Department for
Business, Energy
& Industrial Strategy

EVALUATION OF THE TRANSITIONAL ARRANGEMENTS FOR DEMAND SIDE RESPONSE

Phase 3 - Main Report

August 2018

Contents

1. Introduction

This report presents findings from Phase 3 of the evaluation of the Transitional Arrangements (TA) for Demand Side Response (DSR). It focuses on the second TA auction in March 2017, which was restricted to load turn-down DSR only. The evaluation is realist and theory-based. Initial theory for the second TA auction was developed and then tested against findings from in-depth interviews with TA participants (both aggregators and direct participants) as will as findings from in-depth interviews with aggregator clients and with organisations that might have been expected to participate but did not. These findings have formed the basis of revised theory for the second TA.

This report presents findings from the evaluation of the second auction of the Transitional Arrangements (TA) for Demand Side Response (DSR). This realist, theory-based evaluation was undertaken for the Department for Business, Energy and Industrial Strategy (BEIS) by CAG Consultants, in partnership with Databuild, 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) and the Phase 2 report[2] (findings about delivery of obligations for the first TA scheme). This Phase 3 report covers findings about the second TA auction, while the future Phase 4 report will cover delivery of obligations for the second TA scheme.

Research and policy background

The TA is a pilot and forms part of the Capacity Market (CM) for security of electricity supply. The TA aims 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.

The TA aims to encourage the development of DSR that is increasingly needed to balance supply and demand in a decarbonised electricity grid[3]. 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 use. Under this definition, DSR may be achieved through any combination of

[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]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

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 involves two auctions for specific types of capacity within the CM, the first for delivery of capacity in the 2016/17 delivery year[4], 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 between 2 MW and 50 MW, the second TA scheme is only open to turn-down DSR and has a minimum threshold of 500 kW.

The TA auctions are 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 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 had not occurred at the time of Phase 3 research and will be researched, where feasible, in Phase 4.

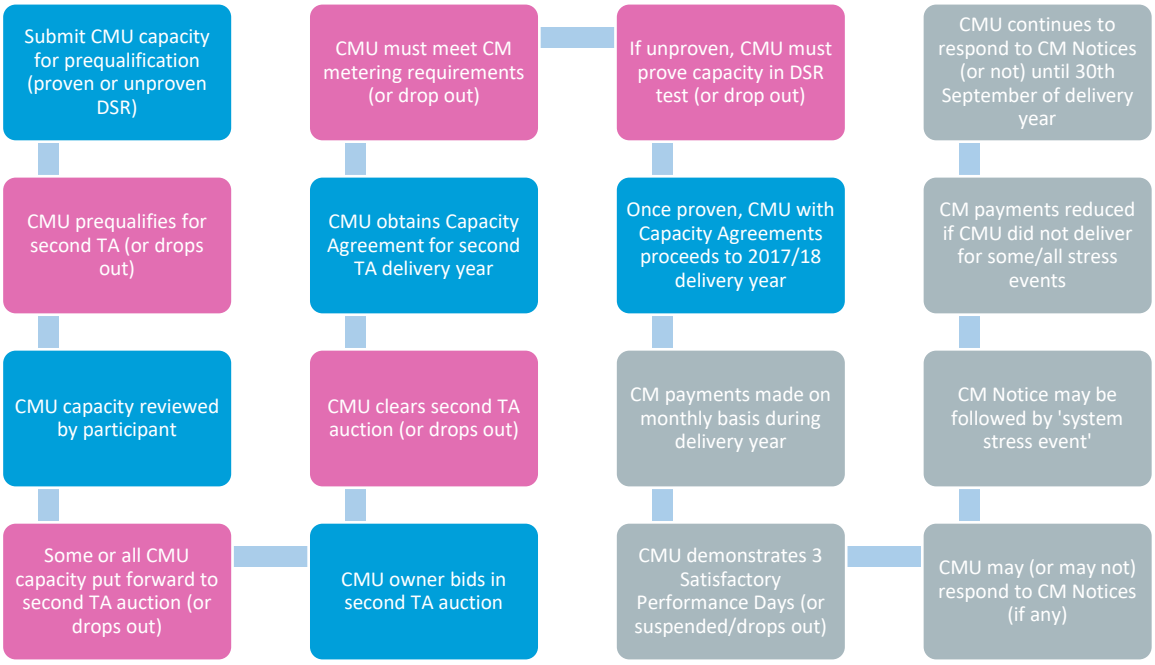


Figure 1.1: Main steps in process for second TA

Further details about the operation of the TA are given in Appendix 1, while a glossary of technical terms is provided in chapter 6 of this report.

[4] The delivery year runs from 1st October of one year through to 30th September of the following year.

Evaluation aims and objectives

The second TA has 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 do not include a significant contribution to security of supply in the delivery year (2017/18), because short-term system tightness has already been addressed through the introduction of the Early Auction alongside the TA. The objectives of the second TA scheme are therefore:

1. To develop a stock of flexible capacity[5] that can 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 is designed to answer five high-level questions (HLQs) posed by BEIS, in which the desired outcomes stem from the two objectives above. The evaluation will also look for any unanticipated outcomes of the scheme. This report presents findings that are relevant to HLQs 1, 2 and 4 for the second TA. It does not address HLQs 3 and 5, which will be covered in the Phase 4 report.

- HLQ 1 - What outcomes can be attributed to the 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 TA contributed to these outcomes and the variation by group and circumstance?
- HLQ 3 - Did the TA represent good value for money to both scheme participants and the consumer?
- HLQ 4 - Which aspects of the 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 DSR and small-scale generation to the CM?

[5] 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 is realist and theory-based. A realist approach[6] 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 futher in the Phase 2 report, chapter 1.

The development of a 'theory' of the TA is central to implementing a realist evaluation as it allows evaluators to rigorously examine the design and execution of the scheme, and test policy assumptions against available evidence. We developed an initial theoretical framework for Phase 3 of the evaluation, as presented in Appendix 4, which sets 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 is expected to lead to particular reasoning and choices being made (i.e. causal 'mechanisms' being activated[7]), leading to desired or undesired policy outcomes. These realist hypotheses are generally known as context-mechanism-outcome configurations or 'CMOs'[8].

Realist evaluation uses the idea of generative causality (i.e. a mechanism or reasoning only fires when the contexts are right). In Phase 3, we used realist analysis to test CMO hypotheses about why organisations participated in the second TA, and why some dropped out during the process. Our approach to analysis is explained further in Appendix 3. During Phase 4, we will undertake more extensive analysis, using generative causation methods, to assess the extent to which the TA contributed to observed outcomes, compared to other external influences.

Methodology

The evidence that we have gathered during Phase 3, and against which the initial theoretical framework has been tested, is set out in Table 1.2. The revised theoretical framework is presented in Appendix 6. Table 1.2 also lists the evidence against which this revised framework will be tested further during Phase 4.

The research summarised in Table 1.2 involved research with direct participants that put their own capacity into the TA, and research with aggregators[9] who put forward capacity on behalf of other organisations.

Phase 3 research involved in-depth telephone interviews with representatives of 35 organisations from October to December 2017, including representatives of nearly all of

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

the second TA participants, a sample of aggregator clients and a small sample of non-participants, including aggregators and organisations that could potentially have participated as clients or direct participants. An email survey was also sent to the ten TA participants and the ten aggregator clients that were interviewed. Nine full or partial responses to this email survey were received. For TA participants, as well as some of the non-participant aggregators, the Phase 3 evidence extended the information already gathered in Phases 1 and 2 of the evaluation. Further detail on sampling is provided in Appendix 3.

Table 1.1: Summary of evidence gathered in Phase 3 and planned for Phase 4 of the evaluation

Evidence source	Phase 3 research tasks	Planned Phase 4 research	Limitations
Scoping interviews and literature review	Telephone interviews with BEIS, Ofgem, National Grid and three other external experts on turn-down DSR, and updates from the literature review on turn-down DSR, to inform theory development.	-	-
Research with participants in second TA	In-depth telephone interviews with eight out of the 11[10] TA participants going forward with Capacity Agreements after the second TA auction (seven out of eight aggregators and one out of three direct participants). Six of the interviewed organisations provided additional quantitative information by email. We also undertook three further in-depth telephone interviews with aggregators that tried but failed to qualify for the second TA.	Further interviews in spring 2018 with 11 TA participants on delivery and additionality issues	There was some respondent fatigue and below-census response rates for TA participants. Nonetheless, response rates for TA participants and drop-out/non-participant aggregators were relatively high (50-100%).

[10] While only ten organisations cleared capacity in the auction, one further organisation bought some TA assets from a direct participant after the auction, bringing the number of participating organisations to 11.

Evidence source	Phase 3 research tasks	Planned Phase 4 research	Limitations
Research with aggregator clients in second TA	In-depth telephone interviews with 10 aggregator clients. These were identified through 15 screening interview responses from 45 known clients, drawn from CMUs going forward to delivery at the time of the research. Three of the 15 aggregator clients provided additional quantitative information by email. Clients were identified by matching meter point postcodes from National Grid with commercially-available company data.	Further interviews in spring 2018 with 10 additional aggregator clients, on delivery and additionality issues.	In Phase 3 research, aggregator client interviews were drawn from three of the eight aggregators participating in the second TA. However, the clients interviewed in earlier phases of this evaluation were drawn from a number of other aggregators that participated in the second TA. Across the Phases 1-3 of the evaluation, we have interviewed clients from six of the eight aggregators participating in the second TA, which gives us confidence that we are using evidence that covers more aggregator clients.
Research with non-participating aggregators	Six in-depth telephone interviews and two email responses from aggregators known to National Grid that did not participate in the second TA. These were purposively chosen to fill gaps in Phase 1 and 2 evidence.	-	-
Research with non-participating individual organisations (potential clients or direct participants)	Eight in-depth telephone interviews with non-participating organisations that might have been expected to participate in the second TA but did not. These were identified through a screening survey with 178 organisations that were	-	There was a limited sample and modest response rate of 22% for the screening survey with potential participants, largely because we lacked identified contacts in these organisations. The screening survey findings were used to characterise and

Evidence source	Phase 3 research tasks	Planned Phase 4 research	Limitations
	publicly known to offer turn-down DSR, had attended DSR information events or that matched known aggregator clients in terms of size and sector.		select non-participants for in-depth interviews. We expected response rates to be low and therefore did not use screening survey responses to generate quantitative findings. While we have confidence in findings on contexts and reasoning for certain cases, from in-depth interviews, the theory for non-participants has not been tested comprehensively across this large and diverse population.
Analysis of data from CM register and CM auctions	Analysis of latest CM register for the second TA and TA auction behaviour, including exit prices	Update Phase 3 analysis for final outcomes of appeals and testing. Review high-level outcomes for DSR in other CM auctions.	-
Analyse costs and revenues for turn-down DSR	Develop typology for turn-down DSR, based on interview and email survey data. Begin to assemble cost and revenue data.	Develop cost/revenue models for different types of turn-down DSR, from Phase 3 and 4 data.	Our characterisation of turn-down DSR in the second TA is currently incomplete. While we will seek further site data during Phase 4, our characterisation of capacity by business activity or asset type can 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

Evidence source	Phase 3 research tasks	Planned Phase 4 research	Limitations
			between different tests and turn-down events.
Analysis of stress event(s)	-	Analysis of stress event performance, with case studies of stress event responses and costs.	If there is no stress event, this analysis will focus on responses to CM Notices, DSR tests or SPDs.
Validation of findings	Internal workshop with BEIS and peer reviewers, plus external workshop with delivery bodies and industry representatives	Internal workshop with BEIS and peer reviewers.	-

The topic guides and email questionnaire for Phase 3 research were agreed in advance with BEIS. The topic guides were designed to gather the evidence required to test the theoretical framework in relation to participation and non-participation in the second TA, participation in the auction and auction bidding behaviour. They also included questions on the organisation's attitude to turn-down DSR, on aggregators' approaches to designing their CMUs, and on the reasons for any capacity dropping out. Those who would not be interviewed again in Phase 4 were also asked about their views about the influence of the TA on their future plans.

The email survey gathered information on metering and DSR testing experiences, 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 detail on the qualitative research and email survey methodology is provided in Appendix 3.

The Phase 3 interviews were recorded and transcribed. We then used spreadsheets to code the Phase 3 interview responses against contexts, mechanisms and outcomes (C-M-Os) in the Theoretical Framework, and to capture additional contexts, mechanisms and outcomes that were supported by the interview evidence but not yet captured by the theory. (See Appendix 3, Table A3.3, for an explanation of C-M-Os). Using evidence from Phase 3, supplemented by contextual information about the respondent organisations from earlier phases of research, from email survey responses, from auction data, from the CM Register and from published statements, we then analysed which outcomes were observed for each organisation, how and why individuals within those organisations made choices that led to these outcomes and what factors influenced their choices. While some of the influencing factors that we examined related to the design of the TA, others were external factors relating to the main CM, other flexibility services, the wider energy market or business contexts. We summarised this analysis in the form of 'tailored' C-M-Os for each case, where a case represented an organisation (i.e. a TA participant, non-participant or aggregator client). Where one organisation exhibited several types of reasoning (e.g. submitting different bids in the auction for different CMUs), we prepared developed separate 'tailored' C-M-Os to cover this. We then looked for patterns across these 'tailored' C-M-Os and analysed the extent of support for different C-M-Os in the initial framework (see Appendix 4). Finally, we refined and expanded the initial set of C-M-Os to create a revised theoretical framework (see Appendix 5), that simplified and generalised the findings from the 'tailored' case-by-case C-M-Os. This preserved anonymity for respondents while retaining the essence of our findings for individual cases.

As explained above, this analysis drew on data from other sources where relevant, in addition to the Phase 3 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 was 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.

In Phase 3, we did not apply 'contribution tracing with Bayesian updating' (a variant of 'process tracing'). This method was applied to analysis of additionality hypotheses during Phases 1 and 2 but was not used in Phase 3 because we were not testing additionality hypotheses during Phase 3. Process tracing will, however, be used as part of the Phase 4 analysis, to weigh the strength of evidence where there is conflicting evidence for particular cases.

The capacity provided through the second TA was characterised by matching meter point data from National Grid with commercially-available company databases, for CMUs that had passed testing at the time of Phase 3 research. This enabled us to identify the company and sector of second TA sites with a high level of confidence[11]. We also used a combination of data sources to identify the types of assets providing turn-down DSR on each site, and – where possible – the scale of capacity typically provided by each site. This assessment was based on email survey data, where available, supplemented by additional information from Phase 3 and earlier interviews where relevant. We aim to fill data gaps in Phase 4 research, as well as gathering further cost and revenue information for second TA CMUs.

[11] 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.

2. Participation in the second TA

The second TA auction, which was restricted to load turn-down DSR, attracted lower volumes of capacity (755 MW) than the first TA (1560 MW). This is because the latter was also open to back-up generation behind the meter and to small-scale, distribution-connected generation. The organisations attempting to prequalify capacity for the second TA were, with one exception, a subset of those that attempted to prequalify for the first TA.

Fourteen organisations attempted to qualify 47 Capacity Market Units (CMUs) for the second TA, putting forward 755 MW of turn-down DSR. The final contracted DSR capacity going forward to delivery in the second TA was 293 MW of turn-down DSR, across 28 CMUs. Although the total volume of capacity was lower than the first TA, the volume of turn-down DSR was considerably higher than the estimated 60-90 MW of turn-down DSR that went forward to delivery in the first TA.

Aggregators that participated in the second TA saw the scheme as an opportunity to build their business or build revenue for existing clients. Aggregators that chose not to participate regarded the TA as incompatible with their DSR portfolio or the timing of their business development.

Individual organisations that participated, whether directly or as aggregator clients, were able to offer turn-down DSR cost-effectively because they had sizeable sites suitable for turn-down DSR and did not face excessive upfront costs or excessive risks to their business. They tended to feel confident about delivering turn-down DSR because they were already doing 'self-despatch' of turn-down DSR for Triad management. Individual organisations that chose not to participate did so for technical reasons (e.g. their sites were unsuitable), because their sites were not cost-effective (e.g. too small to justify upfront costs), or because senior management were not persuaded of the business case for turn-down DSR and perceived it as too risky for their business.

This chapter presents our findings on the volumes of turn-down DSR capacity participating in the second TA, and the types of organisations putting forward this capacity. We define 'participation' as attempting to prequalify for the second TA. Organisations seeking to participate in the second TA had to submit prequalification information for their CMUs to National Grid in autumn 2016. To prequalify, these CMUs had to comprise turn-down DSR only within Great Britain, be between 500 kW and 50 MW, and could not have cleared in the recent T-4 or T-1 auctions (held in December 2016 and January 2017 respectively).

Summary of participation outcomes for the second TA

Fourteen organisations attempted to prequalify for the second TA, putting forward 47 CMUs (755 MW). In March 2018, contracted capacity was 293 MW, across 28 CMUs and nine[12] organisations. Winning capacity was purchased at the auction clearing price of £45/kW, so the total cost was just under £13 million. During the second TA process there were changes to the number of CMUs and volume of capacity originally put forward for prequalification, as shown in Figure 2.1 and Figure 2.2 respectively.

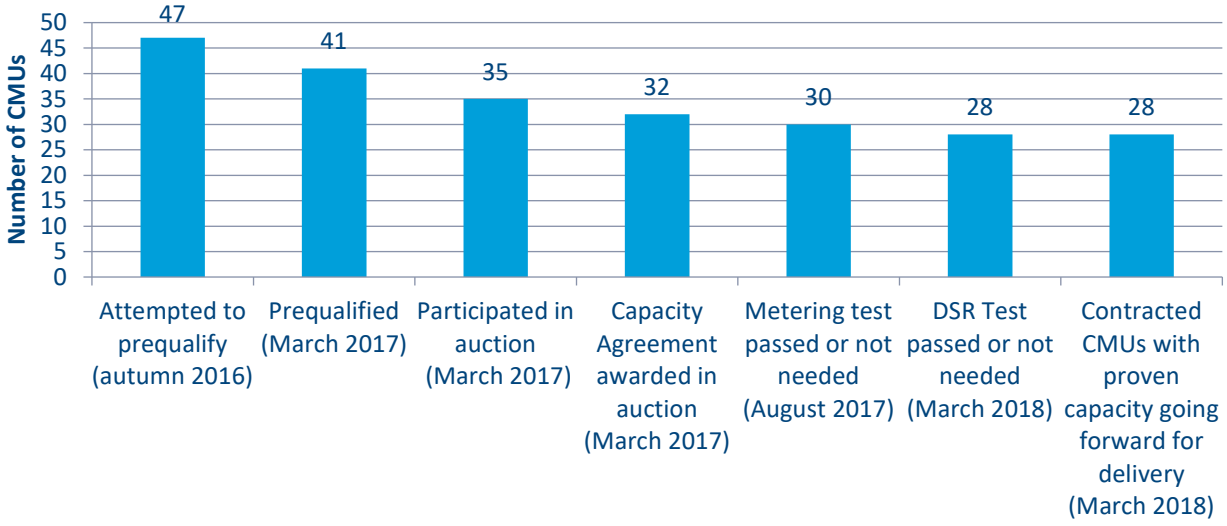


Figure 2.1: CMUs participating in the stages of the second TA (source: CM register)

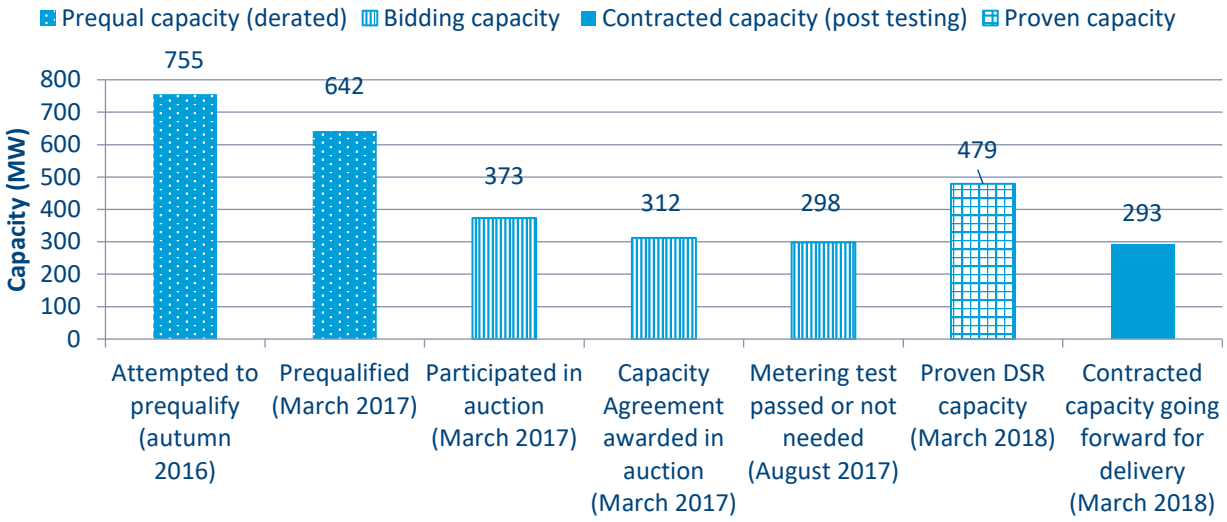


Figure 2.2: MW of capacity participating in the stages of the second TA (source: CM register)

[12] Six of the original fourteen organisations dropped out. But one organisation joined because it bought some TA assets from a participating organisation, making nine organisations going forward.

The changes are summarised here, while the reasons for these changes are explored in chapter 3:

- Six of the CMUs failed to prequalify, meaning 41 CMUs (642 MW) prequalified.
- Six further CMUs (61 MW) were withdrawn before auction, and 18 other CMUs were put into the auction at below their prequalified volume (reducing the volume by a further 208 MW), meaning 35 CMUs (373 MW) entered the auction.
- Three CMUs failed to clear the auction, meaning 32 CMUs (312 MW) were awarded Capacity Agreements.
- Two CMUs (14 MW) had their Capacity Agreements terminated for failing to pass metering testing requirements post-auctio.
- Two further CMUs (1 MW) were terminated for failing to obtain DSR certificates and four CMUs had proven DSR capacity slightly below their auction capacity (4 MW).
- The final outcome was that 28 CMUs (293 MW) were covered by Capacity Agreements and were available for delivery.
- The proven capacity of these 28 CMUs, as demonstrated by their DSR tests, was 479 MW (64% higher than their contracted capacity).

These volumes were lower than the first TA volumes, in which 1560 MW was submitted to prequalification, 1100 MW entered the auction, 803 MW was awarded Capacity Agreements and 620 MW was contracted for delivery after metering and DSR testing. The higher volumes in the first TA were primarily due to the first TA being open to small-scale generation and back-up DSR as well as turn-down DSR. As shown in Figure 2.3, just under 40% of the capacity submitted to prequalification was contracted for delivery after testing in both the first and second TA, but as explained in chapter 3, the reasons for drop-out differed between the two auctions.

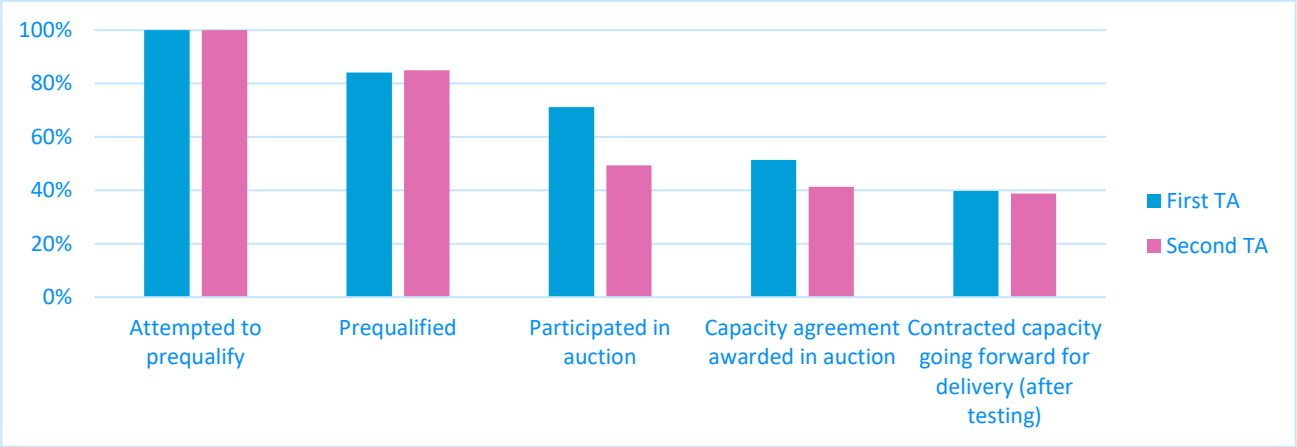


Figure 2.3: Percentage of capacity (MW) going forward to each step in TA process

The factors that influenced organisations to attempt to prequalify (i.e. participate) in the second TA are explained further below.

Who participated in the second TA?

Those attempting to prequalify in the second TA were, with one exception, a subset of the organisations attempting to prequalify unproven DSR for the first TA (see Appendix 2). This is consistent with our findings in Phase 1 of the evaluation that the TA scheme attracted organisations that already had strong links with National Grid and had the organisational capacity to deal with complex CM rules.

The smaller number of participants in the second TA is also consistent with the first TA finding that most aggregators reported that turn-down DSR was more difficult to recruit than small-scale generation and back-up generation behind the meter, so not all aggregators from the first TA saw the second TA as an attractive opportunity.

In the second TA, all 293 MW of the contracted capacity was turn-down DSR, as specified in the TA rules[13]. This compares to an estimated 60-90 MW of turn-down DSR contracted for delivery in the first TA (see Phase 2 report, chapter 6).

While all the DSR capacity in the first TA entered the auction as unproven DSR, two of the CMUs in the second TA (with bidding capacity of 37 MW) were 'proven' (i.e. passed DSR tests) in the first TA. Phase 3 evidence indicates that a total of four 'pure' turn-down DSR CMUs participated in the first TA: in addition to the two that went forward to the second TA, two other 'pure' turn-down DSR CMUs from the first TA were entered for the second TA but failed to prequalify because they cleared in the Early Auction. The remaining DSR

[13] In the first TA, 'DSR' included back-up generation behind the meter ('back-up DSR'), as well as load turn-down ('turn-down DSR'). In the second TA, no generation was allowed, either in front of or behind the meter, so DSR was solely 'turn-down DSR'.

CMUs from the first TA contained some back-up generation sites and therefore could not be carried forward to the second TA as proven DSR[14].

Ten organisations obtained Capacity Agreements in the second auction and one further organisation purchased TA assets after the auction. Of the 11 TA participants post-auction, eight were aggregators and three were direct participants. Post-auction terminations led to two aggregators dropping out of the second TA, so at the time of writing there were nine participants. Six of these were aggregators, contributing 82% of the contracted capacity going forward to delivery, as shown in Figure 2.4.

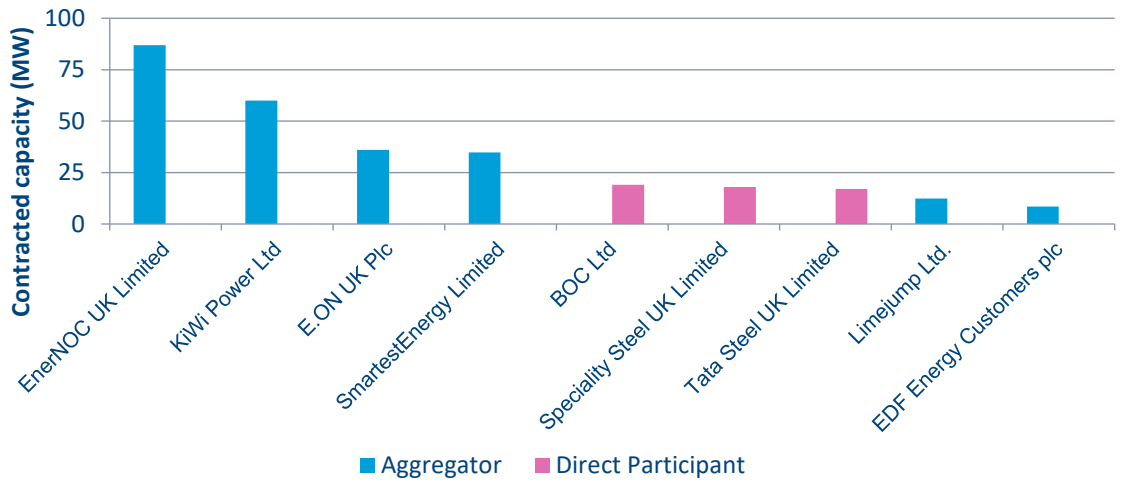
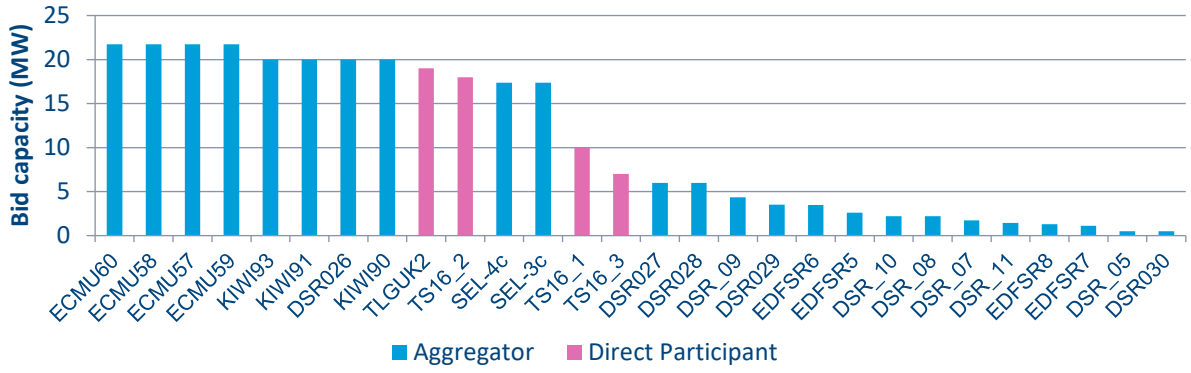


Figure 2.4: Contracted capacity in the second TA, by organisation (source: CM register)

There was considerable variation in the size of the 28 CMUs proceeding to delivery, from 500 kW to over 20 MW, as shown in Figure 2.5 below. The 24 aggregator CMUs averaged 10 MW per CMU, but with a wide variation. There was less variation in the size of the four direct participant CMUs, with an average of 14 MW per CMU.



[14] DSR tests in the second TA were passed at CMU rather than component level, so a CMU only remained 'proven' while all its components remained unchanged. National Grid plans to revise this rule in future CM auctions, to allow certain changes within proven CMUs.

Figure 2.5: Contracted capacity in the second TA, by CMU (source: CM register)

Overview of the aggregator market

During Phase 3 research, we asked aggregators about their attitudes to turn-down DSR. The interview evidence suggested that there were currently three main types of aggregators in the GB market:

- **'Supplier aggregators'** that were electricity suppliers who saw value in the flexibility market and increasingly offered flexibility services as well as supply services, potentially offering clients with high degrees of flexibility access to the wholesale market and balancing mechanism as well as flexibility services.
- '**Specialist aggregators'** that had a core business of aggregating DSR (both back-up and turn-down) for a range of flexibility services, often offering a software platform to support their activities but not possessing an electricity supply licence.
- '**Technology providers'** that previously worked as aggregators but increasingly specialised in offering software or technology platforms for control or automation of demand response (particularly for frequency response services such as dynamic Firm Frequency Response (FFR)), working in partnership with specialist aggregators or supplier aggregators.

This categorisation was not static, as some specialist aggregators had recently chosen to become 'technology providers' while others were choosing to become small-scale 'supplier aggregators'. As highlighted in the Phase 2 report, a number of specialist aggregators wanted access to an electricity supply licence in order to offer their clients access to the balancing mechanism. Some were seeking supply licences in their own right[15], while others had formed partnerships with, or had been acquired by a supplier.

While some of the 'technology providers' specialised in turn-down DSR for dynamic FFR, all of the 'specialist aggregators' and 'supplier aggregators' described their approach to DSR as technology-neutral. They sought to offer clients the best revenue opportunities for the clients' assets, whether these assets provided load turn-down, back-up generation or electricity storage. However, some of the 'specialist' and 'supplier' aggregators were more experienced at providing turn-down DSR and were more confident about the turn-down market and more likely to see the TA as an attractive opportunity for growth.

Figure 2.6 below therefore distinguishes between technology providers that specialise in automatic DSR, some of which focus specifically on turn-down DSR (Group A), and other types of aggregators (specialist and supplier aggregators) that take a more client-centred

[15] https://theenergyst.com/limejump-aggregators-will-need-an-electricity-supply-licence-to-survive/

and technology-neutral approach to flexibility. These aggregators may be confident about turn-down DSR (Group B) or cautious about turn-down DSR (Group C).

It is possible that differences in the technology platforms used by aggregators contributed to differences in their attitude to the turn-down DSR market. For example, there may be differences in the costs of metering or controls required to bring different types of assets into a particular aggregator's technology platform. While we did not find evidence on this point during Phase 3, we will explore this further during Phase 4.

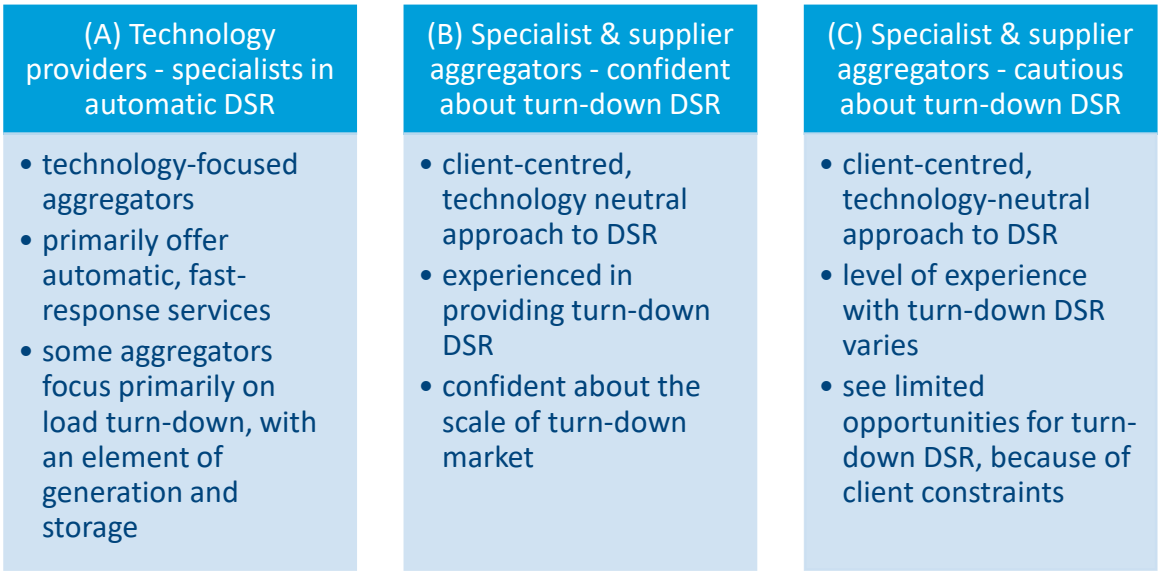


Figure 2.6: Characterisation of aggregator market

How and why did aggregators participate, or not?

The following sections outline the reasoning[16] of different aggregators about seeking to prequalify capacity for the second TA. To avoid disclosure, we have not specified whether particular aggregators were technology providers, specialist or supplier aggregators, but we have highlighted where aggregators were more or less confident about the prospects for turn-down DSR. A revised participation theory, based on these findings, is presented in Appendix 5.

Reasoning: 'We see the TA as an attractive opportunity for growth'

This reasoning was common for aggregators that sought to prequalify capacity for the second TA, who viewed DSR aggregation and/or the CM as strategic opportunities. They saw the TA as an attractive and low-risk route to recruit new turn-down clients, learn about

[16] We asked about participants' reasoning during interviews, so that we could test the causal 'mechanism' in our realist 'CMO' hypotheses.

the CM and build their DSR business, as well as give existing clients more revenue. Common factors for these aggregators were that:

- They had flexibility market experience and organisational capacity to participate in the second TA.
- They wanted to grow their DSR business (including turn-down DSR).
- Other changes made turn-down DSR more attractive (e.g. the Medium Combustion Plant Directive[17] posed a risk for back-up generation; while Ofgem's review of embedded benefits reduced the attraction of distribution-connected generation).
- They expected the TA price to be attractive and saw the TA/CM as relatively low-risk for clients, because of low credit cover (TA only) and infrequent stress events with long notice periods given for turn-down (for the TA and CM).

The Capacity Market is, to be frank, I suppose the easiest market to get into. Especially from a customer's point of view where you've got quite a long timeframe to respond. The amount of responses that you're expected to receive there is fairly low, so from a customer's point of view and from our point of view, it's kind of the easiest DSR market to get into. (Aggregator participant)

There were two main ways in which these aggregators sought to grow their DSR business through the TA:

- **'Business-building'.** These aggregators saw the TA as a low-risk opportunity to build their flexibility business, through acquiring new clients and offering new services to existing flexibility clients. Some mentioned that the TA provided an opportunity to recruit smaller customers, because of the low credit cover. These aggregators had a strategic goal of building their portfolios and were relatively confident about recruiting turn-down DSR.

Like I say, it's a plank of our strategy that we're following to build up demand-side response, so we want turn-down assets – not just for capacity market but other balancing services as well. (Aggregator participant)

- **'Knowledge-building'.** Some of these aggregators were also attracted to the second TA by the opportunity to build experience and learn about the CM. These aggregators had more limited experience of DSR in the GB CM.

[We were attracted to the second TA by] the opportunity to get some capacity, to learn about how the Capacity Market works. ..[it provided a] nascent market for gaining the experience and gaining the contracts in a [..] controlled

[17] http://ec.europa.eu/environment/industry/stationary/mcp.htm

environment that will hopefully enable us to be [..] successful competitors in the market going forward. (Aggregator participant)

Reasoning: 'We see the TA as offering attractive revenue or options for existing clients'

A few of the aggregators that attempted to prequalify capacity for the second TA took a slightly more opportunistic approach. They saw the TA as an attractive and low-risk route to give existing clients more revenue or more options for their flexibility assets. Common factors for these aggregators were similar to those above, but they had a slightly more negative perception of turn-down DSR, relative to generation or back-up DSR (e.g. because of higher upfront costs for smaller loads)

To be specific about turn-down, turn-down is more costly and more expensive to do, because you have smaller loads that need to be brought together. Therefore you have a higher upfront cost to be setting the assets up, relative to having generation of some form or other. (Aggregator participant)

The reasoning of these aggregators in relation to the TA was:

- **'Revenue-building'.** Where an aggregator had existing clients providing turn-down DSR for other services that were compatible with the TA, they saw the TA as a low-cost, low-risk opportunity to grow revenue from their existing flexibility capacity.

We're looking to make sure there's an economic opportunity… the value of the TA was to say, "Look, there's going to be a revenue stream available next year, or the year after, depending on the exact timings." (Aggregator, participant)

- **'Keeping our options open'**. Where an aggregator had existing clients providing turn-down DSR for other services, but were uncertain about their or their clients' CM strategy, they prequalified for the second TA on a precautionary basis, to allow existing flexibility clients the choice of participating in the TA, the Early Auction or other flexibility services.

Reasoning: 'We don't see turn-down DSR in the CM as an attractive strategic opportunity'

We observed this reasoning for aggregators[18] that did not seek to prequalify for the second TA and that did not see it as a good fit with their current flexibility portfolio. Aggregators expressed three main reasons for this, with the latter two groups being more negative about turn-down DSR:

- **'FFR is better for us'.** We observed this reasoning for technology-focused aggregators specialising in dynamic FFR. Their technology focus meant that participation in the second TA was not strategically attractive to them. One

[18] Owing to the limited sample for non-participant aggregators, we cannot be specific about the numbers of non-participant aggregators exhibiting each type of reasoning.

reason given was that they could not demonstrate measurable turn-down relative to a baseline for a half-hour DSR test, because FFR involved multiple responses over much shorter time periods (e.g. split-second). Another reason given was that competition[19] in the aggregation market led them to specialise in technology provision rather than aggregation.

- **'The main CM is better for us'.** We observed this reasoning for aggregators that preferred continuous access to future T-4 revenue streams and that were concerned that participation in the second TA would exclude them from participating in the third T-4 (for delivery in 2020/21). They chose to participate in the Early Auction in preference to the second TA, as this allowed them to put the same capacity into the third T-4. Aggregators with a range of capacity types, and with concerns about the ability of turn-down to cover long stress events, also preferred to offer mixed DSR CMUs in the main CM, comprising both back-up generation and turn-down DSR.

So, that was one part. ... the fact that [the second TA] doesn't offer you the ability to get involved with all those assets, as it is focusing on turndown DSR. (Aggregator, non-participant)

- **'The TA opportunity is not worth it for other reasons'.** Other reasons that non-participant aggregators gave for not seeing the TA as an attractive opportunity were:

  - They had clients with variable baselines that were concerned about the risk of non-delivery arising from the baseline methodology for turn-down DSR in the CM.
  - Those who were less experienced in the CM, were uncertain about TA prices and thought that the lead time from auction to delivery was too short to allow them to sign-up turn-down clients for the TA.

Reasoning: 'We are not ready to participate in the second TA'

We observed this reasoning for non-participating aggregators that were positive about turn-down DSR and would have liked to prequalify capacity for the second TA but were not in a position to do so at the time, because their flexibility business did not exist or was not ready to participate.

The TA was just one of those things where it was a good mechanism and just wrong time for us.... if we'd been in a place where we could've participated, we would've done. (Aggregator, non-participant)

[19] Evidence for competition was also provided by aggregator clients reporting that they were approached by several aggregators.

Overview of individual organisation's journey to turn-down

We asked participant and non-participant organisations whether and how they had considered turning down demand temporarily to reduce electricity costs. We analysed their experiences and developed a typical 'journey to turn down', as shown in Figure 2.7 below. The pink boxes represent external stimuli on this journey. This journey is relevant to organisations with turn-down loads suitable for turn-down in the CM, of sufficient scale to interest aggregators. Only parts of the journey are relevant to those with turn-down loads large enough to participate directly in the TA or CM (500 kW or 2 MW respectively), as they would not be dependent on being approached by aggregators.

For organisations that we interviewed that had loads suitable for turn-down in the CM, large enough to attract aggregator interest, a common finding was that exposure to Triad and distribution system (DUoS[20]) charges was usually the first step on the road to turn-down. Triad charges applied to sites with half-hourly meters (i.e. business sites with a peak load of 100 kW or more, or with new meters in profile class 05-08[21]).

Turning down loads to reduce these charges was an internal management decision that could be made without any external contractual commitment. Having taken that step, these organisations were more confident in their capability to turn down under contract for flexibility services, and were more likely to respond positively to an aggregator's approach about the TA. Our findings on the importance of self-dispatched Triad activity in stimulating and supporting turn-down DSR were consistent with findings in the Power Responsive Annual Report 2017, which estimated that 2 GW of flexible capacity was delivered through 'self-dispatch' in winter 2016/17[22].

Aggregators reported that they were generally interested in sites that offered significant potential for turn-down (e.g. more than 100 kW): not all sites with half-hourly meters that

[20] 'Distribution Use of System' charges (DUoS) – charges paid for use of the electricity distribution network.
[21] For details on meter profile classes, see: https://www.elexon.co.uk/knowledgebase/profile-classes/
[22] See p23 of the Power Responsive Annual Report, published February 2018. http://powerresponsive.com/wp-content/uploads/2018/02/Power-Responsive-Annual-Report-2017.pdf

were dispatching for Triad would have turn-down potential on this scale.

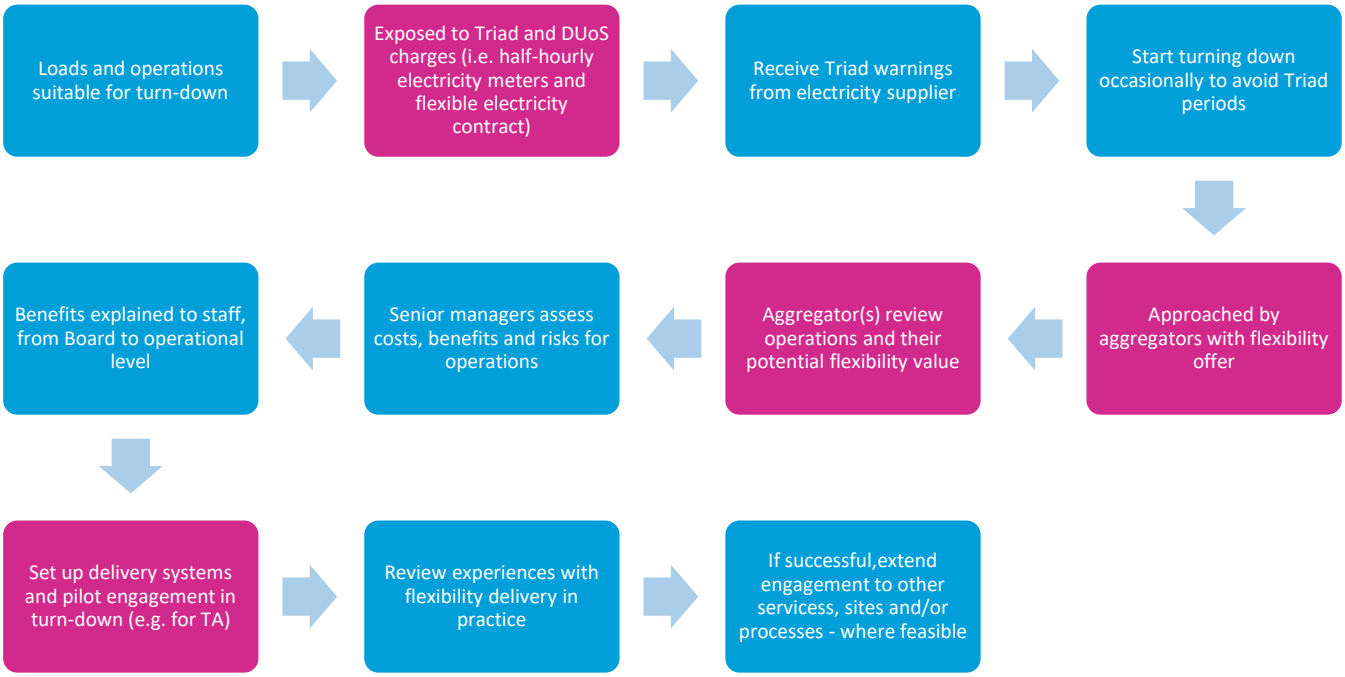


Figure 2.7: 'Journey to turn-down' for individual organisations with loads suitable for turn-down in the CM, at a scale sufficient to interest aggregators

How and why did individual organisations participate or not?

Our realist analysis used interview evidence and testing data to explore the reasoning of direct participants, aggregator clients and non-participating individual organisations (i.e. potential participants or clients) about turn-down DSR and their reasoning (if any) about TA participation. In this section, direct participants and aggregator clients are described jointly as 'participating organisations' to avoid disclosure, because of the small number of direct participants in the second TA.

We identified four main types of reasoning by organisations that put forward capacity to the second TA or might have been expected to do so (e.g. because they had already provided turn-down DSR for other services, had attended DSR events or were similar in size and sector to known aggregator clients).

Reasoning: `We can add TA revenue at little extra cost and low risk'

Individual organisations that we interviewed that exhibited this reasoning, and that went ahead to participate in the second TA, shared some common features:

- They had sites and assets (mainly industrial) that could provide turn-down DSR for stress events at acceptable levels of risk and costs, given their expectation (based on evidence from the first TA and advice from aggregators) of few stress events in 2017/18. This involved having some spare capacity, some surplus stock, some flexibility on the timing of production, or some subsidiary activities within their production process, so that they did not expect turning down these subsidiary activities or time-shifting production to affect their services to customers.

We are not sold out on product. If you are sold out on product, you have no DSR capability, at all. You just can't stop. (Participating organisation)

- They had experience of turn-down for Triad (and in some cases other flexibility services), so could easily add a few turn-downs for the second TA.

It was felt there was no real extra burden being put on the sites, with four tests a year and an indeterminate number of real calls. We were already responding to upwards of 15 Triad forecast periods a year, so it was felt that, yes, we could do that with minimal disruption, and be paid. (Participating organisation)

- Their turn-down DSR site(s) could feasibly meet TA metering requirements (either having simple metering and no renewables[23] or other onsite generation; or being big enough to justify investment in metering and monitoring systems to meet TA requirements).

- They could either add TA revenues at little extra cost ('it's a no brainer'); or they were attracted by the high clearing price of £45/kW for the second TA, that justified the hassle and cost of getting sites set up for the CM (e.g. metering arrangements; contracting with an aggregator).

Our gut feeling is from a production perspective at £16 a [kilo]watt it's probably not for us. Then the second auction cleared at £45 and at that point we went through testing and we are where we are today...(Participating organisation)

- They were in sectors perceived by aggregators to have sites that would be cost-effective for turn-down DSR in the second TA, or they were already aware of the TA themselves. Aggregators marketed actively to organisations in certain sectors such as metal-related, food processing and water industries – see chapter 4 for detail.

- The scale of potential TA revenues was significant for their business (e.g. £10,000 or more) and justified spending management time on risk assessment and initial set-up.

[23] Onsite renewable generation that was already receiving subsidy (e.g. Feed in Tariff) had to be separately metered and subtracted from the TA capacity, to avoid double-subsidy.

- They perceived the risk to TA revenues from non-delivery (including risks arising from changes to their baseline) to be relatively low, at the capacity level for which they contracted.

So, the risks are obvious risks that we can't produce the capacity that we've said we can. But I think we've been quite conservative in what we've put forward. And we've proved that we can do it on the test runs. (Participating organisation)

- They perceived the risks to their core business (e.g. quality of product; timeliness of client deliveries) and the risk of not being able to restart production to be low.

Anytime you interrupt production there's a risk that something will go wrong but it's a small risk, it's a reliable plant. So there is very little risk to us. (Participating organisation)

A further reason given by some participating organisations was that they saw some wider benefit from TA/CM participation (e.g. reputational benefits from helping National Grid with electricity security).

Within this participant group, there were organisations that prequalified turn-down capacity directly for the second TA, and others that participated in the second TA via an aggregator. Our combined findings from Phase 1 and Phase 3 of the evaluation suggest that direct participation in turn-down DSR required considerable experience of providing flexibility, strong organisational capacity and high levels of confidence in understanding the complex CM rules and electricity system, as well as suitable turn-down loads in excess of 500 kW.

Aggregator clients did not need to exhibit all these characteristics. Instead, those participating via an aggregator either had an existing aggregator relationship or were approached by aggregators for the second TA and had the capacity and confidence to engage with them. Reasons for participating via an aggregator included reducing the hassle of participation, giving the organisation more confidence that they could comply with scheme requirements or giving them access to expert market knowledge, in return for sharing TA revenues with the aggregator.

It was the easiest way for us to access it without expending a great deal of internal effort. So we're quite a lean organisation, we don't have loads of spare people who could spend a lot of time looking into something like that. So they have their cuts, if you like, and they do all the administration and the rest of it for us. (Participating organisation)

Reasoning: 'The costs and risks of participating in the second TA are significant for us, but our business is struggling and we need the revenue'

This reasoning was observed for organisations that perceived a significant risk to their business from turning down, owing to the nature of their production process, but chose to

turn down to gain revenue for their struggling business. The choice as to whether to participate directly or via an aggregator depended on similar factors to the group above.

> *As I said, unfortunately our company's struggling, has been for a long time, and any money that we can make out of anything is worth looking at. (Participating organisation)*

Reasoning: 'We operate in a sector that has potential for turn-down in the CM, but don't think the potential returns from the second TA are sufficient to compensate for the investment and risks involved in participation'

This reasoning was observed for non-participating organisations that we interviewed in industrial sectors that looked promising for turn-down DSR in the TA/CM (see chapter 4 for further details). The non-participating organisations that we interviewed in these sectors had been approached by aggregators about the second TA or other flexibility services, but they (or the aggregators that approached them) did not perceive the potential returns to justify the costs or risks involved in TA participation. Despite others in their sector having potential for turn-down DSR, the set-up in their particular organisation was not favourable. Technical reasons given for non-participation by individual organisations included:

- Having production processes that could not readily be stopped and restarted, or that had to run at full capacity 24/7 to meet customer demand.
- Having production processes that could only turn-down for relatively short periods (e.g. 30 minutes or less), that were better suited to providing static FFR services or participating in a mixed DSR CMU in the main CM (alongside back-up generation which could sustain delivery over a longer timeframe).
- Having production processes where the risks associated with turn-down were thought to be unacceptable (e.g. risks to product quality; risk of equipment failure on restarting).

Factors that were mentioned as making some sites not cost-effective for participating aggregators and non-participating organisations in the second TA included:

- Having smaller sites with renewable generation or other onsite generation on them, for which investment in metering for the second TA and/or CM was not cost-effective.
- Having small sites with capacity below 100 kW or 250 kW (unless these sites had straightforward metering and no renewable or onsite generation, or unless they formed part of a larger portfolio for a single client).

> *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 participant)*

- TA revenue being available for only one year, where this was insufficient to justify required investment in metering on small, complex sites.

If we have to invest money, we also want to have a return over several periods of time. We don't have the income and revenue security in order to justify that investment. (Non-participating organisation)

Cultural factors also played an important role. Non-participants that did have suitable production processes, and potentially cost-effective sites, reported that their senior management was not persuaded of the business case for turn-down. A key factor that we observed was that the energy manager did not have enough resources to prepare a full business case to persuade the board (e.g. to reassure them that the risks posed to their core business were low).

[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 is and why we should be doing it. (Non-participating organisation)

Reasoning: 'We are not in a promising sector and don't have the sites, resources or time to participate, either directly or through an aggregator'

Our research was mainly targeted at industrial organisations likely to have suitable sites for turn-down DSR at a scale that justified interest from aggregators. However, we also interviewed some organisations in commercial sectors that had demonstrated an interest in DSR by attending a DSR industry event. These organisations reported that they had a continuous operational requirement and therefore could not turn-down for Triad or the TA/CM, although some had participated in FFR. Contributory factors included:

- Only being able to offer turn-down DSR for short time periods (e.g. less than 30 minutes) because of operational requirements.
- Offering only very small loads (e.g. because of having no central building management system).
- Loads being too small to be cost-effectively aggregated for the CM using current technology.
- Organisations finding it easier to offer flexibility using back-up generation assets than turn-down DSR.
- Organisations not having the scale of resource or time to participate directly, so being dependent on interest from aggregators.

They [aggregators] were all looking for the big loads, the easy wins, and you can understand that from a business-case perspective. It's just going to be a lot easier to do, but we don't have that. (Non-participating organisation)

3. Reasons for capacity changes in the second TA

At the time of writing, 293 MW (40%) of the capacity that participants attempted to prequalify for the second TA was contracted and going forward for delivery. The main drop-out point for capacity in the second TA was withdrawal of capacity before the auction (269 MW). This was because of changes in the circumstances of specific aggregator clients, and downwards revisions in the capacity that aggregators and direct participants thought they could realistically contract in the second TA. The low clearing price in the Early Auction (£6.95/kW), just before the second TA auction, appears to have contributed towards these downward revisions, by dampening interest in the Capacity Market and second TA. The high clearing price in the second TA (£45/kW) was partly attributable to the low liquidity in the second TA auction, caused by capacity withdrawing before the auction. There was little drop-out of capacity during the testing stage, after the auction. This was partly because of learning from the first TA and partly because the high price helped aggregators to attract clients. Most participants overfilled their CMUs (i.e. lined up more capacity than strictly needed) as a precaution against losing capacity during testing or delivery. This resulted in 187 MW of additional capacity that was 'proven' in a DSR test but was not contracted for the second TA.

About the second TA process

The steps in the second TA process were identical to those for other CM auctions. Capacity could drop out at each of these steps:

1. **Prequalification.** This was the process by which National Grid ensured that capacity submitted in autumn 2016 met the CM rules, including specific criteria for the second TA (e.g. turn-down DSR only, CMUs between 0.5 MW and 50 MW);

2. **Submission for auction.** Before the auction, in March 2017, participants could decide whether they wanted to submit some or all of their prequalified capacity to the auction;

3. **Auction bidding.** In this 'descending clock' auction, participants could choose to submit an exit price for each CMU in each round. Their CMU would exit from the auction if the price got down to this level. Participants could also choose not to submit an exit price. Any CMUs remaining without an exit price at the end of the auction would automatically receive a CM agreement at the final auction clearing price.

4. **Passing metering and DSR testing requirements.** Participants that received CM agreements after the auction had until the end of August 2017 to identify meter points for unproven capacity, provide metering assessments, meet metering testing requirements (where required) and undertake DSR tests to prove DSR capacity (for unproven DSR CMUs). DSR test requirements are explained in the Phase 2 report, Appendix 1. Metering tests were required for sites which had bespoke or balancing services metering but were not required for sites with supplier settlement metering. Sub-metering (at similar levels of accuracy) was required for subsidised renewable generation on TA sites, to avoid double-subsidy. To satisfy the turn-down only requirement for the second TA, other onsite generation also had to be separately metered. Alternatively, a declaration[24] could be issued to satisfy National Grid, the Electricity Market Reform Settlements body (EMRS) and BEIS that onsite generation would not be used during a stress event or test. The testing requirements for the main CM are similar, except that separate metering or declarations for non-subsidised onsite generation are not required.

5. **Proven DSR capacities.** The proven capacity of DSR could be higher or lower than the contracted capacity. If lower, the Capacity Agreement volume was reduced to reflect the proven capacity. If higher, the contracted capacity remained unchanged.

We proposed a candidate theory to explain how organisations would behave before and during the auction, and in response to testing requirements. This is outlined in Appendix 4. This theory has been revised in response to the evidence, as set out in Appendix 5.

Summary of outcomes for the second TA

Figure 3.1 highlights that the main stages when capacity dropped out in the second TA were withdrawal of capacity before the auction (269 MW), and failure to prequalify (114 MW). Relatively little capacity was unsuccessful at auction (61 MW) and only 19 MW dropped out because of failure to complete testing (i.e. termination) or lower than expected capacity in DSR tests. This left 293 MW in 28 CMUs going forward to delivery. However, 24 out of these 28 CMUs had higher than expected capacity in their DSR tests, with a reserve of 187 MW of capacity that was proven but not contracted.

[24] In the run-up to the delivery year, there was a lack of clarity between National Grid, EMRS and BEIS about whether a declaration would be sufficient to prove that capacity was provided by turn-down DSR. BEIS issued a clarifying email in late August 2017.

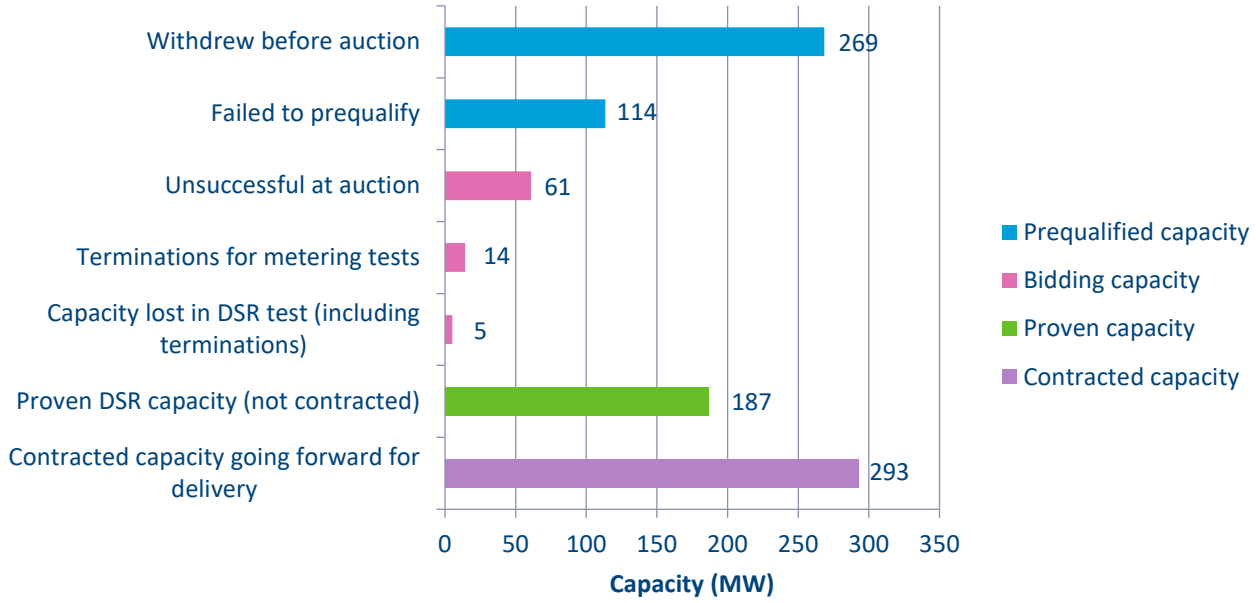


Figure 3.1: Change in capacity provided in second TA, by steps in TA process

While the proportion of capacity dropping out between prequalification and delivery was about 60% in both the first and second TA, more of the drop-out in the second TA was attributable to withdrawal before the auction, and less was attributable to exiting in the auction or failing to meet testing requirements. A comparison of the proportions of capacity dropping out of the two TA auctions for different reasons, after attempting to prequalify, is shown in Figure 3.2.[25]

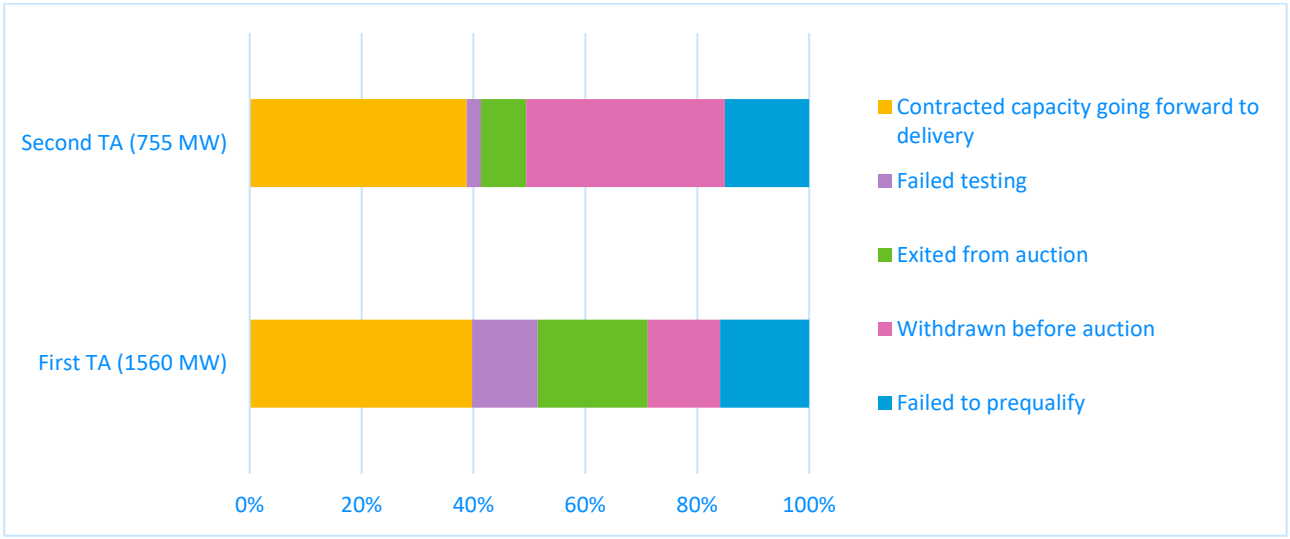


Figure 3.2: Comparison of reasons for drop-out in first and second TA

[25] Totals may not add owing to rounding errors.

There was some evidence that proven DSR capacity also exceeded contracted capacity in the first TA, for certain CMUs. While the surplus in the second TA was 187 MW (64% of contracted DSR capacity), the equivalent surplus in the first TA was 88 MW (29% of contracted DSR capacity. In the first TA, more than half of this surplus was contributed by two direct participant CMUs, but the dominant effect of DSR tests was the loss of proven DSR capacity on CMUs that had recruitment or testing problems.

The reasons for the second TA outcomes are explored below, while the reasons for first TA outcomes are set out in the Phase 2 report. In the section below, we have ascribed findings to aggregators or direct participants, where relevant, except where this could potentially disclose commercially-sensitive bidding strategies.

What capacity failed to prequalify and why?

Fourteen organisations put forward CMUs for prequalification for the second TA, but six CMUs (114 MW) did not prequalify. Four of the organisations did not prequalify any CMUs, leaving ten[26] organisations in the TA. We identified two main types of reasoning by organisations that had CMUs that did not prequalify.

Reasoning: 'CMU not eligible because of errors or problems outside our control'
Factors contributing to this reasoning were that:

- One direct participant put forward generating capacity that was ineligible, because they were not fully aware that the second TA was restricted to turn-down DSR only.
- One aggregator's CMUs could not be entered into the auction because these CMUs (or at least CMUs with the same name) had cleared in the T-4 auction in December 2016. Just before the auction, National Grid noticed this anomaly and disqualified these CMUs.

Reasoning: 'This is not for us at this time, so withdraw'
Factors contributing to this reasoning were that:

- One aggregator withdrew a CMU because their client's capacity was not going to be available in time for the second TA, for reasons related to another flexibility service.
- Another aggregator had turn-down clients that chose to clear in the Early Auction. This made their capacity ineligible for the second TA, but allowed them to put their capacity into the third T-4 auction for delivery in 2020/21,

[26] At this stage, there were ten rather than 11 TA participants because the additional direct participant had not yet purchased TA assets.

from which second TA capacity was excluded. This gave them access to a continuous stream of future T-4 revenues.

What capacity was withdrawn between prequalification and auction, and why?

Four aggregators withdrew some of their CMUs before the auction, comprising 6 CMUs (61 MW), while a further four organisations (two aggregators, two direct participants) reduced the capacity of their CMUs by 208 MW in all (across 18 CMUs). All of the organisations that prequalified were left with some capacity in the auction, so there were still ten[27] organisations at this stage.

We observed three types of reasoning by organisations reviewing the prequalified capacity they would forward to the auction, as follows.

Reasoning: 'Things have changed so we can't offer all the capacity we had planned'
A combination of factors underlay this reasoning by aggregators:

- Some aggregators had taken a speculative approach to prequalification capacity, because they were not sure how much capacity they could put forward. In some cases they needed to recruit new clients, while in others they were not sure how much capacity their existing clients could put forward.

It's prudent to prequalify a bit more unproven DSR capacity and then decide whether to take it into the auction or not, prior to the auction. (Aggregator participant)

- Aggregators were aware of the tight timeframe from TA auction to the deadline for completion of testing at end August.

- Some had dampened price expectations for the second TA, and the CM in general, arising from the low price in the Early Auction in January 2017. This applied to aggregators that needed to recruit clients to fill their CMUs and were less confident they could do this.

I dropped a lot before we went into the auction because of the impact I would say of the [Early Auction] price. Salesmen were not that interested in pushing it as a primary product given the amount of work that's required. (Aggregator participant)

- Some aggregators were affected by changes to particular sites for known clients. For example, an experienced aggregator had a low-risk strategy of

[27] At this stage, there were ten rather than 11 TA participants because the additional direct participant had not yet purchased TA assets.

only putting existing clients forward to the auction. In this case, the volume put into the auction was determined through discussion with that client.

We only will generally go into the auction if we have a back-to-back contract with a customer to fulfil the obligation. (Aggregator participant)

Reasoning: 'The proven capacity of our CMU is higher than we can consistently provide'

At least one direct participant had prequalified capacity based on the proven capacity for this CMU in the first TA DSR test (which exceeded their contracted capacity in the first TA). They reduced the capacity going forward to the second auction to the level that they were confident that they could consistently deliver, at expected TA prices.

Reasoning: 'This is still a good deal for us'

Those organisations that went into the auction with their full prequalified capacity reasoned that the second TA still provided a good deal. They either had known clients that still wanted to participate in the TA or were confident that they could recruit additional clients, provided the clearing price was high enough. An example of a contributory motivation was that an established aggregator wanted to demonstrate their reliability as a DSR provider to National Grid.

What bidding strategies were used in the auction, and why?

The clearing price of £45/kW was significantly higher than previous CM auctions, including the first TA in March 2016 (£27.50/kW), the first four T-4 auctions (ranged between £8.40 and £22.50/kW, between autumn 2014 and March 2018), the Early Auction in autumn 2017 (£6.95/kW) and the first T-1 auction in January 2018 (£6.00/kW). This may partly reflect the restriction to turn-down DSR, which is more challenging for aggregators to recruit than back-up DSR, but it may also reflect lower liquidity in the auction owing to lower prequalification volumes than the first TA and withdrawal of capacity before the auction (for reasons explained above). Our analysis here is based on observed bids and on the bidding strategies reported by auction participants, rather than on theoretical analysis of bidding strategies under price uncertainty, which is beyond the scope of this study.

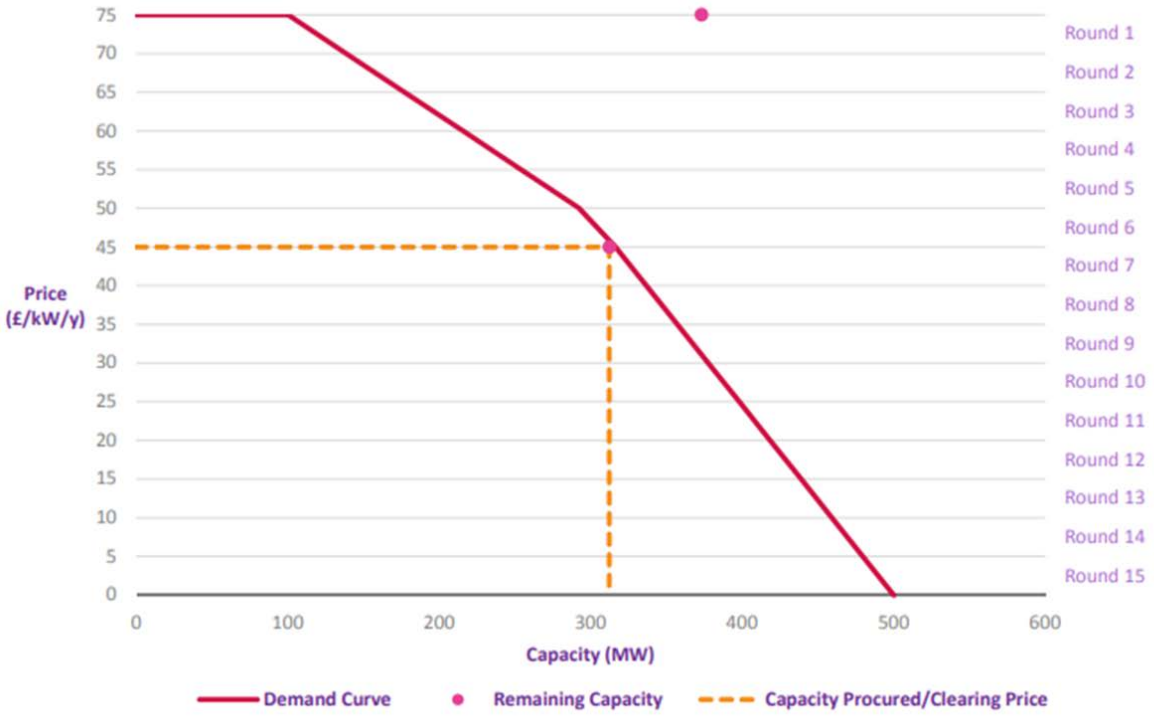


Figure 3.3: Demand curve and clearing price for the second TA auction

Price expectations varied between participants, according to their perceptions of the market. The price expectations of some players were dampened by the low clearing price in the Early Auction.

> *I don't think anyone went into the TA, expecting that kind of revenue outcome, considering the 'T-1' [EA – sic] had cleared at such a low price. (Auction participant)*

However, at least one participant expected the auction to clear higher than the first TA, because it was restricted to turn-down DSR only. All the bidders had low expectations in the number of stress events in the 2017/18 delivery year.

We identified three main types of reasoning to explain bidders' behaviour in the auction, based on motivations reported in interview. While most bidders explained their bidding strategies in terms of covering costs, or generating a minimum level of revenue, it is possible that they may have been less willing to report other profit maximisation strategies. To avoid disclosure of commercially-sensitive bidding strategies, we present joint findings for aggregators and direct participants in this section.

Reasoning: 'We need to submit a high price to cover high costs of recruiting capacity or turning-down'

Three CMUs (61 MW) failed to clear in the auction. The two bidders for these CMUs reported that they submitted exit bids above the clearing price of £45/kW because:

- One participant reported that they had not yet recruited clients and wanted to ensure that they had obtained a high price to recruit high volumes of new clients.

- The other reported that they were influenced by their client's nervousness about the opportunity costs of a four-hour shutdown.

One of these organisations was left with no capacity going forward to delivery, so this reduced the number of participant organisations from ten to nine[28].

Reasoning: 'We need to submit (or be prepared to submit) a medium exit price to cover the costs of recruiting capacity or turning down'

Bidders submitting (or preparing to submit) medium exit prices reported that they sought to cover costs or ensure a given level of revenue above their anticipated costs. The factors that they said underlay their bidding strategies varied:

- Some bidders submitted a range of different prices, which all cleared. They said that this was either because:

o the price was influenced by the opportunity costs of known clients/sites;

As they're the ones who would be turning down and having to put up with those marginal costs of lost production etc. … all the prices are determined by the customers … Obviously, we will help input and help with strategy, but it's ultimately the customer's final decision. (Auction participant)

o or because they felt more confident of recruiting greater volumes of new clients at higher prices.

We don't want to be on hook for delivering potentially large volume capacity requiring, you know, lots of effort to deal with lots of clients at prices that we know we don't stand a chance of even breaking even on. (Auction participant)

- Others did not submit an exit price but reported that they had a price in mind and would have bid if the clearing price had descended closer to their minimum level. They said that they were reluctant to enter a bid until necessary because they lacked confidence in using the bidding system.

I think we just wanted to enter [the price] when we got to the right round in case there was any manual error of exiting prior to that… I think it was just an extra layer of safety to make sure there was no error. (Auction participant)

[28] At this stage, the additional direct participant had not yet purchased TA assets.

Reasoning: 'We can submit a low or minimal exit price, or no exit price'
Two participants reported that they submitted low or minimal prices for their CMUs, because their main interest was covering low levels of cost and using the second TA as a learning experience.

> *It was more of a breakeven goal [..]. Because this was a strategic focus, it wasn't necessarily a money maker for us. It was more getting some experience. (Auction participant)*

Another participant did not submit an exit price but stated that they would have been willing to accept any clearing price, because their clients had low opportunity costs, were already providing turn-down DSR for other services and they had a strategic goal of obtaining CM revenue for these clients, to maintain their client relationship.

> *If you don't secure any capacity to give them a chance to earn revenue, any revenue, and continue to be involved, then you effectively could damage that relationship you've got with that customer. (Auction participant)*

Unobserved reasoning: 'Our bids are designed to maximise our profit'
It is impossible to tell how far bids were really based on 'profit maximisation' rather than generating a given level of revenue or profit over anticipated costs. One bidder did report that their bidding strategy for any CM auction was based on their view of the auction as a whole, as well as their assessment of the costs that they would need to cover.

How did capacity change because of metering and DSR testing, and why?

As explained above, only 19 MW (6%) of capacity dropped out of the second TA post auction. Four CMUs (14 MW) were terminated because of failure to obtain metering certificates. Two further CMUs (1 MW) were terminated for failing to obtain a valid DSR certificate, while four other CMUs lost 4 MW in total because their proven capacity was lower than their contracted capacity.

The volume that dropped out at this stage was much lower than in the first TA, where there was a 182 MW (22%) reduction in capacity at the testing stage, primarily attributable to problems recruiting 'unproven DSR' capacity that could meet metering and DSR tests within the timeframe from auction to delivery.

We have undertaken preliminary research on testing issues for the second TA but have not formally tested theory in this area. Our early findings, which will be explored further in Phase 4, suggest that the reasons for the low levels of drop-out in the second TA, compared to the first, were:

- Organisations having higher levels of awareness about the demands of metering tests, because they had more CM experience.

- The higher price making it easier to recruit capacity, despite the challenges of recruiting turn-down DSR that met metering requirements within a tight timescale.

Unlike the first TA, participants did not mention any capacity dropping out because of problems obtaining grid connection agreements. While some back-up generation in the first TA had not previously been connected to the electricity grid, turn-down loads were – by their nature - all already grid-connected.

Our findings in relation to metering and DSR testing are explored in turn below. We have reported findings for aggregators and direct participants separately in this section, as this distinction helps to explain the findings.

How did metering tests affect outcomes in the second TA?

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 undertake metering tests.

As found in Phase 2, metering tests were perceived as highly onerous, because of the high levels of accuracy required for meters and associated transformers, and because it was difficult to provide the necessary documentation for older equipment. Preparing metering statements was a particularly challenging part of metering testing: several aggregators commented that this required expert input.

> *That metering statement process is so bad now that I would never let a customer try and complete that form on their own, ever. (Aggregator participant)*

Experienced aggregators avoided or minimised metering testing by selecting sites without onsite generation or renewables behind the meter, so that supplier settlement metering could be used and testing was not required. One direct participant had less choice of sites and found requirements onerous.

> *Do I really need to give you the grid references from all my sites? Do I really need to give you the calibration certificate, the current voltage transformers on the meters that were put in 60 years ago when I am getting involved [in the CM]? (Direct participant)*

Where metering tests were undertaken, this was generally for larger sites where the investment of time and money in metering was cost-effective. Metering tests effectively excluded small sites (below about 250 kW) with complex metering or onsite renewable generation from participating. This is broadly consistent with findings from academic research that a UK-based aggregator required a minimum size of 200 kW for turn-down

sites.[29] Some participants commented that changes were needed to make it more feasible for small turn-down sites to participate in the CM.

> *You need a more streamlined metering process if you want to deal with quite large numbers of small assets. I think in load turndown the [next] step[..] will be to expand it to lots more smaller assets, but to do so you need a better process for metering assurance than is currently provided in the Capacity Market. (Aggregator participant)*

The second TA was the only CM auction for which onsite non-renewable generation had to be separately metered. The rules stated that participants could make a declaration that onsite generation would not be used during a test or stress event. At least one participant that faced significant investment in metering made use of a declaration, having ascertained through considerable discussion with National Grid, EMRS and BEIS that this was acceptable. EMRS indicated that another participant decided to invest in metering anyway because they did not receive assurance in time that the declaration would be deemed compliant.

As in the first TA, several comments were made about problems arising from the split of responsibilities between National Grid and EMRS on metering issues, which meant that participants were often referred from one body to another.

> *If you are talking to the delivering body [i.e.National Grid] about something and you ask a question about metering, they say, "We are not going to take a view on that. Go and speak to the settlements people [i.e. EMRS]." Then, the settlements people won't really understand the context that you are asking the question in, because they don't really care about the composition of your CMU, or whatever it is. (Direct participant)*

How did DSR tests affect outcomes in the second TA?

DSR tests resulted in loss of only 5 MW of capacity (4 MW lost because of proven capacity being lower than contracted capacity, and 1 MW lost because of DSR tests not being passed). Respondents reported that DSR tests were more straightforward for single-site CMUs, as the test could be timed to obtain good baseline demand and to suit operations for one site. Despite this, at least half of this lost capacity was in single-site CMUs, for reasons that we do not yet understand and will explore further in Phase 4. The reasoning behind 'single-site' and 'portfolio' approaches to CMU design is explained in chapter 4.

Those with large 'portfolio' CMUs (i.e. those containing many components) reported that DSR testing was challenging, because of coordinating turn-down across multiple components. They 'overfilled' their CMUs, allowing a margin in case some capacity dropped out because of baseline changes or operational issues. However, little capacity

[29] M. Curtis, "Demand side response aggregators: How they decide customer suitability," *2017 14th International Conference on the European Energy Market (EEM)*, Dresden, 2017

dropped out in the DSR tests for these CMUs, and the proven capacity for most CMUs exceeded their contracted capacity, as discussed further below.

Some aggregators chose to test some or all of their CMUs jointly, consistent with their 'portfolio' approach to CMU design. This approach was only chosen by aggregators that had experience in coordinating turn-down across numerous components, and that had these CMUs ready for testing well ahead of the end August deadline. Their reasoning was that over and under performance would net out between different components and CMUs.

> *It's important that we are pooling our performance across all of our CMUs, so that any unexpected over performance from one CMU or one component will benefit the portfolio as a whole. (Aggregator participant)*

Those aggregators that left DSR testing until August, because they were inexperienced or were still waiting for metering test certificates for some sites, had problems obtaining DSR test certificates before end August. This was because of the time required by EMRS to process their DSR test data, after completion of the tests themselves. However, this resulted in little capacity being lost because almost all of this capacity passed on appeal.

What CMUs were overfilled and why?

There was consistent evidence of 'overfilling' CMUs to avoid DSR test (and potentially delivery) failure. The DSR test results, combined with interview evidence, indicate that those with 'single-site' CMUs understated the maximum capacity that their site could potentially deliver to mitigate risks associated with baseline demand and site operations. Those with a 'portfolio' approach to CMU design may also have been conservative in estimating the delivery capacity of individual components, but they also included additional components in their CMUs by way of insurance for under or non-delivery of any one component for a particular turn-down event.

While some overfilling was evident in the first TA (see beginning of chapter 3), the extent of overfilling was greater in the second TA. As shown in Figure 3.4 below, most CMUs demonstrated proven capacity in the DSR test that was more than 20% above their contracted capacity secured at auction, and nearly half demonstrated more than 50% excess. Overall, proven capacity was 64% higher than contracted capacity.

For direct participant CMUs, proven capacity was on average more than 100% higher than contracted capacity. We will seek further information in Phase 4 to understand more fully why overfilling was high amongst direct participants.

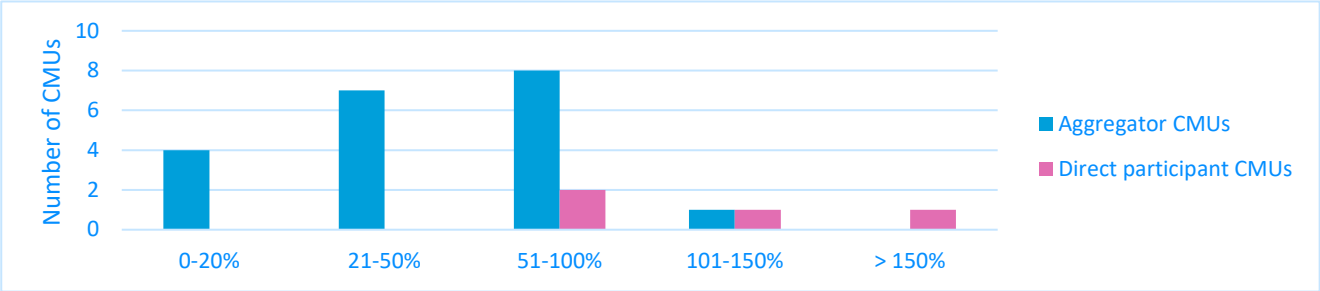


Figure 3.4: Excess of proven capacity over contracted capacity for overfilled CMUs in second TA (as a proportion of contracted capacity) (%)

The high level of overfilling was driven largely by learning from the first TA (in which many aggregator CMUs failed to demonstrate their proven capacity) but was enabled by the high clearing price for the second TA (which facilitated recruitment by aggregators and provided an incentive for participants to demonstrate their full capacity).

We worked with a much higher buffer this year. (Aggregator participant)

Participants were aware that under-performance in the DSR test would result in them losing revenue for any lost capacity. They also stood to lose their credit cover if their proven capacity for a CMU fell below 95% of their contracted capacity. Both the test results and interview evidence suggests that they therefore made every effort to deliver the required volume for DSR tests.

It is worth noting that a stress event would occur at shorter notice than a DSR test, probably at a more inconvenient time and possibly for longer than the 30-minute DSR test window. If a CMU under-performs in a stress event, it would lose only a proportion of its revenue from the second TA (whereas all revenue is lost if a CMUs fails its DSR test) so the motivation to deliver maximum volume might be lower.

While overfilling of CMUs increases the reliability of delivery, it also reduces the CM revenues that can be obtained for a given volume of capacity. One aggregator commented that they resented not being paid for the surplus capacity in their overfilled CMUs.

4. Early findings on topics to be researched further in Phase 4

Phase 3 findings indicate that there were two strategies for CMU design: simple CMUs with single-sites or single-clients which were simple to administer and test; and portfolio CMUs with multiple clients and sites, which were more complex to administer and test but potentially offered more diverse and therefore more reliable capacity. The implications of these CMU designs for reliable delivery will be researched further in Phase 4.

Phase 3 research also indicates that almost all of the capacity in the second TA was provided by industrial sites rather than commercial sites, such as offices. While a large proportion of sites were provided by the water, sewage and food processing industries, many of these sites were small (e.g. a few hundred kW or less) and involved pumps, motors and drives. Fewer sites were provided by heavy industry (e.g. metal-working, construction) but these sites tended to be bigger (e.g. many hundred kW or several MW) and to involve process heating or bespoke industrial processes. The nature of turn-down DSR provided in the second TA, and the costs associated with delivery, will be researched further in Phase 4.

This section presents early findings from Phase 3 research on evaluation topics that will be researched further in Phase 4, namely:

- How does the design of CMUs affect the reliability of delivery?
- What types of turn-down DSR have participated in the second TA?

How does the design of CMUs affect the reliability of delivery?

During Phase 3 research, we asked participants how they designed their CMUs. The findings on size and composition of sites, overfilling and preferences for simple sites that do not need metering tests have been presented above.

We consider here whether the number of ‘components’ in a CMU will affect the reliability of delivery during a stress event.

TA participants could choose whether to put forward one large site or several smaller sites in each CMU, provided the total capacity was between 500 kW and 50 MW. Aggregators could have separate CMUs for each client’s sites or could create diverse CMUs with a mix of capacity from different clients.

Our analysis of CMU data from National Grid indicated that the 28 CMUs going forward to delivery comprised 304 separate components, each consisting of a separate location or site. Figure 4.1 below shows that over a third of these CMUs comprised only one site, but that some CMUs had large numbers of sites. On average, there were 11 sites per CMU.

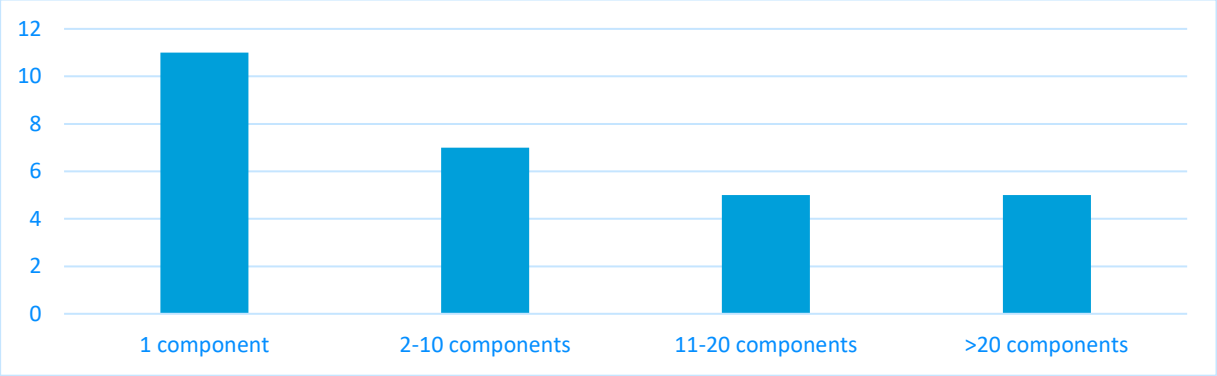


Figure 4.1: Number of components per CMU (in CMUs going forward to delivery)

We asked TA participants about their rationale in putting forward single-site CMUs or multi-site CMUs containing a diverse portfolio of sites and clients. Two distinct strategies were apparent, a 'simple CMU' approach and a 'portfolio CMU' approach, as explained below. A couple of organisations put forward CMUs of both types (e.g. because they could not find enough sites above the 500 kW threshold to use the 'simple CMU' approach for all their CMUs, or because some of their recruitment activity was sub-contracted) but most had a clear preference for one strategy or the other.

TA participants made little comment during Phase 3 research about the effect of the reduction in the minimum CMU size from 2 MW in the first TA to 500 kW in the second TA. The Phase 2 report put forward evidence that this change was welcomed, and Phase 3 research confirmed that some aggregators did put forward small CMUs below 2 MW in the second TA.

'Simple CMU' approach

Where aggregators or direct participants had larger sites, above the 500 kW threshold, some chose to put these into 'single-site' or 'single-client' CMUs. These participants reasoned that single-site CMUs were preferable because the site could choose the timing of shut-down for the DSR test, to fit their own operations. Also, they reasoned that it would generally be possible for a single-site CMU to demonstrate Satisfactory Day Performance delivery using periods when the site was turning-down for other reasons.

> *For example, if you're doing your load turndown with a factory, to arrange simultaneous shutdowns of lots of sites can be quite intrusive in their operations. To do it with one site is far less intrusive because you can choose it at a convenient time, can't you? (Aggregator participant)*

The aggregators pursuing this strategy also reasoned that it was easier to allocate penalties if a single-site or single-client CMU failed to deliver.

If you have two customers in an aggregated CMU and one of those customers fails, the whole CMU fails, so both customers get penalised. Obviously then, as the aggregator in the middle, we've either got the choice of [taking] on that risk or we have to pass that risk on to customers. Yes, it's a lot trickier. (Aggregator participants)

At least one aggregator welcomed the reduction in minimum CMU size from 2 MW in the first TA to 500 kW in the second TA, because it was more feasible for them to create relatively simple CMUs in which responsibility for failure was clear.

It was great that we could offer smaller CMUs because basically, it is my experience that it's much better not to have a lot of small customers in, say, a 2 MW or a 5 MW CMU. It's better that they're responsible for their own success or failure. (Aggregator participant)

There was some evidence that the aggregators taking this approach thought that it would not be cost-effective for them to aggregate many small sites, because of the capital investment they would need to make in administration, metering and/or controls. This may suggest that they did not have control systems well-suited to handling many small sites.

We've been pursuing a model of getting fewer but larger customers with big industrial sites, with a large amount of capacity in one unit, rather than chasing or pursuing customers with lots and lots of very small amounts of capacity, using capital-intensive systems for activating small amounts of capacity. (Aggregator participant)

'Portfolio CMU' approach

Some aggregators and one direct participant put forward medium to large DSR CMUs that consisted of several smaller sites. This approach was taken by organisations that had the resources, technical capability and confidence to manage delivery across multiple sites cost-effectively. Aggregators taking this approach were more selective in putting forward sites that did not require metering tests or costly investment in metering.

Up to a point, a 'portfolio approach' is a natural consequence of putting forward smaller sites since, under TA rules, a site below 500 kW derated capacity could not be put into a CMU by itself. However, the organisations taking the 'portfolio CMU' approach created multiple-site CMUs even if the sites themselves were above 500 kW, provided this did not breach the maximum capacity limit for the TA.

These participants said that they explicitly chose a portfolio approach as this made their CMUs less likely to fail during a DSR test or stress event: under-delivery by one customer or site could be compensated by over-delivery on other sites. They reasoned that the diversity of types and sizes of customer, and differing seasonality of their electricity demand, made their CMUs more robust.

I think it comes down to a mixture of size of customers, and flexibility, at different times of the year. (Aggregator participant)

One of the reasons for including relatively small sites (down to 100-200 kW) was that, while large loads were more cost-effective to recruit, their delivery could be unpredictable.

> *A lot of these loads they can [be] volatile, steel mills in particular will be fairly volatile loads, they have relatively irregular collection patterns. We need to balance that with other loads, to make sure that we have a stable availability. That's why it's really important to have a good diversity, in order to have a full, robust portfolio. (Aggregator participant)*

There was some suggestion, from Phase 2 case studies and from Phase 3 interviews, that aggregators taking a 'portfolio CMU' approach chose to overfill their CMUs to a greater extent than those taking a 'simple CMU' approach, both to reduce the risk of penalties and to avoid issues associated with allocating any penalties between clients. However, this is not fully consistent with DSR testing results, which show a mix of overfilling outcomes on both 'portfolio' and 'simple' CMUs.

Implications for reliability of delivery

These findings suggest that portfolio CMUs may offer more reliable delivery than simple CMUs, at a CMU level, because they offer diverse capacity. However, it is not clear whether this has implications for the reliability of delivery from the second TA portfolio as a whole, looking across all CMUs. The implications of the 'simple CMU' and 'portfolio CMU' approaches for delivery will be researched further during Phase 4, through research on SPD outcomes and any responses to CM Notices during the winter period.

What types of turn-down DSR have participated in the second TA?

We analysed CMU data provided by National Grid to identify the location of meter points in the second TA. We then used commercial databases to identify the organisations operating at these locations and the sector in which they operate. This gave us a comprehensive database of the sites participating in the second TA.

We combined the site data with further information gathered from aggregators, direct participants and aggregator clients on the nature and volume of capacity available at each site. This additional data was primarily based on email survey returns, supplemented with interview data where available. This data is incomplete and will be extended in Phase 4, but early findings are presented here. Site-level capacity estimates were provided by TA participants, presumably based on performance in recent DSR tests. While this provided one 'snapshot' of delivery capacity, the level of capacity delivered by particular sites may vary between turn-down events.

Preliminary analysis by business activity

The analysis of sites indicated that the turn-down capacity in the second TA was almost entirely provided by industrial rather than commercial sites such as offices. This is consistent with interview findings from aggregators who reported that they targeted industrial sectors and potential clients with good potential for flexibility, using their own

expertise in DSR, their knowledge of supply data (for aggregators that were also electricity suppliers) and their existing client base.

A full breakdown of sites by sector is provided in Figure 4.2 below.

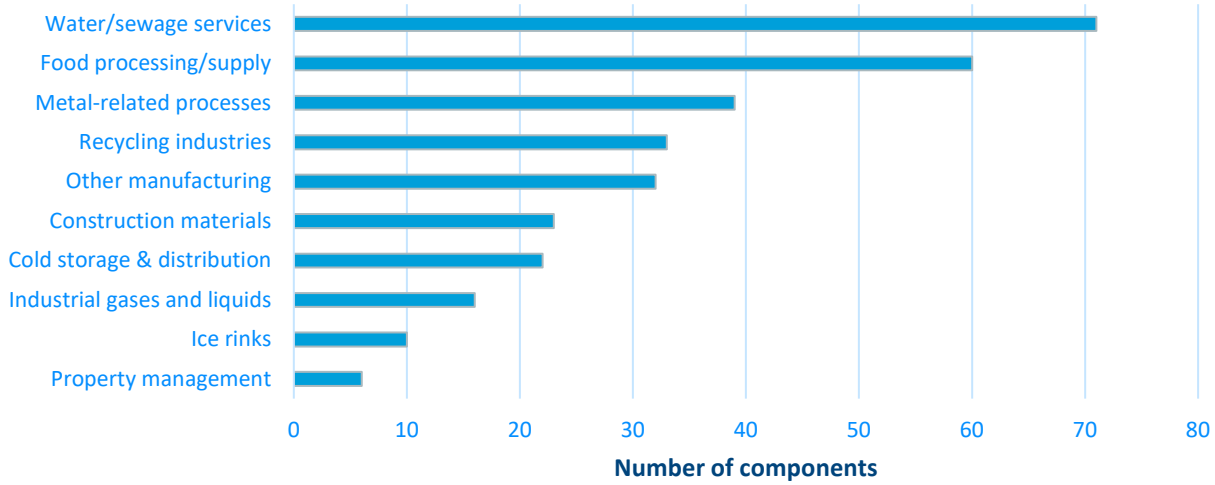


Figure 4.2: Breakdown of sites in second TA, by sector (source: National Grid meter data)

We estimated typical capacity per site for different types of turn-down assets, based on email survey, CM register and interview information. Figure 4.3 presents early evidence that process heating and 'other' bespoke industrial processes typically contribute much higher capacity per site than refrigeration, chillers, pumps, other motors and drives. This suggests that sectors such as water and sewerage put forward many small pumping sites, while sectors such as metal-related processes and construction materials put forward a smaller number of sites involving much larger process-heating loads.

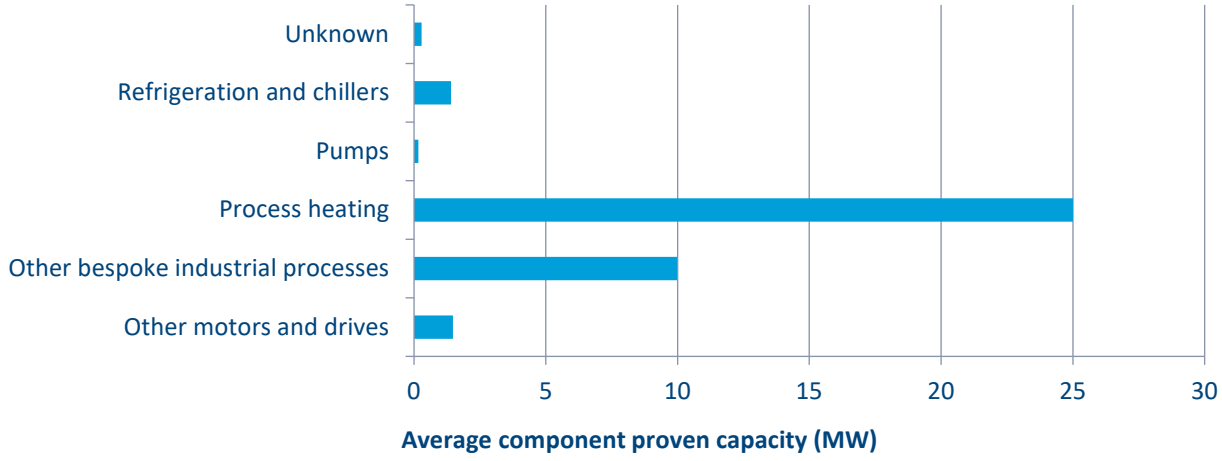


Figure 4.3: Estimated capacity per site by type of turn-down asset, where known (source: email survey data)

An estimated breakdown of capacity by sector and by type of asset will be developed in Phase 4.

Load shifting vs load shedding

We attempted to assess the degree of 'load shedding' (i.e. reducing overall energy consumption) as opposed to load shifting (i.e. moving energy consumption to another time) that is involved in turn-down DSR put forward by the second TA. In particular, we sought to assess the extent to which aggregator clients would expect to undertake load shedding that involved some loss of production that was not made up at other times, as this would have implications for the opportunity costs of turn-down DSR for the organisation's main business.

We therefore asked about the implications of turning down for an organisation's main business, in the screening survey with aggregator clients, the email surveys and the in-depth interviews with aggregator clients. Similar questions will be asked of additional aggregator clients and of direct participants during Phase 4.

Of the fifteen aggregator clients that responded to the screening survey, only three suggested that they were shedding load (i.e. reducing their overall energy consumption) with an impact on production. The remaining twelve were either load shifting, such that they were moving their consumption to another point in time, or load shedding without any impact on production. Of the three participants that reported shedding load with an impact on production, further research would be needed to investigate how much production is actually affected. One of these participants claimed to shed load for the TA only because they were in financial distress and desperately needed TA income.

The limited sample to date provides little evidence of output actually being lost. However, the perceived risk of lost production may in practice be a significant barrier to participation, even if that loss is unlikely to occur. It is also possible that, for a site with multiple assets, a turn-down response may involve a combination of load shedding and load shifting, depending on the levels of stock at different stages of the production process.

The distinction between load shedding and load shifting, and its implications for opportunity costs, will be researched further during Phase 4.

Next steps

The capacity estimates and typologies presented here will be refined further in Phase 4 research. We have started to develop cost and revenue estimates for different types of turn-down assets, including both upfront costs and delivery costs. These cost estimates will also be researched further during Phase 4.

5. Conclusions

It is too early to assess fully the second TA's contribution to its objectives. However, early findings from Phase 3 suggest that the second TA stimulated learning about turn-down DSR for some participants and also encouraged some aggregator clients to expand from 'self-despatch' of turn-down DSR for Triad to delivery of flexibility for the Capacity market – at least on a trial basis. The extent of the TA's contribution to its objectives, relative to other external factors and trends, will be assessed further during Phase 4. Phase 4 will also provide some insights into the value for money provided by the second TA, and the prospects for turn-down DSR in the future Capacity Market.

What contribution has the second TA made to its objectives?

As outlined in chapter 1, the second TA had two policy objectives:

1. To develop a stock of flexible capacity[30] that can 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.

We have not formally tested the second TA's contribution to these objectives, as this will be undertaken during Phase 4 of the evaluation when more delivery information is available. However, Phase 3 evidence has highlighted some preliminary findings that are relevant to the TA's contribution to its objectives.

Firstly, there is evidence from interviews that those aggregators that were cautious about turn-down DSR, and already had large flexibility programmes in GB, largely put forward existing clients for the second TA. However, those that were more confident about the market for turn-down, and sought to grow their flexibility programmes in GB, used the second TA to recruit new clients for turn-down as well as other flexibility services. This is consistent with evidence from Phase 3 interviews with clients of these aggregators, which revealed that some were previously only doing Triad and were treating TA as a pilot for further provision of turn-down DSR in flexibility services and the main CM.

> *I think this year is really the test for us, and [we want to] see how we perform, and how it works for us, but I think if it's successful this year then it will be something that we will continue to take part in and to do. (Aggregator client)*

[30] Flexible capacity means electricity generating capacity and demand that is able to increase or decrease in response to signals, to help balance supply and demand of electricity across the GB grid.

Secondly, there is also evidence from Phase 3 interviews that the second TA stimulated aggregator learning on turn-down DSR. Some aggregators saw turn-down DSR as increasingly important in the Capacity Market in future, given the constraints that the Medium Combustion Plant Directive posed for diesel back-up generation.

> *In hindsight, we're very glad to have done that because of the new draft legislation in place, which affectively wipes out most diesel, most back-up generation for Capacity Markets in the future. We're happy to have learnt what we have from the turndown Capacity Market. (Aggregator)*

The extent of the TA's contribution to its objectives, and its contribution compared to other external factors and trends, will be assessed further during Phase 4. The other potential influences that we will consider in this assessment include:

- The current restructuring of National Grid's flexibility services, which affects the scope for stacking flexibility revenues from different sources.
- Increased awareness of flexibility across potential participant groups, to which the Power Responsive campaign has also contributed.
- Increased funding and support for flexibility industries as proposed in BEIS's Industrial Strategy.
- New constraints on diesel generation arising from the Medium Combustion Plant Directive.
- The effect of Ofgem's review of embedded benefits which have removed Triad benefits for distributed generation and may potentially reduce Triad benefits for 'behind the meter' generation, which would also affect revenue stacking.
- Increased access to wholesale market and balancing mechanism revenues by flexibility providers.
- Improved technology for the automatic aggregation of small loads (e.g. Heating, Ventilation and Air Conditioning - HVAC).
- Cheaper battery technology and increasing participation in flexibility markets by electricity storage.

Other comments about second TA

Participants and non-participants made a number of broader comments about the second TA, which are relevant to the evaluation questions. One non-participant commented that the high price in the second TA auction did not represent good value for money for the consumer, in terms of the price per unit of flexible capacity.

Other views expressed were that:

- The Government should not have changed the rules between the first and second TA, by restricting the second TA to turn-down DSR only, as this meant that the TA was effectively two separate one-year programmes rather than a smooth two-year programme.
- Participants in the second TA should be allowed to participate in all future T-4 auctions, on the basis that volumes and prices in future T-1 auctions are more uncertain.
- Metering rules should be changed to reduce barriers to participation by small sites.
- The CM should offer longer-term contracts for DSR, as many aggregators offer multi-year agreements.
- A third TA should be considered, to give further support for turn-down DSR in the CM.
- If the government seeks to prioritise turn-down DSR in the main CM, it should introduce incentives (e.g. less derating for turn-down DSR CMUs).

These views will be considered further during Phase 4, which will provide some insights into the ‘value for money’ provided by second TA and future prospects for future turn-down DSR in the main Capacity Market.

6. 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 Capacity Market 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 the evaluation. See '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.
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 intervention appears to be

Term or acronym	Definition
	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'.

Term or acronym	Definition
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.
Fast Reserve	A service tendered monthly by National Grid that procures large blocks of reserve capacity (exceeding 50MW) that can respond within 2 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 2 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.'[31]
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).

[31] 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
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.
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.
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 Capacity Market. (See http://www2.nationalgrid.com).

Term or acronym	Definition
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'.
Proven DSR	A unit of DSR capacity that has passed the DSR test required to participate in the GB Capacity Market.
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[32] 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, programs 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.

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

Term or acronym	Definition
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.
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)

Term or acronym	Definition
	• 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)
TA	Transitional Arrangements for DSR and small-scale distribution-connected generation – a pilot consisting of two one-year ahead CM auctions in 2016 and 2017 that are 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