scheme contributed to its objectives. We concluded that the second TA scheme did make a significant contribution to both its objectives. However we need to consider other aspects of value for money to assess the cost-effectiveness of the second TA scheme as a means of generating this contribution.

Comparison with costs of capacity provided in second TA

The cost estimates presented in chapter 3 indicate that the initial capital cost of providing capacity for the second TA scheme ranged from minimal to £571 per MW, with the average cost to TA participants being £144 per MW (or £0.1 per kW). The cost of initial staff time inputs by aggregators and direct participants ranged from 'minimal' to £19,300 per MW, with the average staff cost being £4,800 per MW (i.e. £4.80 per kW). These costs do not include opportunity costs, costs incurred by aggregator clients or time inputs during the delivery year (e.g. SPD management), but they are well below the second TA clearing price of £45 per kW. The uncertainties in these cost estimates are explained in Appendix 6: it is possible that some other aspects of costs have not been recognised or reported.

Chapter 3 also presents TA participants' predictions of the ongoing staff time required for future CM participation by aggregators and direct participants. The predicted cost of ongoing staff inputs ranged from 'minimal' to £7,800 per MW per year, with the average predicted staff cost being £2,600 per MW per year (equivalent to £2.60 per kW). Again these costs do not include opportunity costs, costs incurred by aggregator clients or time inputs during the delivery year. With these caveats, the average ongoing staff cost is below the clearing prices for the Early Auction and for the T-1 and T-4 auctions held in 2018, but the higher end costs are close to or above the clearing prices in these auctions.

We know from the CM register and from interview data that some organisations taking part in the second TA scheme chose not to clear in the T-1 auction held in 2018, which cleared at £6/kW. This may indicate that these organisations simply had higher ongoing costs (within the range indicated above). Interview evidence suggests that other possible reasons for perceptions of higher costs (or the risk of higher costs) include:

- The perceived hassle and management risk of participating in the CM (e.g. time and effort spent understanding the requirements and checking compliance, in addition to the actual time spent organising tests, contracts etc).

The main Capacity Market auction prices, year ahead prices, have been very low. I think the value that we'd be getting from the next few years didn't look like it would be very cost effective. I know our costs are limited anyway… but equally, we're not offering many megawatts, so we need to make sure it's worthwhile for us to do it. (Case study, Phase 4)

- Potential risks or loss of production from turn-down, particularly if an extended turn-down was required. Considerations include risks to product quality, safety and customer satisfaction. While participants had the option of switching back on after a given period to mitigate such risks, this would involve some loss of TA revenues if there was an extended stress event.

At the end of the day, we will make a calculation against what the penalty is in the capacity mechanism, and the business will choose to serve its primary focus, if pushed. (Direct participant, Phase 4)

We've set a limit that we'll only turn down for 30 minutes maximum. We've agreed that with the aggregator. They balance that as part of their wider portfolio…It's purely down to safety and operational constraint. (Aggregator client, Phase 3)

- Opportunity costs for turn-down (as described in chapter 2) and the small risk of major failure arising from a turn-down event (e.g. old equipment failing to switch back on after turn-down).

If you have a motor failure on one of these [multi-megawatt] motors, then you're talking £0.5 million for a replacement motor. Then there are all the on-costs from the rest of the stuff – it may go into seven figures. We specify our kit, it's designed to start and stop many times, so it's designed to do that, and it's all about the straw that breaks the camel's back. (Direct participant, Phase 4)

While the risk of repeated stress events, and loss of TA revenues, was mentioned as a risk in the first TA research, respondents interviewed about the second TA scheme were aware that there had been no stress events since the start of the CM and did not raise this as an issue since they did not expect repeated events.

In summary, the revenues of £45/kW available to participants in the second TA scheme exceeded the actual costs of providing capacity and allowed a significant margin to cover the perceived risks of participation. The next sub-section compares clearing prices and clearing rates in the second TA with other CM auctions.

Comparison with clearing prices in other CM auctions

The clearing prices for and volumes of DSR clearing in all the CM auctions up to September 2018 are shown in Figure 4.1 below. The second TA auction clearing price of £45/kW was significantly higher than other auctions. Direct comparisons are problematic because the second TA auction procured turn-down DSR only while the other auctions procured DSR from both back-up generation and turn-down.

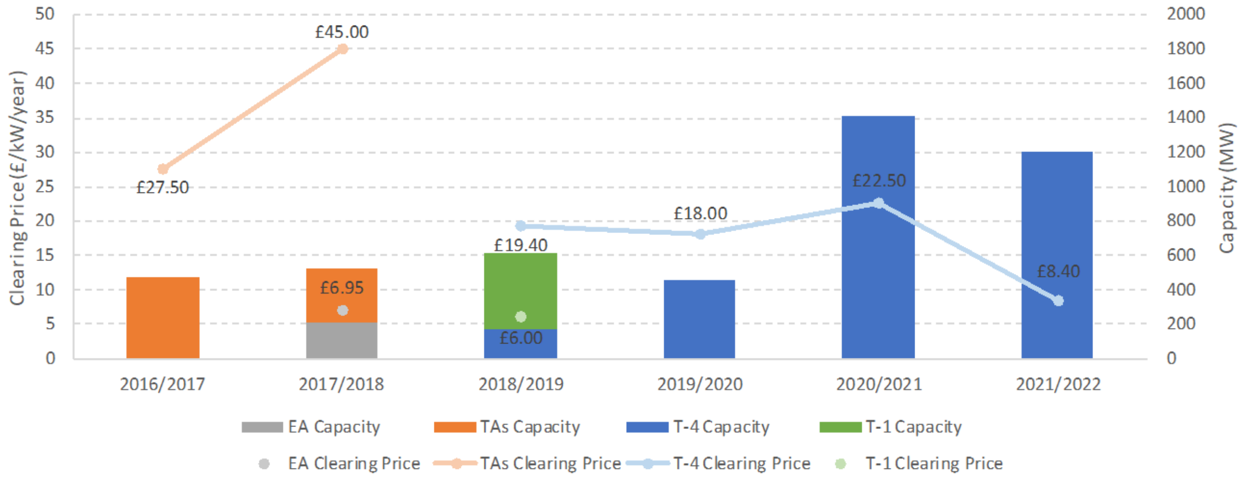

Figure 4.1: Clearing prices and DSR capacity awarded in recent CM auctions (source: CM register)

A small number of aggregators chose to put DSR capacity (including turn-down DSR) into the EA and/or T-4 auctions in place of the second TA, clearing at prices significantly below the second TA auction. Interview evidence from Phases 3 and 4 indicated that this was a choice made by:

- New supplier-aggregators that were not ready for the second TA auction.
- Experienced specialist aggregators that wanted to avoid exclusions between the TA and specific T-4 auctions, which would interrupt the annual stream of flexibility revenues available to their long-term clients.
- Specialist turn-down aggregators that offered frequency services with automatic controls from multiple small sites that had back-up generation on the same site (as it was not cost-effective to invest in separate metering for the second TA scheme only).
- Aggregators that did not have the skills and capability to put forward pure turn-down capacity or that preferred a strategy of putting forward mixed DSR to the CM.

As would be expected, higher clearing prices resulted in a higher proportion of DSR clearing in the second TA auction compared to other CM auctions. Figure 4.2 shows the proportion of DSR capacity clearing against clearing prices for DSR in the EA and first and second TA auctions (with the usual caveat that the second TA auction was restricted to turn-down DSR only). While 83% of capacity in the second TA auction cleared at £45/kW, only 29% of DSR capacity cleared in the EA auction at £6.95/kW.

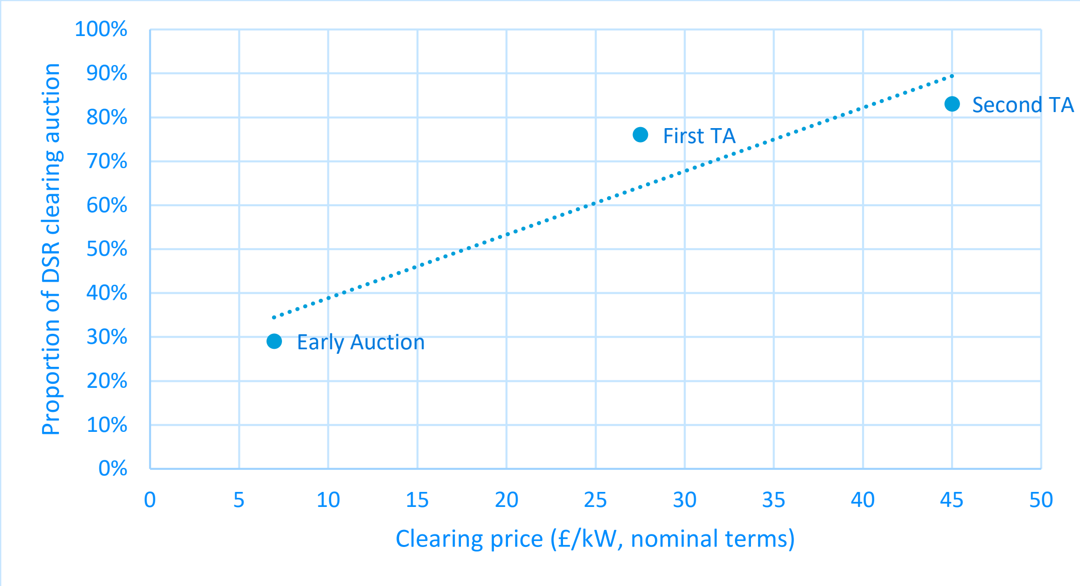

Figure 4.2: Proportion of DSR clearing in 1-year ahead auctions (source: CM register)

While there is evidence of some capacity from the second TA scheme proceeding to participate in the subsequent T-1 auction at a clearing price of £6/kW, there is also evidence of flexible capacity from the second TA failing to clear at this price. Interview data suggests that most participants and clients in the second TA scheme would seek a price higher than the recent T-1 and T-4 auctions (e.g. prices in the teens, above £10/kW, with a few participants reporting that they would look for £20/kW or £30/kW).

In summary, the second TA clearing price of £45/kW was high relative to other CM auctions and this seems to have stimulated recruitment of turn-down DSR capacity for the second TA. If future clearing prices for the main CM remain at or below £10/kW, some of this new turn-down capacity is unlikely to clear – unless other conditions in the CM are changed. Potential changes to CM rules are discussed further in chapter 6.

Comparison with other flexibility services in GB

The challenge of comparing prices for different flexibility services in GB is that different conditions and requirements apply to each service. The services vary widely both in the type of flexibility purchased (e.g. DSR technology, number and lengths of delivery periods, notice period, capacity thresholds, level of user commitment/control, contract length and so on) and in the types of incentives offered (e.g. availability payments, utilisation payments). The current revision of flexibility services under the National Grid's Service Needs and Product Strategy (SNAPs) review means that service specifications and prices are both fluid.

However, turn-down DSR was initially procured by National Grid through 'Demand-Side Balancing Response' (DSBR) contracts that were not dissimilar to TA requirements in terms of number and lengths of turn-down events and notice periods, although payments comprised both a utilisation and availability payment (in contrast to the TA which involves availability payments only). National Grid advised that the effective costs of DSBR contracts were around. £20/kW all-in, based on the total payments for DSBR (£2.309 million) divided by the total quantity (112MW) procured in 2015/16. However, the precise contractual terms for DSBR are not publicly available and it is unclear how far DSBR services can be compared to capacity provided by the TA schemes.

Other flexibility services are less directly comparable to the services provided by the TA schemes. For example, Short-Term Operating Reserve (STOR) contracts are open to both back-up generation and turn-down capacity and typically require a higher number of turn-downs during the winter period. Firm Frequency Response (FFR) services are also open to a range of technologies and typically require more turn-downs of short duration, at short notice. The revenue data presented in chapter 2 above shows that dynamic FFR services potentially offer higher values to those with suitable capacity (e.g. capacity suitable for frequent turn-downs using automatic controls).

While CM clearing prices have varied considerably since the introduction of the CM, ranging from £45/kW for the second TA auction to £6/kW for the Early Auction, there has also been variation in the prices of other flexibility services. For example, accepted tenders for STOR varied from zero to above £20/kW in the period April 2016-March 2017, while accepted tenders for dynamic FFR have varied from £20-80/kW over this period[27]. Over a multi-year timeframe, it is understood that both STOR and FFR prices declined significantly, with high prices at the time of launch followed by price decreases as competition increased.

Comparison with prices for turn-down DSR in international markets

We have also compared the second TA clearing price with prices for pure turn-down DSR in CMs in other countries. The charts below show clearing prices and turn-down DSR volumes in the well-established CM in Pennsylvania-Jersey-Maryland (PJM) in the United States. Figure 4.3 shows that the capacity cleared in the first and second TA was much lower than the capacity cleared in PJM but that clearing prices were similar to DSR in the PJM market.

[27] National Grid (February 2018). Power Responsive - Demand Side Flexibility- Annual Report 2017. (p14, Chart 2). Available at: http://powerresponsive.com/wp-content/uploads/2018/02/Power-Responsive-Annual-Report-2017.pdf

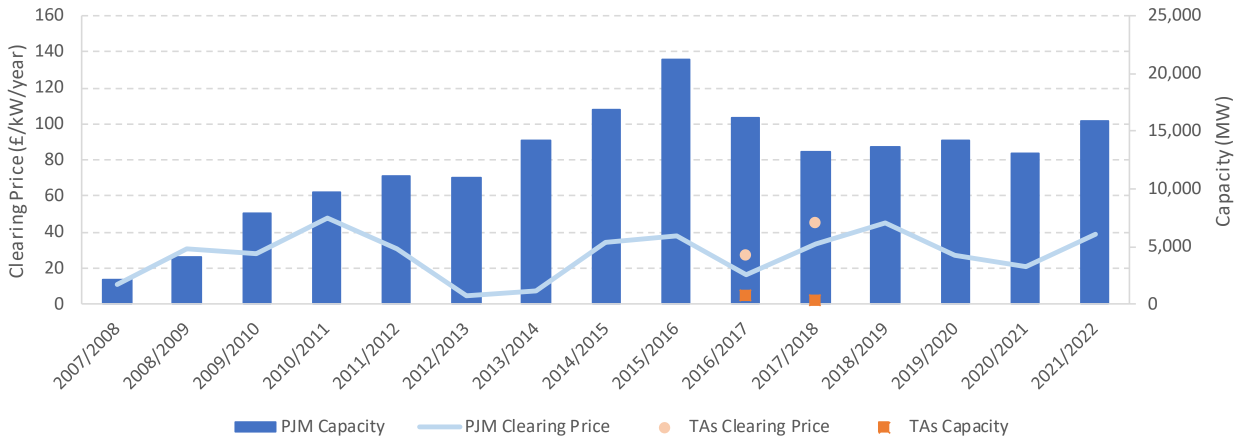

Figure 4.3: Clearing price[28] and awarded capacity in PJM CM auctions (source: PJM 2018; CM registers)

Comparison with the ISO-New England CM in the United States also shows that clearing prices are similar to the first and second TA auctions. Again, cleared volumes in the ISO-New England CM were higher than TA volumes.

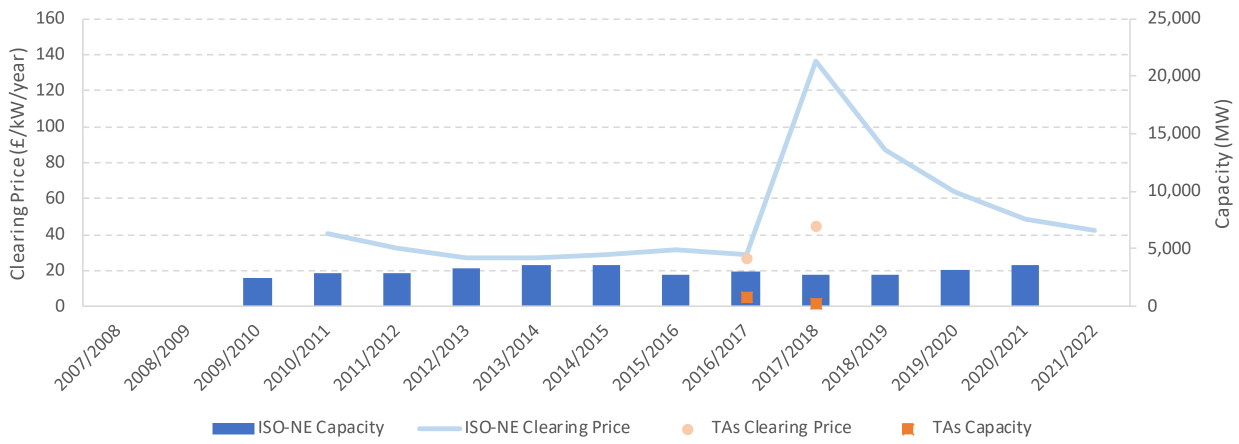

Figure 4.4: Clearing price and awarded capacity in ISO-NE CM auctions (source: Iso-NEW results; CM registers)

Prices in international CMs for DSR reflect subsidy programmes that vary greatly by location. For example, the Massachusetts/Rhode Island demand response programme, which has only operated for one year and is open to commercial DSR only, paid customers $20/kW plus $0.75/kWh to reduce during events. The system operator website suggests that the combined price received by customers was around $35/kW/year (equivalent to £26/kW at current exchange rates). Similarly, a demand response pilot programme has been running in New Mexico. This three-year programme, designed to encourage residential and small commercial customers to accept demand response on air

[28] Clearing prices in Figures 4.3 and 4.4 were converted at a fixed exchange rate of £1.32/$.

conditioning units with smart thermostats, pays a $125 (£90) enrolment incentive and $25 (£18) annual participation incentive.

This international evidence suggests that the price paid for turn-down DSR in the second TA was broadly comparable to prices paid in other CMs overseas.

Comment on auction parameters for second TA

Finally, we consider whether the second TA auction would have been more cost effective if the auction parameters had been changed. In principle, small perturbations to the demand curve could have significantly changed the auction clearing price.

BEIS were aware before the second TA auction that a significant volume of DSR had been withdrawn before auction. Figure 2.1 shows that 269 MW of pre-qualified capacity was withdrawn, leaving a reduced volume of 373 MW in the auction. If BEIS had reduced the target volume for the auction, and if bidding behaviour had remained unchanged (a strong assumption), this could have resulted in a lower clearing price. A similar adjustment took place before the first TA, but Phase 1 of the evaluation reported negative comments from aggregators about this adjustment. It is possible that a reduction in the target volume would have resulted in auction participants withdrawing further volumes from the auction (i.e. changing their behaviour).

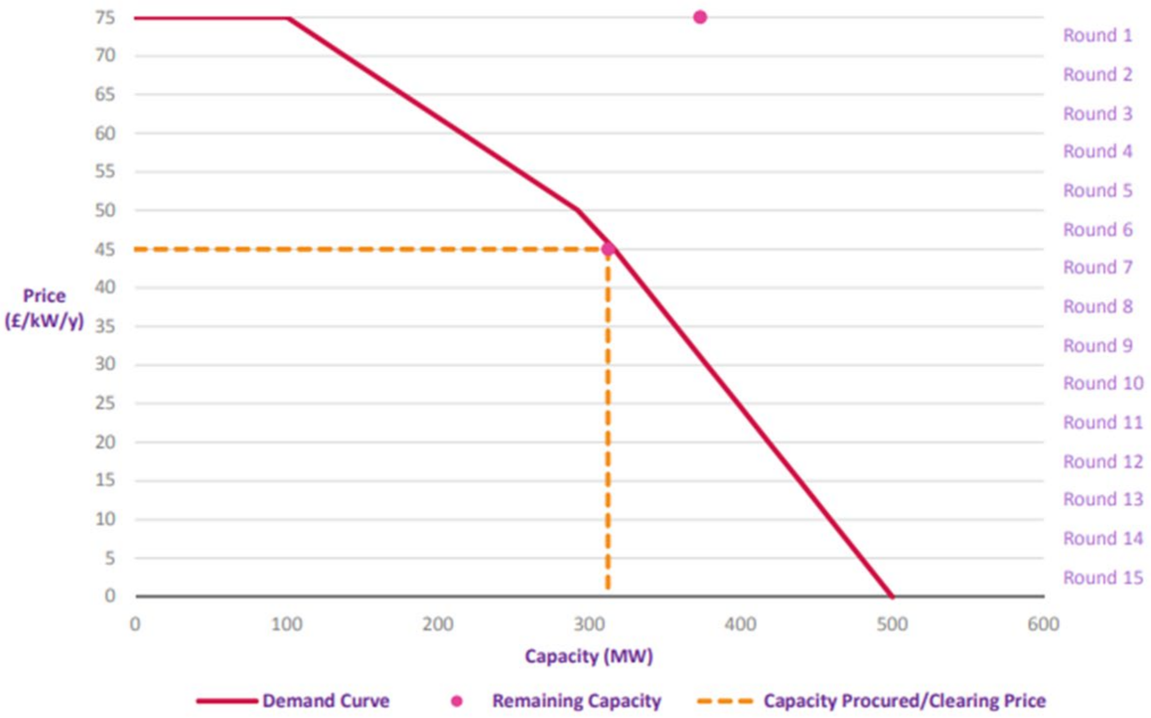

Figure 4.5: Demand curve and clearing price for the second TA auction (source: National Grid)

Summary on value for money of second TA

In summary, the second TA auction appears to have been:

- Expensive by comparison with recent CM auctions in GB, albeit with different auction objectives.
- Slightly more expensive than DSBR, but cheaper than frequency services in GB, although these services differ significantly in their requirements.
- Comparable to prices paid for DSR in international capacity auctions.
- Possibly more expensive than it would have been if the target volume in the auction had been reduced by BEIS before the auction, because the supply curve appears to have been steep around the clearing price. However, it is uncertain whether this would have prompted further withdrawals of capacity pre-auction.

5. Findings on HLQ4: influence of TA design and implementation

SUMMARY

Our findings about participant experiences of the second TA scheme, post auction, were that:

- **Metering tests were considered onerous, so participants with previous CM experience designed their CMUs to avoid the need for metering testing. This impacted on the type of sites that participants included in their CMUs (e.g. sites with renewables or onsite generation were avoided wherever possible).**

- **There was evidence of some participants designing CMUs to ensure that the 30-minute DSR and SPD testing requirements were met (with a potential impact on their ability to respond to an actual stress event, that might extend beyond 30 minutes).**

- **As noted in the Phase 3 report, timescales for DSR testing of capacity between the auction and start of the delivery year were tight. This, combined with problems with capacity recruitment, resulted in some capacity being lost before the delivery year (e.g. four CMUs withdrew after auction).**

- **SPD processes meant that SPD suspensions occurred even in cases where participants appeared not to be at fault (e.g. because of difficulty in establishing and maintaining the flow of meter data from sites to National Grid and EMRS, the body responsible for compiling meter test data on behalf of National Grid).**

- **DSR and SPD tests were not 'real-world' tests and did not fully assess the extent to which CMUs would be able/ready to respond to a system stress event. While participants were confident they could respond, there was evidence that some CMUs were primarily designed to meet testing requirements rather than necessarily respond to system stress events (which were regarded as an unlikely occurrence in the 2017/18 delivery year).**

This chapter focuses on aspects on TA participant experiences post-auction (the focus of the Phase 4 interviews). In particular, it focuses on:

- The reasons and factors behind CMU withdrawals from the TA post-auction.
- CMU 'white labelling' arrangements.
- DSR testing and SPD monitoring experiences.

- Factors likely to have affected participants' ability to respond to a system stress event.

CMUs exiting the TA post-auction

Four CMUs from three participants exited the second TA in the period between the auction and the beginning of the 2017/18 delivery year. Table 5.1 sets out the reasons why these CMUs exited and factors that contributed to these withdrawals (based on interviews with two of the three participants with CMUs that dropped out).

Table 5.1: Reasons and factors behind CMU withdrawals from the second TA post-auction

Reasons for CMU withdrawals	Factors contributing to CMU withdrawal
Not recruiting sufficient client capacity[29]	• Inexperience in aggregating turn-down DSR capacity • Recruitment more challenging than the aggregator envisaged • The short timescale to recruit capacity between the auction (March) and the delivery year (October) • Unsuccessful recruitment strategy (contracting out recruitment activity to a partner who failed to recruit any clients).
Not passing the DSR test	**Component-related factors** • Late installation of metering equipment, resulting in a late DSR test (shortly before the deadline) • Large variability in electricity demand (resulting in issues in demonstrating turn-down against the baseline) **CMU design** • Small CMU size (at the 0.5MW capacity threshold) i.e. there was no margin of error in the event of a poor DSR test result[30]. **Participants' approach and experience** • Tests being carried out late in process (insufficient time to undertake retests) • Lack of experience with testing processes and timescales • Ineffective preparation and planning, resulting in a test not carried out at optimum time of day to maximise turn-down against the baseline • Administrative errors (resulting in CMU ineligibility)

[29] Note that there were other instances of TA participants not recruiting sufficient capacity for their CMUs. In these instances, however, the TA participants 'sub-contracted' their CMUs to other TA participants. In these 'white label' CMUs, the contracted capacity was fulfilled with clients managed by other aggregators, as discussed in the next sub-section.
[30] In contrast, CMUs that did not prove their contracted capacity in the DSR tests, but nonetheless proved more capacity than the minimum 0.5MW threshold, were still eligible to participate in the delivery year.

CMU ‘white labelling’

The qualitative interviews highlighted instances where ‘white label’ arrangements were in place between TA participants and other aggregators. Under these arrangements, CMUs registered under one aggregator were effectively being managed by another aggregator.

Key reasons why aggregators had sub-contracted out their CMUs were:

- They were unable to fill one or more CMUs with their own capacity.
- They had strategic arrangements in place with partner organisations to manage one or more of their CMUs.
- Their business plan was to become an ‘aggregator of aggregators’, so this sub-contracting arrangement provided potential learning opportunities.

Reasons why aggregators had sought to manage other aggregators’ CMUs included:

- They had spare capacity that was willing and able to participate in the CM.
- The high clearing price for the second TA had attracted their clients.
- They had a strong motivation to participate in the CM in 2017/18 to help grow their aggregation business and learn about the CM.
- The CM rules allowed these arrangements to be developed.

Metering testing

The majority of the Phase 4 interview research did not explore metering test experiences in any detail. This was in part because the majority of participants did not require a metering test. For many of those that did, metering testing had already been explored in Phase 3 research interviews.

Testing overview

To be able to participate in the TA, all participants had to have a metering system installed that was compliant with CM regulations and rules. All TA participants had to prepare metering assessments, to provide National Grid/EMRS with information on metering arrangements on their sites. Only those with bespoke or balancing services metering, or with onsite generation (including renewable generation), had to complete a metering statement and undertake metering tests. For CMUs with supplier settlement metering configurations, no metering statements or tests were required.

Outcomes

Six CMUs that participated in the second TA completed metering statements and tests prior to the start of the delivery year, on at least one site. Interview evidence indicated that one additional CMU had to undertake a metering test during the delivery year as a result of a renewable energy asset being added to one of the CMU sites. The other CMUs had supplier settlement metering configurations and therefore did not require testing.

Reasons for outcomes

The Phase 4 findings confirmed the findings from Phase 3. These were that:

- A low proportion of CMUs had to undertake metering testing because participants with experience of the CM generally avoided or minimised metering testing by selecting sites that had supplier settlement configurations.
- Where metering tests were undertaken, this was generally for larger sites where the investment of time and money in metering was cost-effective.
- Where metering statements and tests were required, the process was considered to be onerous.

As highlighted in Table 5.1, late installation of metering equipment contributed to one CMU exiting the TA. The CMU's DSR test was subsequently undertaken late and, after it failed the test, there was no time to undertake a retest.

DSR testing and SPD monitoring

Testing and monitoring processes

DSR testing overview

A DSR test is a test specified in the CM rules. It is taken by participants with unproven DSR CMUs to prove that a CMU can achieve its stated capacity. The test involves compilation of baseline data over a six-week period, followed by collection of meter data by EMRS to confirm reduction below the baseline level at agreed times. The date and settlement period of the DSR test are chosen by the capacity provider. CM rules state that the DSR test must take place at least one month prior to the start of the delivery year. CMUs that fail to prove their contracted capacity in a DSR test can retake the test, provided it is done within the testing timescales outlined above[31].

SPD monitoring overview

Participants in the CM have to demonstrate that they have met the capacity obligation which they acquired at auction on three separate dates. This capacity has to be demonstrated for at least one settlement period on each of those dates. The data from the three SPDs is compared to the six-week baseline demand for the unit. SPDs must be submitted between 1st October and 30th April of the relevant delivery year. If satisfactory

[31] For more information about the DSR test process see: National Grid (2017), Capacity Market DSR Testing Process, June 2017. Available from: https://www.emrdeliverybody.com/Lists/Latest%20News/Attachments/107/DSR%20Test%20Guidance%20Document.pdf [Accessed 21 June 2018]

performance is not demonstrated by the relevant deadline, three additional SPDs must be submitted after 1st May, or at any time in a subsequent delivery year. Capacity payments are suspended from 1st May from until the later of 1st June and the day in which the SPDs are all demonstrated.[32]

Outcomes

Of the 28 CMUs participating in the TA for the 2017/18 delivery year, 26 were unproven DSR CMUs and therefore required DSR tests. Figure 2.1 shows that little capacity was lost in the DSR tests for the second TA.

As highlighted in chapter 2, six CMUs were suspended for failing to demonstrate three SPDs before the end of April 2018. Two of these were rectified, so only four CMUs (8 MW) were still suspended in September 2019. As the majority of the fieldwork for Phase 4 was undertaken prior to end April 2018, SPD monitoring results had not been confirmed and these suspensions had not yet been announced at the time of the research interviews.

In the majority of cases, participant experiences were similar for both DSR testing and SPR monitoring. This is because of the significant overlap in the two processes. As a result, this section combines findings in relation to both DSR testing and SPD monitoring.

Factors impacting on testing and monitoring experiences

Table 5.2 presents the factors that impacted participant testing and monitoring experiences.

Table 5.2: Factors impacting on TA participant DSR testing and SPD monitoring experiences

Component-related factors	
Capacity 'buffers'	Clients and direct participants with a conservative approach to putting forward capacity for the TA ensured that their components had sufficient 'reserve' capacity to meet testing and monitoring requirements even if circumstances were far from ideal. These capacity buffers contributed to the 'overfilling' of CMUs, as they meant that components could prove more capacity than they were contracted to deliver.
Corporate and site-level commitment	Corporate and site-level commitment to ensuring they met testing and monitoring obligations contributed to over-delivery of capacity against contracted requirements.
Turn-down DSR experience	Components with pre-existing experience in providing turn-down DSR capacity – for example through Triad management or balancing services - were well-equipped to meet their DSR testing and SPD monitoring obligations.

[32] For more information about the Satisfactory Performance Days monitoring process see: National Grid (2017), *Capacity Market Satisfactory Performance Monitoring*, 3 April 2017. Available at: https://www.emrdeliverybody.com/Lists/Latest%20News/Attachments/97/Satisfactory%20Performance%20Guidance%20Document.pdf [Accessed 21 June 2018]

Component-related factors	
Demand cycles	Components with cyclical, or 'spiky', demand cycles were often able to meet turn-down DSR test and SPD monitoring requirements as part of their normal production schedules.
Compatibility with Triad	A number of respondents highlighted that their components were able to perform well for DSR testing and SPD monitoring because turning down for these tests was compatible, or aligned with, their Triad management schedules.
Seasonality	Seasonal differences in production cycles sometimes aided, or hindered, the ability of components to meet DSR test and SPD monitoring requirements. For some components, demand was higher during certain periods of the year, which meant they could demonstrate greater levels of turn-down (e.g. for agricultural sites used to heat produce, meaning electricity demand was higher in the winter, or for cold stores, which had higher demand in the summer).
Non-seasonal component changes	Changes to components during the delivery year could result in the need for components' metering requirements to change. (e.g. one CMU needed to undertake a metering test during the delivery year because of the introduction of renewables to the site). Demand levels could also vary throughout the year, affecting component capacity.
Technical factors	Technical issues – for example with equipment signalling, turn-down timings and data processing - led to challenges in providing sufficient turn-down.
Extent of 'overfilling'	As highlighted in the Phase 3 evaluation report, some participants chose to 'overfill' their CMUs, whereby they created CMUs that had more capacity to turn-down than the contracted amount. This strategy helped some to 'de-risk' TA participation and enabled them to successfully meet DSR testing and SPD monitoring requirements, as CMUs could demonstrate contracted capacity even if one of more of their components underperformed.
Number of CMU components	Participants with multiple components within one CMU, or across multiple CMUs using joint testing, could face additional challenges when conducting DSR tests and SPD monitoring, compared with single component CMUs. The key challenge was coordinating with, and aligning, multiple components to ensure they could all turn-down effectively at the same time on the same day.
Testing approach (individual vs joint)	Participants with more than one CMU were permitted under CM rules to undertake joint CMU testing. Joint CMU testing allowed TA participants to 'de-risk' DSR testing and SPD monitoring by spreading the risk of under-delivery across a number of CMUs. This allowed any CMUs that

Component-related factors	
	over-delivered to compensate for CMUs that under-delivered.
Client notice and consultation	Aggregators and clients highlighted that advance notice of, and consultation about, planned DSR testing and SPD monitoring times helped components to perform effectively for DSR turn-down. Sites could ensure that procedures were in place, production schedules were optimised and staff were all aware of what was required. Consultation between aggregators and clients helped to ensure tests and monitoring were conducted at optimal times for CMU components, and in a manner which respected client business demands.
Timing of tests and monitoring	Not conducting DSR tests and SPD monitoring turn-downs with sufficient time ahead of the deadlines was an issue that contributed to challenges in completing the tests.
Capacity Market experience	Inexperience in the CM was a factor underlying a number of the challenges that TA participants faced. Inexperience, for example, had led to a lack of clarity about CM baseline rules, hindering DSR test effectiveness.
General	Issues with data flows were the most common issue raised by TA participants in relation to undertaking DSR tests and SPD monitoring. Data flow issues resulted in delays in progressing through the DSR testing process. With regards to SPD monitoring, whilst the deadline for this had not passed at the time of all of the interviews, several TA participants reported that CM payment suspensions were likely as a result of data flow issues.
Setting up and maintaining data flows	Setting up data flows was a common challenge amongst TA participants. Complaints included delays with suppliers setting up the data flows, or data flows not being set up even when suppliers had confirmed they were. Participants also highlighted that data flows had been cancelled by suppliers without any notification. Furthermore, these cancellations had not been picked up EMRS.
Data flows through EMRS and National Grid	TA participants also raised a number of issues in relation to data flows through EMRS and National Grid. These included:

[33] TA participants with supplier settlement metering configurations were required to set up data flows with their electricity suppliers so that EMRS could access and review half-hourly meter data for each CMU. The process was for TA participants to notify their supplier that they required their metered volumes to be used in the Capacity Market. The supplier then needed to inform the appointed Half-Hourly Data Aggregator (HHDA) to provide metered volumes to EMRS. The metered volumes could then be downloaded by the appointed Half-Hourly Data Collector (HHDC) who could then pass them on to the appointed HHDA. The HHDA could then submit the metered volumes to EMRS. For more information see: EMR Settlement Limited, Low Carbon Contracts Company and Electricity Settlements Company (2017), *WP195 – Capacity Market and CFD Metered Data, EMRS Working Practice*, Version 4 – 6 December 2017. Available at: https://www.emrsettlement.co.uk/documentstore/workingpractice/wp195-capacity-market-cfd-metered-data.pdf [Accessed 24 June 2018]

Component-related factors	
	• EMRS sending incomplete data to National Grid. • Long turnaround times by EMRS. • Inaccurate data calculation by National Grid. • EMRS using estimated meter readings, rather than actual meter data, which adversely impacted on initial DSR test results. • EMRS not identifying halts in data flows; where suppliers had cancelled data flows with no notice, TA participants were frustrated that these interruptions had not been picked up EMRS. • One of the issues compounding the data flow problems was that TA participants themselves had no visibility of the data being sent to the EMRS by the HHCAs. This lack of transparency had hindered TA participant's efforts to understand if, how and why data flow data flow issues had occurred.
Not 'real-world' tests	The rules for both DSR tests and SPD monitoring allowed participants to select when they demonstrated their turn-down. This benefitted them because they could select optimal times for turn-down (i.e. times when they were most likely to be able to turn-down effectively). Thus the tests and monitoring did not simulate potential 'real-world' system stress events, which could happen at any time and with only four hours' notice.

CMU capability for system stress events

System stress event overview

A system stress event occurs in response to a national shortage of generation resources. The National Grid defines a system stress event as occurring when: "i) a demand control event has occurred and ii) that demand control event has been confirmed after post-event analysis, conducted by National Grid, to have been definitively triggered by a national shortage of generation resources."[34]

System stress events are thus only confirmed post-hoc. If the National Grid, as System Operator, thinks there is a risk of a forthcoming system stress event, they will issue a Capacity Market Notice (CMN). A CMN is a signal four hours in advance that there may be less generation available than National Grid expects to need to meet national electricity demand. It is intended to be a signal that the risk of a system stress is higher than under normal circumstances. Where a CMN was active, not subject to a cancellation until a later time, and was found to have developed into a system stress event, participants in the

[34] https://gbcmn.nationalgrid.co.uk/faq/system-stress-events/what-is-a-system-stress-event

second TA would be obliged to respond by delivering reductions in their demand equivalent to the capacity they had contracted to deliver for each CMU.[35]

Participant expectations about system stress events

Expectations differed amongst TA participants and aggregator clients about the number of system stress events that might occur during the delivery year.

Typical advice by aggregators to clients was to plan for around five system stress events per year. This advice appeared to be given mainly to manage client expectations and to ensure that they were prepared for this eventuality. TA participants generally thought that the actual likelihood of a system stress event occurring was low.

Participant confidence about ability to respond

TA participants were generally confident that they would be able to respond effectively if a system stress event occurred. Participants that were less confident (all aggregators) highlighted that:

- Turn-down DSR is inherently less reliable than existing generation, and
- Turn-down DSR capacity in the TA had not been tested under ‘real-world’ system stress event conditions, so its ability to respond was unknown to some degree.

There was also a view amongst both aggregators and aggregator clients that their ability to respond would depend on the circumstances at the time (e.g. time of day, time of year, opportunity costs of turning down, length of system stress event and so on).

TA participants also highlighted significant variation in the length of time for which components in the TA were capable of turning down. ‘Best case’ estimated turn-down times ranged from a maximum of 30 minutes for one component to up to 48 hours.

Factors impacting on confidence about ability to respond

Table 5.3 presents the factors that impacted participant confidence about their ability to respond to a system stress event.

Table 5.3: Factors impacting on TA participant confidence about ability to respond to a system stress event

Component-related factors	
Capacity ‘buffer’	As highlighted in Table 5.2, components with significant capacity ‘buffers’ were more confident they could effectively turn-down if required. One respondent highlighted that, with this approach, a best-case scenario might mean that the

[35] https://gbcmn.nationalgrid.co.uk/faq/system-stress-events/what-are-capacity-market-participants-obligations-during-a-system-stress-event

Component-related factors	
	component would not need to take any action if a system stress event was called. A worst case, however, would result in having to turn-down by three times as much as its contracted capacity, because of the negative DSR baseline effect (see below). But their conservative approach to defining their capacity would allow them to deliver even under this worst-case scenario.
Extent of flex	The extent to which a component had 'spare' capacity, or flex, at any given time within its natural operating processes (i.e. the extent to which a component was at 'full-stretch' or not). Greater amounts of flex meant higher certainty of being able to respond to a turn-down request.
Technical constraints	The extent to which the processes involved had technical limitations which would impact on the ability to respond for different lengths of time (e.g. cold stores had a built-in buffer and could hold temperatures for longer periods of time, while chillers were more temperature sensitive and had less ability to hold temperatures for a long period without any energy input).
Corporate commitment	Extent of corporate and site level commitment to comply with a request to turn-down for a system stress event (see also Table 5.2).
Turn-down DSR experience	Extent of prior experience at providing turn-down DSR. The greater experience, the higher the chances of being able to respond.
Interaction with other services	Extent of component interaction with other services (including Triad, balancing services, red zone management, etc), and the extent of compatibility between these services and the TA requirements (see also Table 5.2).
Variability of demand profiles	Components with variable demand profiles would have different levels turn-down potential depending on the: • Time of day • Day of the week • Time of year Components with more stable demand profiles across the day and week were more confident in their ability to respond regardless of when a system stress event happened.
Potential opportunity costs	The costs of lost production or other services was a big determinant of the extent to which a component could turn down. The greater the

Component-related factors	
	opportunity costs, the less likely the component would be to turn down. These costs could vary throughout the year depending on customer demand levels and other business requirements.
Stock levels	For some processes, the levels of stock supply were an important factor affecting how long a component could potentially turn down for. The greater the stock levels, the longer the turn-down potential.
Regulatory constraints	Regulations could also impact of the extent or length of potential turn-downs. For chilled stores, for example, regulations about the storage of food at certain temperature was a constraint to turn down.
Extent of 'overfilling'	The extent to which a CMU was 'overfilled' with capacity. The greater the overfilling, the greater the chances of being able to respond (see also Table 5.2).
Diversity of CMU components	Having a diverse range of component types was regarded as a strategy to de-risk the ability of CMUs to respond.
Criteria for component selection	The extent to which components and CMUs were selected and designed to meet the DSR test and SPD requirements, versus the extent to which they were selected and designed to provide reliable capacity for a potential system stress event. The interviews highlighted that aggregators placed differing levels of emphasis on these choices. For some aggregators, fulfilling the test requirements was the key emphasis.
Oversight and understanding of component capability	Extent to which an aggregator or direct participant understood their components and their turn-down capabilities (see also Table 5.2).
Contractual requirements	Extent to which contractual requirements between aggregators and clients would incentivise clients to turn down for a system stress event.
Client expectation management	The expectations that aggregators set with clients about the likelihood of system stress events occurring and their potential length. As well as the extent to which clients understood or believed them.

Component-related factors	
Client management strategies	The extent to which clients were supported to ensure they were able to respond successfully if required.
Number of activations required under the TA	Participants highlighted that a relatively low number of turn-down activations was required in the TA. This was felt to increase the likelihood of components being able to respond to a system stress event: as the turn-down burden on participants was low, it would not be a significant burden to clients to perform a one-off turn-down in addition to minimum testing requirements.
Notice period	The four-hour notice period for a potential system stress event was an important facet of the CM for some components, which highlighted that they would not be able to respond if the notice period was short.
DSR baseline methodology	There was a view from TA participants that the DSR baseline methodology meant that it was harder for turn-down DSR to participate in the CM than generation. Generating capacity could simply remain on, or turn on, in order to respond to a system stress event. Whereas for turn-down DSR, even if it was completely off during a system stress event, compliance was based on whether this represented a reduction in capacity against its baseline equivalent to its contracted capacity. If the component also happened to have a low or zero demand baseline, therefore, it would not be compliant.

6. Findings on HLQ5: Implications for future of turn-down DSR in the CM

SUMMARY

We found that the turn-down DSR participating in the second TA was almost exclusively provided by large industrial loads. While the high price for the second TA stimulated extensive marketing by aggregators, interviews with industrial non-participants in Phase 3 indicated some unrealised potential. However, we found that commercial and public sector loads, such as HVAC, can typically tolerate only short turn-down periods (e.g. 30 minutes or less) and tend to require automatic controls. We did not observe aggregation of commercial loads for the second TA except in rare cases where the loads required no investment in metering or controls and were also generating revenue from other services (e.g. frequency services).

While modest volumes of turn-down DSR cleared in the 2018 T-1 auction at £6/kW, these were single large sites with low costs and/or access to revenue from other flexibility services. Interview and auction evidence indicated that future CM prices of £10-20/kW, closer to those observed in the first three T-4 auctions, would be needed to support recruitment of new turn-down DSR and investment in new controls for small sites. Viability could be adversely affected if participants became unable to stack CM revenues with other revenues.

Potential changes that could encourage participation of turn-down DSR in the CM, in combination with higher prices, included: limits to the duration of turn-down DSR offered; streamlining of metering requirements for DSR in the CM, particularly for small sites; reduced credit cover; and less stringent baseline requirements for DSR in the CM. More flexibility for changing the composition of proven DSR CMUs, as planned by Ofgem, will also be supportive. Longer agreements (2-3 years) would be welcomed but we understand that BEIS, in designing the TA, regarded this as unjustifiable because of the low up-front costs of DSR.

Changes that could discourage participation by some types of turn-down DSR but might be advantageous to the CM's role in supporting security of supply include: higher penalties for non-delivery; testing regimes that more accurately reflected potential delivery during a stress event; and reduced notice periods for a system stress event.

External factors that may stimulate future growth in turn-down DSR include new emissions regulations for diesel generation, new control technologies making smaller sites cost-effective and new/cheaper battery technologies.

This chapter considers the implications of the evaluation findings for the future contribution of turn-down DSR to the CM and other flexibility services. BEIS posed a number of specific evaluation questions about the contribution of turn-down DSR to the CM, which are explored in turn below. These are:

- What is the potential size of the turn-down DSR market and its potential to contribute to security of supply in the years following the TA?
- Will turn-down DSR be self-sustaining and self-perpetuating (i.e. viable) in the CM?
- Will turn-down DSR evolve using new technologies in the CM, with no further Government intervention and support?

Our response to these questions, as set out below, is limited insofar as it is based on the evidence emerging from the evaluation. A fuller response to these questions is beyond the scope of the current evaluation.

What is the potential size of the turn-down DSR market and its potential to contribute to security of supply in the years following the TA?

While we interviewed non-participant aggregators and potential client organisations during Phase 3, we have not undertaken a statistically reliable survey of industrial and commercial organisations that would have potential to provide turn-down DSR. Surveys undertaken by the Energyst magazine reported an increase in turn-down DSR between 2016 and 2018[36]. However, care must be taken in interpreting this finding as the sample was small and only involved Energyst readers, so may not be representative of the wider industrial and commercial market as a whole. The Association for Decentralised Energy (ADE) estimated the potential size of the overall DSR market (from turn-down, back-up and battery storage in the UK) as just over 5 GW for industry and over 1.5 GW for services.[37] In June 2018, BEIS commissioned independent research on non-domestic turn-down DSR which is expected to generate additional insights on the market.

[36] "60% of DSR participants use back-up or onsite generation. However, around three quarters say they decrease consumption. A third say they increase consumption or turn loads on. Some participants combine all three types of response. The generation versus load finding is almost the exact opposite of last year's survey (76% via generation, 59% decrease consumption)." Energyst DSR Survey 2017. In the Energyst DSR Survey 2018, 85% of respondents involved in DSR reported decreasing consumption, compared to 40-50% reporting use of distributed generation and less than 5% storage. (Available at https://theenergyst.com/dsr).
[37] Association for Decentralised Energy and Renewable UK (June 2018). *Industrial flexibility and competitiveness in a low carbon world.*

Potential size of the turn-down DSR market

Our analysis of TA data indicates that the turn-down DSR brought forward by the second TA was provided almost exclusively by industrial rather than commercial loads. These two sectors are considered in turn below.

Industrial turn-down

The sectors and types of industrial loads that participated in the second TA are described in chapter 2. These industrial loads were large (typically 100 kW or more) and shared some common features in that they involved:

- Batch production or a production process with some element of spare capacity relative to customer demand or final production; and/or
- Ancillary processes that could be stopped temporarily without affecting the quantity or quality of overall production; and/or
- An element of storage (e.g. storage of final product; or storage of intermediate products; storage of heat or cold; or storage of water).

We have strong evidence that, while there were a number of drivers and enabling factors for increased turn-down DSR[38], the high price in the second TA stimulated extensive marketing by aggregators. This was corroborated by evidence from aggregator clients in the second TA that were approached by several aggregators.

I believe we were approached by four [aggregators]. Well, we were approached by about a dozen, but we were seriously approached by four. We had serious discussions with three of the four, and the company that we finally went with were offering us a slightly better percentage return to us than the other two companies. (Aggregator client, Phase 3)

But there was evidence from Phase 3 interviews with non-participants that there were still some industrial companies with apparently suitable loads that had not yet started participating in flexibility because their senior management were concerned about the risks to their main business and were not persuaded of the benefits of DSR.

Yes, I mean, most of our sites have got very large chilled storage, so enormous fridges with several megawatts of cooling ...We have a lot of HVAC that could go onto some kind of turn-down, and then we've got our CHP engines and other back-up generation which could be used…. [We] just probably didn't have enough resource to really put the business case and keep pushing our board, and maybe there's just a lack of an understanding at senior level as to what it [DSR] is and why we should be doing it. (Potential TA participant, Phase 3)

[38] See chapter 3 for discussion of the influence of the second TA compared to other drivers such as Triad charges and the Medium Combustion Plant Directive and other enabling factors such as the Power Responsive Campaign and recent changes to rules about the compatibility of frequency services and the Capacity Market.

Others had potential to participate more but were relatively new to the concept of DSR (and specifically turn-down DSR) and were reluctant to accept to offer services such as frequency response that required automatic control of assets by an aggregator.

It has been suggested that we may go as far as frequency response at a later date once the production sites are comfortable with the whole concept. Initially, the production sites were not comfortable with having some of their equipment going offline, effectively outside of their control.(Aggregator client, Phase 3)

We also found evidence of experienced turn-down DSR providers planning to optimise their flexibility offer and 'design in' flexibility to their plans and future investments, which could result in them offering more DSR in future.

I mean if we can connect DSR into our planning and scheduling, so we could sort of engineer it into our processes, really, then it's becoming much more interesting… Participation in DSR schemes will become an ingrained aspect of what we do instead of being something imposed on it.(Aggregator client, Phase 3)

In summary, while volumes of industrial turn-down DSR appear to have increased significantly in recent years, there is still some potential for greater volumes to come forward – particularly from smaller sites that could not be included cost-effectively in the second TA.

Commercial turn-down

Detailed analysis of the capacity put forward for the second TA revealed that very few sites involved turn-down of commercial loads. The commercial sites that did participate involved automatically-controlled HVAC loads within large buildings turning down for short periods. The second TA sites were unusual in that sub-metering had already been installed for general energy management purposes, making participation in the CM cost-effective. There are few commercial or public sector buildings with loads big enough to provide turn-down DSR cost-effectively in the CM.

As soon as demand management was becoming available, emerging in the marketplace as an option for us, we did look round and think, "Right, what can we do with regards to this?" It turns out that, really, we don't have any single large pieces of equipment that could be easily plugged into this. That was quickly put into the side...(Public sector non-participant, Phase 3)

We provide a fully-funded solution for a large commercial building if you can shed 400kW or more." Then once we did an audit and looked at the case studies, we thought, "That's a bit too much. You can't find a lot of those buildings." [..] now, we provide a fully-funded solution for 200kW or more [for frequency services, not CM]. (Non-participant aggregator, Phase 3)

Commercial turn-down is active in the flexibility market, despite its absence in the second TA. Published case studies and interview evidence from both TA and non-TA aggregators indicate that a number of commercial companies with large electrical loads and automatic controls provide turn-down for dynamic or static frequency response services.[39] These services generally involve turn-down for short periods (e.g. split-second for dynamic Firm Frequency Response; several minutes for static Firm Frequency Response; and a maximum of 30 minutes for Frequency Response by Demand Management (FCDM)). This is consistent with ADE findings[40] that 'service sector' DSR (in the commercial and public sectors) is generally best suited to delivery DSR for periods less than 30 minutes.

In theory, it would be possible to for aggregators to provide longer periods of turn-down (as required for the CM) by sequencing multiple short turn-downs by commercial loads, or (following recent CM rule changes) to meet CM requirements while contracted for frequency services. But interviewees advised that it was not cost-effective to bring in small sites below 100 kW to the CM, because of the cost of installing controls and metering requirements.

To be honest, if it's one client that's bringing you 100 kilowatts it's not worth their while, because the revenues they're going [to] get are just not worth the hassle of going through the process to get live. (Aggregator, Phase 3)

At the low prices observed in the recent CM auctions (£6-9/kW compared to £45/kW for the second TA), aggregators reported that it was more problematic to bring new clients on board at these prices, particularly for smaller distributed loads which did not already have controls and metering in place.

At those low prices for smaller distributed load, you're almost looking for opportunities for loads where everything requires a control already in place, because the pricing is so low you're looking to almost utilise existing flexible capacity, existing infrastructure to leverage that so you can use it to take part in [CM]. The potential revenues in the near term make it hard to make a case to spend any money to go into the market.(Aggregator, Phase 4)

While public sector organisations participated with back-up generation in the first TA, there was no public sector turn-down DSR in the second TA. The Power Responsive campaign has expressed particular concern that potential for DSR in the public sector remains untapped, despite potential for providing flexibility using both back-up generation and HVAC.[41]

In summary, evidence from participants and non-participants in the second TA suggests that commercial and public sector loads, such as HVAC, can typically tolerate only short turn-down periods and tend to require automatic controls to ensure that temperatures stay within agreed limits. Given recent CM rule changes, these loads can now stack CM revenues with frequency service revenues. However, greater participation from these

[39] For example, several commercial turn-down case studies are published by the Power Responsive campaign at: http://powerresponsive.com/case-studies
[40] ADE and RenewablesUK (June 2018) (op cit) – Figure 10, page 24.
[41] "Public sector sites are not yet participating at scale (including: MoD, NHS, and Local Authorities); with seemingly untapped potential available (e.g. back-up generation, HVAC)." http://powerresponsive.com/wp-content/uploads/2018/02/Power-Responsive-Snapshot-Asset-Led-Perspectives-on-DSF.pdf , February 2018.

sectors will require high revenues to cover the cost of investment in controls and sub-metering for small sites, or advances in technology that bring down the cost of these investments.

Other influences

As outlined in chapter 3, there are many factors that will affect the scale of the future market for DSR, and particularly turn-down DSR. These include:

- Possible reductions in Triad revenues as a result of Ofgem's ongoing review of embedded benefits (which could significantly reduce the main source of flexibility revenue for most DSR providers).
- CM rules, and the extent these are changed by the upcoming CM review (see next sub-section).
- National Grid's evolving SNAP, including product road maps for frequency response and reserve products.
- Ofgem proposals to broaden supplier licensing for aggregators, providing wider access to the Balancing Mechanism.
- Recent and upcoming tenders for flexibility services by the distributed network operators (DNOs), as part of their transition to become distribution service operators (DSOs).
- The Medium Combustion Plant Directive which requires new diesel plant to meet stricter emissions standards and makes turn-down DSR relatively more attractive.
- Cheaper and better electricity storage technologies and increasing take-up of electric vehicles.
- Cheaper and better control technologies for HVAC (as discussed above) and for smaller residential and commercial loads.
- The introduction of the European balancing platform (project TERRE) in December 2019.

The overall economic situation will also have an impact on provision of turn-down DSR. Providing turn-down DSR is less feasible during an economic up-turn, when plants are running close to their capacity, as there is less scope to shift load and make up production at another time.

Potential for turn-down DSR to contribute to security of supply

In the absence of a stress event, we do not have direct evidence about the extent to which turn-down DSR has contributed to security of supply via the second TA. As outlined in the reliability section in chapter 5, qualitative research indicates that experienced aggregators and DSR providers built a significant safety margin into their contracted capacity, to ensure that they could deliver if required. But there remains a risk that any given DSR provider might not be able to deliver their obligated capacity for a particular event because of unforeseen circumstances or because of their level of baseline demand. This risk would increase with the length of a given turn-down event and would be decreased by the level of penalties that applied for non-delivery. Further insights into the implications of increasing penalty levels are given in the next sub-section.

Will turn-down DSR be viable in the CM?

We have considered whether turn-down DSR will be self-sustaining in relation to a number of potential reforms that are being considered in the upcoming CM review.

Viability in the main CM

Some of the turn-down DSR that participated in the second TA did not clear in the recent T-1 auction at £6/kW. While some turn-down DSR, with low costs and/or revenue streams from other services, could still participate, aggregators suggested that few new clients would be attracted at these low prices.

> *It [price expectations] doesn't come so far down that we have a bunch of customers that would want to participate at £6. I think £6 is extremely low, and that's a very difficult sell.(Aggregator, Phase 4)*

> *Clearly there's no fixed price because it varies from customer to customer, but I think something above £12-£15, at those kind of prices certainly load turn-down becomes more of interest for the larger sites, it starts to become more viable for smaller assets. It's not a yes/no answer, but £6 is too low really to be of interest for a lot of these markets, and too low to be viable, it's quite challenging; unless it's something on top of a lot of other revenue streams for assets. (Aggregator, Phase 4)*

Interview data, combined with exit price data from the second TA, indicated that most participants in this auction were looking for prices between £10-£20/kW. A few were seeking prices in excess of £20/kW or £30/kW. This suggests that turn-down DSR will only be self-sustaining in the CM if future clearing prices rise above the levels seen in the 2018 auctions, closer to the levels seen in the first three T-4 auctions (£18-22.50/kW).

Adjustment for duration of turn-down DSR offered

Phase 1 of the TA evaluation found that the uncertain length of turn-down was a barrier to participation in the CM[42]. However, this was reported less in recent phases of the evaluation, possibly because there was less expectation of extended turn-downs given the lack of system stress events to date. While the interview evidence in the Phase 3 and 4

[42] "Some participants and non-participants (particularly aggregators) reported in interview that uncertainty about the number and length of stress events adversely affected sign-up by direct participants and aggregator clients." (Phase 1 report, 2016)

research suggested that the majority of components in the TA could turn-down for several hours if needed, there was also evidence that there were components (participating via aggregators) that could only offer turn-down for a maximum of 30 minutes because of operational constraints.

We therefore anticipate that some aggregators and clients might be interested in the option of a slightly different product that involved time-limited turn-down.

Higher penalties

Our knowledge of international capacity markets suggests that the reliability of capacity offered by DSR, including turn-down DSR, would be improved by the imposition of higher penalties. According to economic theory, penalties should ideally be set at the value of lost load.[43] But interview evidence suggested that higher penalties could deter direct participants from participating in the future CM.

> *[Interviewer: Are there any other factors that would influence whether you entered into the Capacity Market in future?] Respondent: I guess you'd have to look at penalties for non-delivery. At present, there are limited penalties, shall we say. ... If the penalty was severe, then there are certain scenarios where under the current rules we can demonstrate that we had delivered and then we wouldn't want to expose ourselves to penalties where you're not in control of what you're doing.(Direct participant, Phase 4)*

The perceived risk of non-delivery by direct participants was strongly linked to risks around baseline calculations. While an organisation might be confident of being able to turn down in response to particular CMN, those with variable baselines were less confident about being able to demonstrate sufficient turn-down relative to their calculated baseline. For example, their baseline might be depressed because their plant had not been operating at full capacity in the run-up to the CMN, or because it had frequently turned down for Triad during this period. One case study respondent reported that they tended to have low demand on a Monday which could cause baseline problems if a stress event fell on a Monday.

> *So, if a Stress Event was a Monday and we look back at previous Mondays, then we're comparing effectively a low demand against other days of low demand and we probably would not see ourselves as having achieved the demand, even though we've effectively given the demand up by not running. (Case study, Phase 4)*

A well-informed client echoed these concerns and reported that they chose to participate via an aggregator rather than directly because of concern about the risk of non-delivery against their baseline (given their frequent turn-downs for Triad). Other aggregator clients seemed less aware of CM penalties per se, although most were aware of some potential loss of revenue if they failed to turn-down when requested by their aggregator.

---

[43] Estimates of the Value of Lost Load range from £3,000/MWh to nearly £17,000/MWh. (Elexon, 2018) https://www.elexon.co.uk/wp-content/uploads/2017/09/33_278_10_VoLL-Review-Process-Paper-v1.0.pdf

In Phase 3 interviews, aggregators reported more concern about failing tests (which could result in the loss of all TA revenues) than about failing to deliver for a given stress event (which would only result in the loss of a proportion of TA revenues). They assessed a stress event to be unlikely in the 2017/18 delivery year and chose to ‘overfill’ their CMUs to minimise the risk of under-delivery in DSR tests, SPDs and – potentially – stress events. As described in the Phase 3 report, those aggregators submitting ‘portfolio’ CMUs reduced the risk further by putting forward a mix of types of loads in their CMUs.

In summary, higher penalties would increase the reliability of turn-down capacity put forward in the CM but might act as a disincentive to participants, particularly direct participants with variable baselines (e.g. because of delivery for Triad).

Reduced notice period for CMN/stress event

The four-hour notice period for a CMN was an important enabling factor for some industrial providers of turn-down DSR. The notice period was valued by industrial plants offering manually controlled DSR, particularly those running batch production processes that could not be interrupted mid-batch without high financial loss or operational consequences. These industrial plants were able to optimise their production to prepare for load shedding or load shifting (e.g. by building up stocks or heating up/cooling down elements of their process).

Because we get it early then, that allows us to plan the production through the day, so it can maximise what we actually shed. (Case study, Phase 4)

In contrast, organisations offering turn-down using loads suitable for automatic controls (which were less prevalent in the second TA) were less concerned about the notice period.

We'll respond within the time limit that we're given. Yes, if we only had 30 minutes response time, then we would still respond. Yes, that would be fine. (Case study, Phase 4)

Longer agreement length

Throughout Phases 1-4 of the evaluation, some aggregators argued that DSR contracts in the CM should be longer than one year. Those that did not have access to other sources of revenues, to stack with the CM, voiced this view more strongly. They were dependent on a continuous multi-year stream of CM revenues to offer their clients. Interview evidence supported this, in that these aggregators offered their clients multi-year contracts, with some elements of flexibility about how much capacity they submitted in a given year and (for larger clients) what clearing price they would accept.

Multi-year contracts between aggregators and aggregator clients were preferred by many aggregators as they provided clients with a pipeline of future opportunities and allowed both sides to recoup recruitment and set-up costs. However, interview evidence indicated that those aggregators that were less dependent on CM revenues, because they had other streams of flexibility revenue, signed one- or two-year contracts with an option to renew.

There was an inter-relationship between the length of contract won at auction and the price sought. Those aggregators who felt that turn-down DSR required ongoing support suggested that they would prefer a multi-year opportunity, even at a reduced price.

> *I think we still need to have some ring-fenced options to encourage more DSR. I'd be happy if the price turned out higher than the regular T-1 to reflect the effort required for people to participate. …I think that we could do with something that runs [for] the next, you know, five years to really get DSR turn-down off the ground.* *(Aggregator, Phase 4)*

However, we understand that, in designing the TA, BEIS did not see long-term contracts as appropriate unless projects faced large upfront capital costs, significantly above the levels observed for the second TA.

Earlier delivery assurance/testing

There was little evidence from the evaluation about the possible impact of earlier delivery assurance being required by BEIS. The timescale for proving up capacity in one-year ahead auctions by the start of the delivery year was already tight, particularly for those aggregators that recruited some or all of their clients after the auction. There were mixed views about early delivery assurance for four-year ahead auctions. One small aggregator commented that early 'proving up' of unproven CMUs in the main CM was advantageous to them as it allowed them to recycle credit cover and use it to reinvest in new, unproven capacity. Conversely, interview evidence from non-TA aggregators in Phase 3 indicated that some new entrants valued the four-year lead-time as it allowed them to develop their strategy and capacity for aggregation of DSR.

Proposed changes to CM rules

Ofgem are currently changing the CM rules to allow more flexibility to change and retest an individual component within a proven CMU, under certain circumstances, without being required to retest the whole CMU. Interview evidence indicated that the current lack of flexibility leads to aggregators submitting multi-component CMUs as unproven DSR, even if many of the components have previously participated in a proven CMU. This is supported by the low prevalence of proven DSR in CM auctions to date: Phase 3 and 4 interview evidence indicates that the only proven CMUs put forward to date have been those put forward by direct participants or aggregators putting forward single-client CMUs.

Other potential changes to CM rules

Interview evidence suggested that some other potential rule changes could reduce barriers to DSR participation in the CM, particularly for turn-down DSR, would be:

- To streamline metering requirements for DSR in the CM, particularly for smaller sites that cannot participate cost-effectively at present.
- To review baseline requirements for DSR in the CM (e.g. assess the accuracy of baseline predictions for different types of DSR loads in different seasons, using variants of the baseline formula). TA participants suggested some changes (e.g.

shortening the six-week baseline period; and changing the 'pre-CM Warning' adjustment for those providing other services to National Grid). But it was difficult to tell whether these changes would reduce the barriers for DSR participants generally, beyond those adversely affected by the current rules. Our comparison of the current methodology with the recommendations of a study[44] of baseline methodologies suggested that the current methodology was reasonable in its approach to allowing for variable loads.

- To reduce credit cover requirements for DSR, which would encourage new entry and innovation, particularly for smaller participants with more limited access to internal finance.

Evidence about the effect of these aspects of CM rules is presented elsewhere in this report and in reports of earlier Phases. There may be interactions between the different changes discussed above. For example, lower annual prices in the main CM might deter fewer potential DSR participants if agreements were longer or if baseline requirements were less demanding. Similarly, rigorous metering accuracy is only meaningful if baseline measurements are themselves meaningful.

Will turn-down DSR evolve using new technologies in the CM, with no further Government intervention and support?

This report has discussed a range of factors that may stimulate future growth in turn-down DSR, including MCPD restrictions for diesel generation, new control technologies making smaller sites cost-effective and new/cheaper battery technologies increasing the flexibility options available to industrial and commercial organisations. CM rules now allow HVAC assets to meet CM requirements while contracted for frequency services, which should further facilitate participation by HVAC and commercial DSR. Interview and auction evidence indicate that CM prices would need to rise to £10-20/kW to support recruitment of new turn-down DSR and investment in new controls for small sites: this is closer to price levels seen in the first three T-4 auctions, and above price levels seen in 2018 CM auctions. However, technological change may reduce investment costs over time. Interview evidence indicates that TA aggregators were looking for turn-down DSR opportunities from asset types that did not participate in the second TA scheme.

So I'm personally working with [specialist aggregator] who are focused on the HVAC demand turn down. They work with what's their availability in the day, keeping comfort levels in that building within a dead band. So that's quite an exciting one because - back in the day - it was all about back-up generation… We're certainly starting to open up different opportunities there and looking at electric vehicles and batteries… (Aggregator, Phase 4)

[44] EnerNOC (2011). *The Demand Response Baseline,* White Paper. Available at: https://library.cee1.org/sites/default/files/library/10774/CEE_EvalDRBaseline_2011.pdf

7. Conclusions

Based on our assessment of the evidence set out in this report, we have summarised our evaluation findings for the second TA against the high-level evaluation questions posed by BEIS, as follows:

HLQ 1 - What outcomes can be attributed to the second TA and were they as intended by BEIS? What outcomes occurred for whom and under what circumstances?

The second TA procured 293 MW of turn-down DSR, compared to 60-90 MW of turn-down DSR in the first TA. This was put forward by six aggregators and three direct participants and comprised 333 individual sites, which were almost entirely industrial. There was very little participation by commercial sites in the second TA.

HLQ 2 - Through what levers and causal mechanisms has the second TA contributed to these outcomes and the variation by group and circumstance?

The high price, low credit cover and small volume threshold in the second TA scheme encouraged both new and existing aggregators to prioritise turn-down DSR and recruit new clients that were previously turning-down for Triad only and were not previously contracted for flexibility services. This provided a safe environment for learning about both turn-down DSR and CM participation amongst those with less experience of one or other of these processes. However, there was less additionality for experienced players (for whom the high price in the second TA was mainly a windfall) and for non-participant aggregators with sites providing mixed back-up and turn-down DSR capacity that could not easily be submitted to the second TA. All participants in the second TA went on to participate in the main CM auctions in 2018, although only large, single-site CMUs and turn-down capacity with access to revenue from other flexibility services cleared at the lower prices in these auctions. The extent to which the second TA capacity will obtain agreements in future CM auctions will depend on future clearing prices.

HLQ 3 - Did the second TA represent good value for money to both scheme participants and the consumer?

Although the second TA was expensive by comparison with recent CM auctions in GB, it had different auction objectives. Second TA prices were broadly comparable to turn-down DSR in international markets, but it is difficult to draw direct comparisons with other flexibility services in the UK as they have different requirements.

HLQ 4 - Which aspects of the second TA's design and implementation account for the findings of HLQ 2 and 3?

Participants found the second TA onerous compared to other flexibility services, particularly in terms of metering tests (which most participants avoided by selecting site with supplier settlement metering) and setting up and maintaining meter data flows for DSR tests and SPDs. These issues apply equally to the main CM.

HLQ 5 - What are the implications of the findings for the future contribution of turn-down DSR to the CM?

Our research suggests that turn-down DSR will only be viable at scale in the main CM if future CM prices exceed £10-20/kW. Currently, prices around these levels are required to support recruitment of new turn-down DSR and investment in new controls and metering for smaller sites. The future viability of turn-down DSR in the CM will also depend on participants' continued ability to stack CM revenues with revenues from Triad cost-avoidance and frequency services.

8. Glossary and definitions

Term or acronym	Definition
Aggregator	An intermediary organisation that provides a service of collating capacity (from generation and/or DSR) for National Grid balancing services or the Capacity Market (CM), from a range of other organisations, in return for a share in the revenues generated.
Aggregator client	An organisation that contracts via an aggregator to access National Grid balancing services or the CM, rather than participating directly in these services.
Back-up generation	Generator (often diesel-powered) designed to be used if there is a power cut or problem with mains power. Usually located onsite 'behind the meter'.
Balancing services	System services contracted by National Grid. Those mentioned in the TA evaluation reports comprise: • 'Reserve services' that provide reserve capacity to balance electricity supply and demand (through generation or demand response). Examples include STOR and DSBR (see below). • 'Frequency-related services' that provide very short-term changes in electricity demand or supply to help maintain the frequency of the grid for Great Britain (GB) at 50Hz. Examples include FFR and FCDM (see below).
Baseload generation	Electricity-generating equipment normally operated to serve loads on an around-the-clock basis.
Bayesian updating	The evaluation used an analysis tool called 'contribution tracing with Bayesian updating'. Bayesian updating refers to the specification of 'prior' probabilities for each hypothesis, and to the updating of these to 'posterior' probabilities, based on certain evidence tests. See Appendix 6 of Phase 2 report for more detail.
Capacity	The CM was established for the purpose of ensuring adequate capacity to meet the demands of consumers for the supply of electricity in Great Britain. Capacity can be in the form of electricity generation plant or reduction in demand for electricity.

Term or acronym	Definition
Capacity Agreement	A capacity agreement comprises the rights and obligations accruing to a capacity provider under or by virtue of the CM Regulations and the Rules in relation to a particular capacity committed CMU and one or more delivery years.
Capacity Market (CM)	A series of auctions administered by National Grid, through which it procures future electricity capacity. The main auctions, known as 'T-4', are held annually 4 years ahead of the delivery year. Adjustments are made through annual 'T-1' auctions, one year ahead of the delivery year. The Transitional Arrangements involve two additional auctions that are designed to encourage growth in specific categories of capacity, to enable them to participate in the main CM in future.
Capex	Capital expenditure.
CHP	Combined heat and power (a plant that produces heat as well as electricity).
CMN	Capacity Market Notice. The automatic warning issued by National Grid, warning that a stress event may occur in four hours' time. The criteria for issuing a CMN are automatic and depend on the predicted balance between electricity supply and demand four hours ahead.
CMOs	Context-Mechanism-Outcome combinations. These are realist hypotheses about how the policy is expected to work, which are tested during realist evaluation. See separate definition of 'realist evaluation'.
CMU	Capacity Market Unit is a unit of electricity generation capacity or electricity demand reduction that participates in GB's CM. To pre-qualify for the first TA, a CMU had to be between 2 MW and 50 MW. For the second TA, the minimum threshold was 500 kW. A CMU may consist of a number of sites or components.
Component	A single site within a Capacity Market Unit (CMU). Some CMUs have only one component, while others have 20 or more. There is no lower limit on the capacity offered by a component, but there is an upper limit in that the sum of capacity offered by a TA CMU's components cannot exceed 50 MW.
Context	The circumstances which affect whether a policy 'works' and for whom. Consideration of 'context' forms an important part of realist approaches to evaluation. In realist terms, contexts are circumstances which affect whether and which mechanisms 'fire', and therefore whether and how the policy works.
Contribution analysis	Contribution analysis involves a structured process to develop and test a 'contribution story' (i.e. a coherent narrative that explains how a policy

Term or acronym	Definition
	intervention appears to be influencing change and assesses the likelihood that the intervention is contributing to observed results).
Contribution tracing	Contribution tracing involves the formulation and testing of competing hypotheses which could explain observed outcomes. The method involves explicit assumptions about the weight attached to different types of evidence. It aims to increase the transparency and replicability of qualitative analysis. See Appendix 6 of Phase 2 report for more detail.
Delivery year	The contractual year for delivery of CM obligations, which runs from 1st October of one calendar year through to 30th September of the following year.
Derated capacity	Volume of generation or demand reduction capacity after a reduction to account for outage rates, maintenance down time and so on, which varies by technology type. National Grid publishes lists of standard derating factors by technology.
Distributed generation	Generation units which are connected to the distribution network, rather than the transmission network.
Distribution network	The electrical network that delivers electricity to the bulk of consumers (excluding a small number of consumers that are connected directly to the transmission network).
Direct participant	An organisation that participates in National Grid balancing services or the CM directly, rather than via an aggregator.
DNO	Distribution Network Operator. DNOs own and operate the distribution network of towers and cables that bring electricity from the national transmission network (see the National Grid) to homes and businesses.
DSBR	Demand-side balancing reserve (interim balancing service for winter 2015/16). One-to-one agreements between organisations and National Grid in which the organisation was paid to reduce demand at certain times. National Grid announced in August 2016 that they would not tender for DSBR in winter 2016/17.
DSR (Demand-side response)	DSR means the activity of reducing the metered volume of imported electricity of one or more customers below an established baseline, by means other than a permanent reduction in electricity use. See also ‘Turn-down DSR’ below.

Term or acronym	Definition
	The TA evaluation reports focus on DSR by industrial and commercial rather than domestic consumers, as domestic DSR is much less well-developed in GB.
DSR Test	Test specified in CM rules, to demonstrate that a DSR CMU can reduce electricity usage by a given amount, relative to baseline demand. The test involves compilation of baseline data over a 6-week period, followed by collection of meter data by EMRS to confirm reduction below the baseline level at agreed times.
DTU	Demand Turn Up (the opposite of demand or load turn-down). Contracts with National Grid to make use of excess electricity generated by the distribution system (largely from solar power) when not otherwise needed.
DUoS	'Distribution Use of System' charges – these are charges for use of the electricity distribution network.
Dynamic FFR	See FFR.
Early Auction	An additional one-year ahead CM auction that was held in January 2017 and procured capacity for delivery in 2017/18. This auction cleared at £6.95/kW.
Electricity Settlements Company (ESC)	Government body set up to deal with paying capacity providers and recovering the costs from electricity suppliers.
Elexon	Organisation responsible for administering the GB electricity market Settlement and Balancing Code. Contracted by the ESC (see above).
Embedded benefits	Benefits negotiated between consumers and suppliers, when DSR or small-scale generation by electricity consumers helps suppliers to avoid network costs.
Embedded generation	Similar meaning to 'distributed generation'
Electricity Market Reform Settlements Limited (EMRS)	A wholly owned subsidiary of Elexon which the ESC contracts to settle CM payments and Contracts for Difference, and to collect and store metered data. Sometimes referred to as 'the settlements body'.
Enhanced Frequency Response (EFR)	A faster frequency response product tendered by National Grid in 2016, which requires organisations to interrupt their electricity supply within less than a second. Some of the service providers offer battery storage.

Term or acronym	Definition
Fast Reserve	A service tendered monthly by National Grid that procures large blocks of reserve capacity (exceeding 50MW) that can respond within two minutes. Pump storage is currently the main provider of Fast Reserve.
FCDM	Frequency Control by Demand Management (similar to FFR). A bilateral agreement between an organisation and the National Grid, that requires the organisation to interrupt its supply for 30 minutes, at two seconds' notice.
Firm Frequency Response (FFR)	Firm Frequency Response. A monthly tendered service through which National Grid procures a very short period of generation or demand reduction, at 30 seconds' notice, to support the 50Hz frequency at which the system operates. National Grid procures two kinds of FFR: • Static (or non-dynamic) FFR, which is a discrete service, involving responses of a few seconds triggered by a defined frequency deviation. • Dynamic FFR, which is a continuously provided service used to manage normal second-by-second changes on the system.
Flexibility	Ofgem defines flexibility as 'modifying generation and/or consumption patterns in reaction to an external signal (such as a change in price) to provide a service within the energy system.'[45]
Flexible capacity	Electrical capacity (generation or load) that can offer flexibility to the electrical grid (see 'flexibility').
Frequency-related services	Services procured by National Grid to support the 50Hz frequency at which the system operates. These involve short-term changes to generation or demand at short notice, and usually require an automated response.
GB	Great Britain (the area covered by the electricity grid in England, Scotland and Wales).
Generative causation	Generative causation assessment methods involve a forensic examination of causality using strategic data collection and logic, rather than a probabilistic assessment of causation through statistical correlation. The basic explanatory structure in realism is that a mechanism (M) acting in context (C) will generate outcome (O). These causal propositions (CMOs) are the starting point and end product of investigation in realist evaluation. See Appendix 5 of the Phase 2 report for more information.

[45] Ofgem (2015), 'Making the electricity system more flexible and delivering the benefits for consumers.' *Ofgem Position paper*. Available at: https://www.ofgem.gov.uk/ofgem-publications/96959/flexibilitypositionpaperfinal-pdf Accessed 13 September 2016.

Term or acronym	Definition
Hassle costs	'Hassle costs' are the cost directly associated with TA participation. This could include marketing effort by aggregators, the cost of time spent on the TA application, auction and testing processes, and the cost of new metering or controls specifically required for the TA.
HLQ	High Level Evaluation Question – one of the main questions that BEIS has asked this evaluation to research.
HVAC	Heating, ventilation and air conditioning equipment, usually within a building
Long-term STOR	Longer term version of STOR (see below), which was contracted by National Grid on a once-off basis but is now closed to new entrants. Holders of Long-term STOR contracts must declare that they will surrender these contracts if they obtain a capacity agreement for the same capacity.
Mechanism	A change in people's reasoning, brought about through the resources provided by a policy, which leads to a policy outcome. Identification of causal 'mechanisms', which operate in particular 'contexts', forms an important part of realist approaches to evaluation.
Missing money	In our analysis of costs and revenues associated with electrical capacity put forward for the first TA, 'missing money' is defined to be the minimum revenue that a participant would require for their participation in the TA to break even. See Phase 1 report.
NAO	National Audit Office
Net CONE	The 'net Cost of New Entry' is one of the parameters used to define the demand curve in a CM auction. It is set to reflect the estimated cost of marginal plant at the target capacity entering the auction.
National Grid	The National Grid runs Great Britain's electric high-voltage transmission network, is System Operator for the electricity system, commissions balancing services for the supply of flexible capacity and administers the GB CM. (See http://www2.nationalgrid.com).
Outcome	A change in the state of the world, brought about as a result of a policy or other influences. Realist approaches to evaluation attempt to identify the 'contexts' and 'mechanisms' that lead to a particular 'outcome'.

Term or acronym	Definition
Proven DSR	A unit of DSR capacity that has passed the DSR test required to participate in the GB CM.
Realist contribution analysis	We have used the term ‘realist contribution analysis’ to describe contribution analysis that is undertaken in the context of a realist evaluation. Contribution analysis involves the specification of a theory of change, assessment of the evidence base, gathering of new evidence, theory testing and then refinement of the theory of change. While contribution analysis is often used to assess the ‘average’ contribution of an intervention, across a scheme as a whole, we have applied this method using a realist approach and have assessed the TA’s contribution on a case-by-case basis. See Appendix 6 of Phase 2 report for more detail.
Realist evaluation	A realist approach[46] to evaluation emphasises the importance of understanding not only whether a policy contributes to outcomes (which may be intended or unintended) but how, for whom and in what circumstances it contributes to these outcomes.
Realist hypotheses	Realist evaluation involves developing theories about programmes and policies. These theories involve the development of clear hypotheses about how, and for whom, programmes might ‘work’. The implementation of the programme, and the evaluation of it, then tests those hypotheses.
Realist synthesis	A realist synthesis is the synthesis of a wide range of evidence that seeks to identify underlying causal mechanisms and explore how they work under what conditions, answering the question "What works for whom under what circumstances?" rather than "What works?"
Red zone management	Consumers avoiding (or generators targeting) the times when distribution costs (or payments) are highest – i.e. the periods defined as ‘red’ or ‘super red’ in the peak demand traffic light system.
Reserve services	Contracts between National Grid and organisations that can provide capacity held in reserve, in the form of generation or DSR.
Satisfactory performance days (SPDs)	CM participants are obliged to provide evidence of three half-hour settlement periods during the winter of a delivery year, on different days, in which they met their full capacity obligation.

[46] Pawson and Tilley (1997) (op cit). Pawson (2006) (op cit).

Term or acronym	Definition
Small-scale generation	For the purposes of the first TA, generation units less than 50MW that are connected to the distribution grid.
SNAPs (System Needs and Product Strategy)	National Grid's consultation on the future of balancing services during 2017. National Grid has now published a Product Roadmap for Frequency Response and Reserve. Further details of National Grid's review are available at: https://www.nationalgrid.com/uk/electricity/balancing-services/future-balancing-services
Static FFR	See FFR
STOR	Short-Term Operating Reserve - a reserve service run by National Grid through which organisations bid to provide generation or DSR to National Grid during peak demand periods (STOR windows). STOR is procured via tenders three times a year. A response time of at least 20 minutes is required.
Stress event	Period in which the electricity supply/demand balance is too tight (as determined by the System Operator's algorithms). Organisations holding capacity agreements are committed to provide capacity during stress events, or face penalties as set out in the CM rules.
Supplementary capacity market auction (also known as the 'Early Auction')	See Early Auction above
T-1	A one-year ahead CM auction fine-tunes the procurement of capacity in the main (T-4) CM auction for a given year. The first T-1 auction was held in January 2018 and secured agreements for the 2018/19 delivery year at a clearing price of £6.00/kW.
T-4	The main CM auction, held annually 4 years ahead of the delivery year. At the time of writing, four T-4 auctions had been held: • the first T-4 auction was held in December 2014, procuring capacity to be delivered in 2018/19 (clearing price £19.40/kW) • the second T-4 auction was held in December 2015, procuring capacity to be delivered in 2019/20 (clearing price £18.00/kW) • the third T-4 auction was held in December 2016, procuring capacity to be delivered in 2020/21 (clearing price £22.50/kW) • the fourth T-4 auction was held in February 2018, procuring capacity to be delivered in 2021/22 (clearing price £8.40/kW)

Term or acronym	Definition
TA	Transitional Arrangements for DSR and small-scale distribution-connected generation – the TA involved two one-year ahead CM auctions in 2016 and 2017 that were designed to encourage growth in specific categories of capacity, to enable them to participate in the main CM in future. The first TA auction was held in January 2016, procuring capacity for the 2016/17 delivery year (clearing price £27.50/kW) The second TA auction was held in March 2017, procuring capacity for the 2017/18 deliver year (clearing price £45.00/kW)
Transmission network	The high voltage power lines linking power stations to the distribution network. Some major electricity consumers are connected to the transmission network.
Triad avoidance	Consumers trying to reduce their electricity demand during three peak demand periods (or 'Triads'), in order to reduce their transmission charges. Transmission charges are based on demand during Triad periods. The Triad half hours are calculated from metered data (i.e. they are not known in advance) so Triad avoidance requires prediction of when the Triad periods might occur.
Triad targeting	Distributed generators trying to earn revenue by targeting generation at the Triad periods – the transmission charging methodology rewards them for doing so.
Turn-down DSR	Temporary reduction in electricity demand to avoid peak demand periods or to respond to National Grid instructions (sometimes called load reduction or curtailment). May also involve shifting electrical demand away from the peak demand period (sometimes called load shifting). This report focuses on DSR by industrial and commercial customers, as domestic DSR is much less well developed in GB.
Unproven DSR	A unit of DSR capacity that has not yet passed a DSR test, as specified by CM rules.

This publication available from www.gov.uk/beis

Contact us if you have any enquiries about this publication, including requests for alternative formats, at:

Department for Business, Energy and Industrial Strategy
1 Victoria Street
London SW1H 0ET
Tel: 020 7215 5000

Email: enquiries@bis.gsi.gov.uk