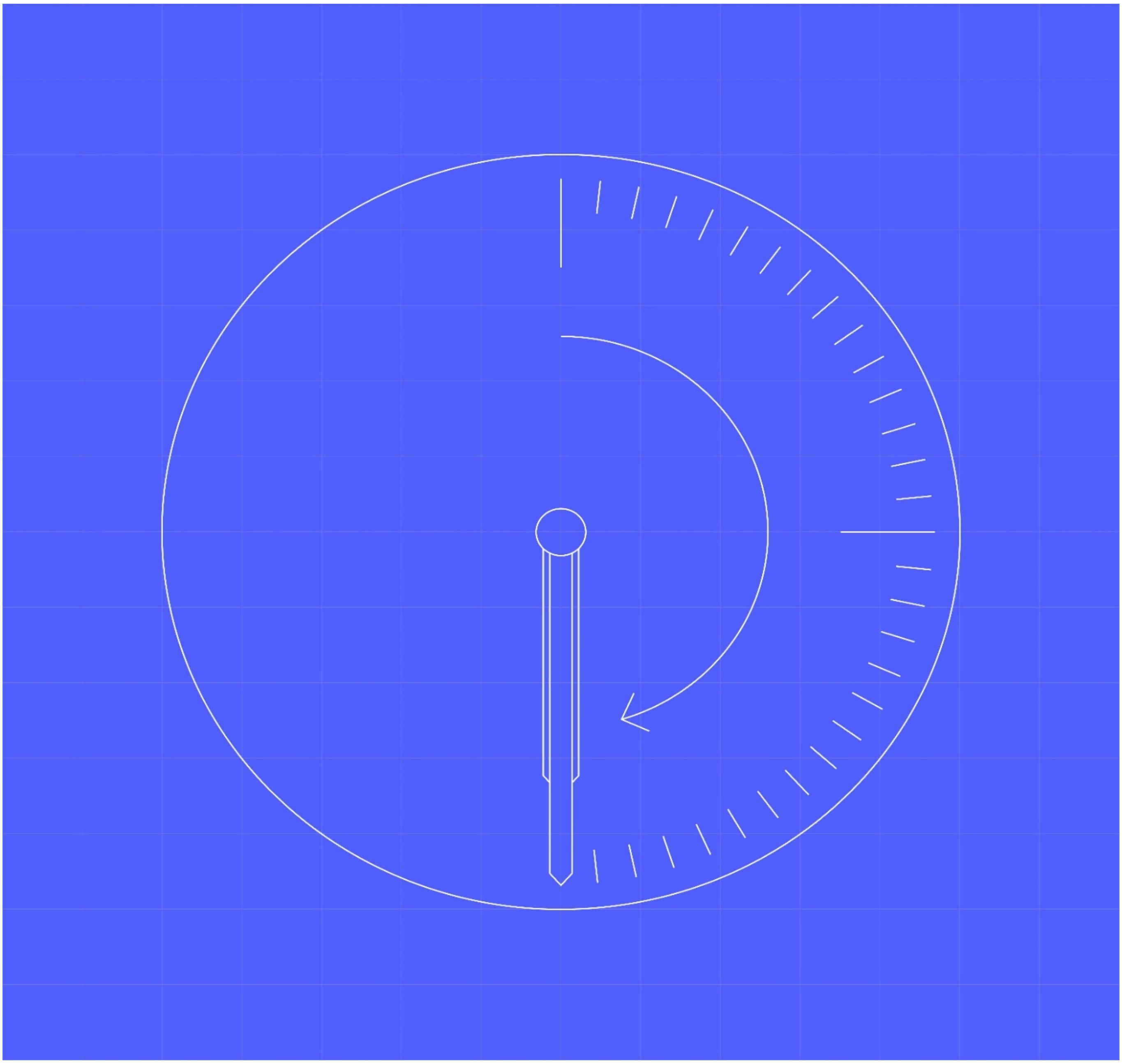
MHHS PROGRAMME
Industry-led, Elexon facilitated

MHHS – CR013 Report



Document Owner:
SASWG Team

Document Number:
DEL1483

Version:
0.2

Status:
Updated draft for review at October PSG

Date:
05/09/2023

Classification:
Public

1 Contents

1.1 Change Record

Date	Author	Version	Change Detail
29/08/2023	SASWG Team	0.1	Draft
05/09/2023	SASWG Team	0.2	Updated draft for review at October PSG

1.2 Terminology

Term	Description
AA	Annualised Advance
BSC	Balancing and Settlement Code
CR	Change Request
CVA	Central Volume Allocation
DNO	Distribution Network Operator
DUoS	Distributed Use of System
EAC	Estimated Annual Consumption
EES	Electricity Enquiry Service
FIT	Feed-in Tariff
GCF	Group Correction Factor
GSP	Grid Supply Point
HH	Half Hourly
kWh	Kilowatt Hour
LLF	Line Loss Factor
LSC	Load Shape Category
LSS	Load Shaping Service
iDNO	Independent Distribution Network Operator
MDS	Market-wide Data Service
MHHS	Market-wide Half Hourly Settlement
MPAN	Meter Point Administration Number
NHH	Non-Half Hourly
PC	Profile Class
PrA	Profile Administrator
PSG	Programme Steering Group
SAA	Settlement Administration Agent
SASWG	Settlement Analysis Scoping Working Group
SDS	Smart Data Service
SIC	Standard Industrial Classification
SVA	Supplier Volume Allocation
RF	Final Reconciliation Settlement Run
ToU	Time of Use

2 Executive Summary

Change Request 013 (CR013), Determining scope of examination of Settlement impacts resulting from Market-wide Half Hourly Settlement (MHHS) Programme, was raised in November 2022 to scope a potential future piece of analysis that would analyse the impacts on Settlement processes of moving to MHHS arrangements. MHHS will impose significant changes to Settlement processes, however relatively little work has been done to investigate this impact to date.

The CR013 work was progressed through a new Working Group convened under the MHHS Programme: the Settlement Analysis Scoping Working Group (SASWG). Through the SASWG, five key questions or hypotheses were chosen for investigation based on the areas Participants predicted the most significant impact, and this has formed the recommended scope of the future analysis.

These hypotheses cover topics such as the impact of new technologies on Load Shapes, the possible commercial Settlement benefits available to Suppliers during Migration, and the impact on Settlement accuracy of Final Reconciliation (RF) coming forwards to 4 months. The full title of each hypothesis can be found in Section 4 (Scope). Should the analysis go ahead, not all five hypotheses need to be investigated together. Participants can pick and choose those that they are most interested in seeing modelled.

Through the SASWG, the data required to model each hypothesis was also agreed on. This is presented in Appendix 3 (Data Required and Data Sources). The majority of this data would need to come from Suppliers and Elexon and the analysis would not be possible without it.

The costs and timescales for two different approaches have been estimated for each hypothesis: Option 1 involves the use of complex models provided by Elexon. Option 2, on the other hand, does not rely on any externally provided models, and instead solely consists of more basic Excel spreadsheet-style modelling. Option 1 would involve longer timescales and higher costs, but, in theory, a more reliable and accurate output than Option 2 and vice-versa. The pros and cons of each approach are weighed up in full at the beginning of Section 5 (Timescales) and Section 6 (Costs).

In terms of next steps, there are three different avenues to deliver the analysis.

1. Proceed with analysis under the MHHS Programme;
2. Proceed with analysis under an Elexon Issues Group;
3. Do not proceed with analysis undertaken by a Central Party, leaving Programme Participants to use this report as a basis for any internal analysis they choose to undertake.

A recommendation in terms of a preferred next steps approach has not been provided as part of this report. Instead, it will be up to Participants to collectively agree the most appropriate route forward and bring this proposal back to the Programme, Elexon or other third party as appropriate.

3 Introduction, Objectives and Approach

3.1 Introduction

The MHHS Programme will impose significant changes to how Settlement processes work as it moves from a predominantly non-half hourly (NHH) Settlement regime using estimation algorithms to the use of actual consumption reads for the majority of sites.

The system and operational processes to deliver this change are well understood and have been the subject of significant development. By contrast, comparatively little work has been undertaken to understand the impact on Settlement itself.

CR013, Determining scope of examination of Settlement impacts resulting from MHHS Programme, was raised in November 2022 and sought to investigate exactly this – What are the impacts on Settlement processes of moving to the new MHHS arrangements and of those of concern, do they warrant further investigation.

Before undertaking a piece of analysis to understand exactly what the impact, in quantitative terms, will be on Settlement processes, it was agreed the preliminary piece of work needed to be a scoping exercise – this is CR013. The scoping phase would identify the areas of Settlement that would be impacted and therefore required further analysis, in addition to costing the work and setting out the potential next steps options to deliver the analysis itself.

It was therefore agreed that CR013 would *exclusively* focus on the scoping of the analysis, and, if upon the completion of CR013, it was decided to proceed with the analysis, this would need to be progressed under a separate and subsequent CR, or another route of facilitation if the analysis was to be progressed outside the MHHS Programme.

The delivery of CR013 will require the production of a report (*this document*) that sets out the scope, approach, timescales, cost and resource requirements for the subsequent analysis. In effect, this should resemble a Project Brief for the analysis.

All deliverables will go to the Programme Steering Group (PSG) for final approval.

3.2 Objectives

3.2.1 CR013 Objective

The objective of CR013 is to ultimately surface and provide the information required to inform a decision from Programme Participants on whether and where to proceed with the subsequent analysis. This information includes the below and will be contained within the final CR013 report:

- Scope;
- Approach;
- Data required and source of this data;
- Format of final outputs of analysis;
- Timescales;
- Costs;
- Next steps options to deliver the analysis.

3.2.2 Subsequent Analysis Objective

Within the CR013 work, there has naturally been discussion not just around the objective of CR013, but also around the objective of the subsequent analysis. This thinking is presented below. However, it is expected that if the analysis is to go ahead, the objective of the analysis will be further deliberated and updated by the analysis team in the early phases of the work.

The objective of the analysis is to model systems and processes in the Settlement Design with key variables modelled to identify and size unintended consequences of the MHHS Design on Settlement processes.

Equipped with this knowledge and an improved understanding of what level of variance (from legacy Settlement processes) is expected as part of migration to MHHS and what level of variance is cause for concern and action, Industry will be able to enter Migration with increased assurance and confidence. If they so choose, this knowledge will also enable Industry to take actions pre-Migration to reduce their level of risk to any impacts.

Note, the Programme does not plan to take any centrally coordinated actions to reduce risk to Settlement impacts, nor will it prescribe actions to Participants. The findings of the analysis will be shared with Industry and then it is up to individual Participants to decide what action they do or do not take.

The objective of the analysis is not to find issues in the baselined Design and suggest alternative solutions – this is what testing is for.

The objective of the analysis is also not to help Suppliers' forecasting. The findings of the analysis may incidentally be used by Suppliers to help their forecasting, however the scope of the analysis will not be driven by this requirement.

3.3 Approach

The recommended approach to the analysis is to start small and simple and iterate. This approach was almost unanimously supported by Participants at the SASWG meetings.

It may be tempting to immediately scope out a large complex model that simulates the entire Settlement system. And whilst this may ultimately be necessary, the recommendation is to start small and simple, and initially seek to model the key questions or hypotheses separately. Afterwards, these separate hypotheses could then be combined into a larger model if necessary. There is a very material risk of building a large expensive model that does not provide any meaningful answers.

The recommendation to proceed with a start small and simple approach is based on the five arguments below:

1. **Speed of initial results** – The start small and simple approach will yield initial results far quicker than attempting to build one large model. An initial and simple prototype could be achieved within a few weeks of work and would enable the running of basic scenario testing. This would quickly provide the materiality of the different hypotheses and would inform any future modelling prioritisation. These initial prototypes could then be re-built in a more sophisticated technology if greater insight is required.
2. **Cost and effort risk management –** Adopting the start small and simple approach would more prudently manage the cost and effort risk associated with the analysis. Cost and effort can be invested progressively and this investment can be assured based on the results seen to date.

   With the one large model approach, potentially very large amounts of cost and effort would need to be invested up-front with no results yet produced to assure this investment. Should the analysis team decide to later change approach, the chances of recovering the spend already invested would be low.
3. **Cumulative impact of assumptions –** If proceeding down the one large model route, the model would need to simulate the entire Settlement system. Given the breadth of what this would need to cover, it will be underpinned by many different assumptions as there are many unknowns as part of MHHS Settlement. When many assumptions are present in a single model, the variance contained within each assumption is multiplied each time a new assumption is applied. This effect could lead to a highly significant level of variance in the final output of the model.

   Instead, if the hypotheses were each modelled separately, this level of variance is contained as there are fewer assumptions overall as they are spread across a larger number of smaller models, rather than all being contained within a single large one. This does not completely eliminate the variance (which will always exist in these assumptions), but it does contain it.
4. **Netting out of directional impacts** – When each hypothesis is investigated in isolation, it will produce a result, of varying scale, either impacting a key metric positively or negatively. Several hypotheses will be investigating the impact on the same key metric e.g. Group Correction Factor (GCF). In the one large model scenario, when the hypotheses are combined, the directionality of each hypothesis' impact is lost as they may counteract one another and net each other out. The true impact will only be seen if the hypotheses are split out.
5. **Challenging to trace impact back to a single hypothesis** – In the one large model scenario, there will be many variables changing simultaneously which, collectively, produce a different end result. For example, if penetration of Smart meters was increased and this produces a different Settlement accuracy figure, if more than one hypothesis has penetration of Smart meters as either an input or a parameter, it will be challenging to disentangle the effects of these multiple hypotheses and understand which one is driving the change in the observed end result.

When modelling each hypothesis in isolation, this is not a concern as only one variable could be changed at a time and so it is possible to be instantly certain which parameter is driving the change in the observed end result.

With the start small and simple approach, it is important to define the hypotheses very clearly and this collection of hypotheses then defines the scope of the analysis.

Each hypothesis can almost be considered a 'mini-model' in its own right and each one will have its own inputs, parameters, calculations and outputs. There will be common inputs, calculations etc. that underlie multiple hypotheses, however the exact combinations will be unique.

4 Scope

4.1 In Scope

To define the scope of the CR013 work, initial consideration took into account all areas of the balancing and Settlement regime that are impacted by the move to MHHS arrangements. It also considered Central Volume Allocation (CVA) consumption, although the focus has been on Supplier Volume Allocation (SVA) processes.

This initially broad approach was refined down to five hypotheses. This refinement was based on the areas where Participants anticipated the largest impacts and wanted to see modelled, in addition to restricting the scope only to those areas that could feasibly be modelled to a good level of accuracy and for which the data to undertake the analysis could be made available.

These five hypotheses will form the scope of the analysis and are listed below in descending priority order. This prioritisation has been informed by the areas where Participants predicted the most significant impact, in addition to the relative ease with which each hypothesis could be modelled and analysed.

More detail on each hypothesis, including the proposed methodology for how each is analysed, is provided in Appendix 2 (Hypotheses, Methodology for Analysis and Outputs) of this report.

Where 'Load Shapes' are referred to below, these are defined as a set of average consumption or export data for a categorisation of Metering System and are derived and provided by the Load Shaping Service (LSS). The Load Shapes are used by the Data Services and Market-wide Data Service (MDS) for calculation of UTC Settlement Period level data where UTC Settlement Period level data is unavailable, invalid or missing.

1. New technologies commercially enabled by MHHS will cause Load Shapes to become increasingly unrepresentative of the Traditional / NHH Smart / NHH Advanced population they are applied to.
2. The rules in place throughout Migration may result in market distortion, and therefore there may be opportunity for Suppliers to benefit commercially by modifying their Migration plan to benefit from the state of legacy and / or MHHS Settlement at that particular point in time.
3. Increase in metered Export sites from <50k to ~1million MPANs will result in reduced GCF, especially in summer.
4. Smart / Advanced meter Load Shapes are not representative of the Traditional / NHH Smart / NHH Advanced population they are applied to. This will require modelling separately for Domestic and Non-Domestic.
    i. Domestic
    ii. Non-Domestic
5. Bringing RF forward to 4 months will have a negative impact on overall Settlement accuracy.

4.2 Out of Scope

Conversely, there were scope areas that were considered, but that are not recommended for inclusion as part of the analysis. For completeness and in case the decision to exclude these areas is ever re-visited in future, the key out of scope areas have been included below along with an explanation for why each was omitted.

N.B. For the avoidance of doubt, the areas listed below are only a selection of the key out of scope areas. Any area that is not included within the preceding 'In Scope' section can be considered out of scope.

1. **LSS sample sizes are not sufficiently large on Day 1 of Migration and create unrepresentative Load Shapes**

   The LSS has a minimum sample size that needs to be achieved for each of the individual Load Shape Categories (LSCs). If this minimum sample size is not achieved (e.g. on Day 1 of Migration) or no data is available at all, the LSS will apply a default / back-stop Load Shape. The methodology through which these default / back-stop Load Shapes are calculated are covered in Sections 6 and 7 of the LSS Method Statement v5.0.

2. **Line Loss Factor (LLF)**

   No Participants expressed a desire to see this area further investigated or modelled when raised in the SASWG meetings.

3. **Distributed Use of System (DUoS) Billing**

   No Participants expressed a desire to see this area further investigated or modelled when raised in the SASWG meetings.

   It was agreed that should DNOs / iDNOs decide at a later date that they would like to see this area investigated, they can take the outputs of CR013, and potentially the subsequent analysis, and use these to progress any DUoS specific analysis internally.

4. **CVA Arrangements**

   No Participants raised specific standalone areas that would need to be considered as a result of MHHS impacts on CVA arrangements, that were not already covered through consideration for the SVA market.

5. **CR018 (Registration Service Operating Hours)**

   CR018 sought to reduce the operating hours of the Registration Service from the original scope of the MHHS TOM, and was approved. For more context on CR018, please visit the Programme Change Control section of the MHHS website.

   It was proposed that the approval of CR018 could have subsequent detrimental impacts upon Settlement, primarily through delayed appointment of the Smart Data Service (SDS). However following review, it was agreed that none of the impacts were material enough to justify modelling.

5 Timescales

Note: All timescales presented within this section are high-level, ballpark figures. Should the analysis go ahead, it is recommended that the analysis team undertake a more detailed exercise to define the analysis timescales in more detail.

Two options are presented below:

1. **Option 1 – Use Elexon Models**

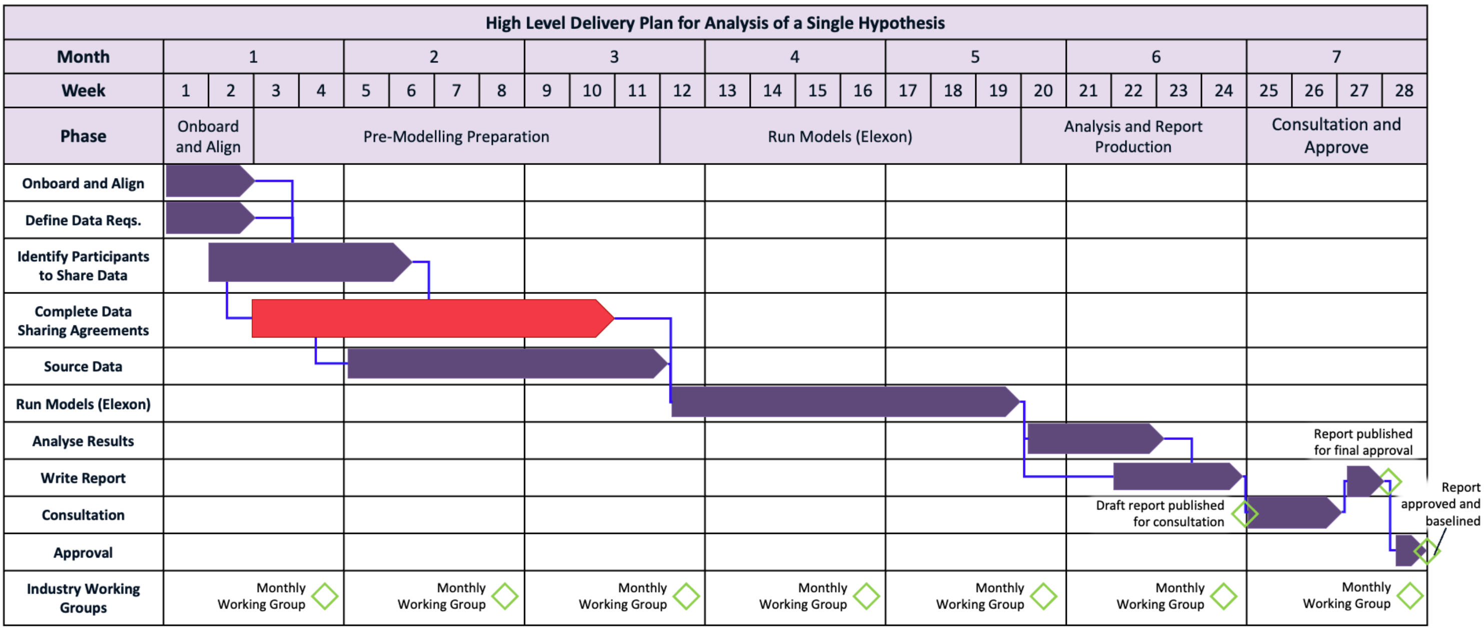
   - Longer timescales;

- The delivery plan to work to if Participants wanted to use more sophisticated models which would only be available through Elexon;
- In theory, should produce a more accurate and reliable output than Option 2;
- May delay analysis start as Elexon's LSS Simulator, for example, will not be available until earliest July 2024.

2. **Option 2 – Basic Spreadsheet Modelling**
   - Shorter timescales;
   - The delivery plan to work to if Participants wanted to prioritise time and cost savings and were willing to sacrifice, in theory, the accuracy and therefore value of the final output;
   - The modelling in this option does not include the use of any sophisticated models;
   - Modelling would rely solely on Excel spreadsheets and a small number of SQL databases;
   - Could start much sooner or as soon as CR, or another route of facilitation if analysis was to be progressed outside the MHHS Programme, is raised and approved and resource is available.

Option 2 is available and possible for all hypotheses (i.e. all hypotheses can feasibly be modelled through the basic approach and none of them are dependent on the availability of more sophisticated models). From our investigation, there would be no significant compromise on any of the hypotheses' modelling methodology by choosing Option 2.

5.1 Option 1 – Use Elexon Models

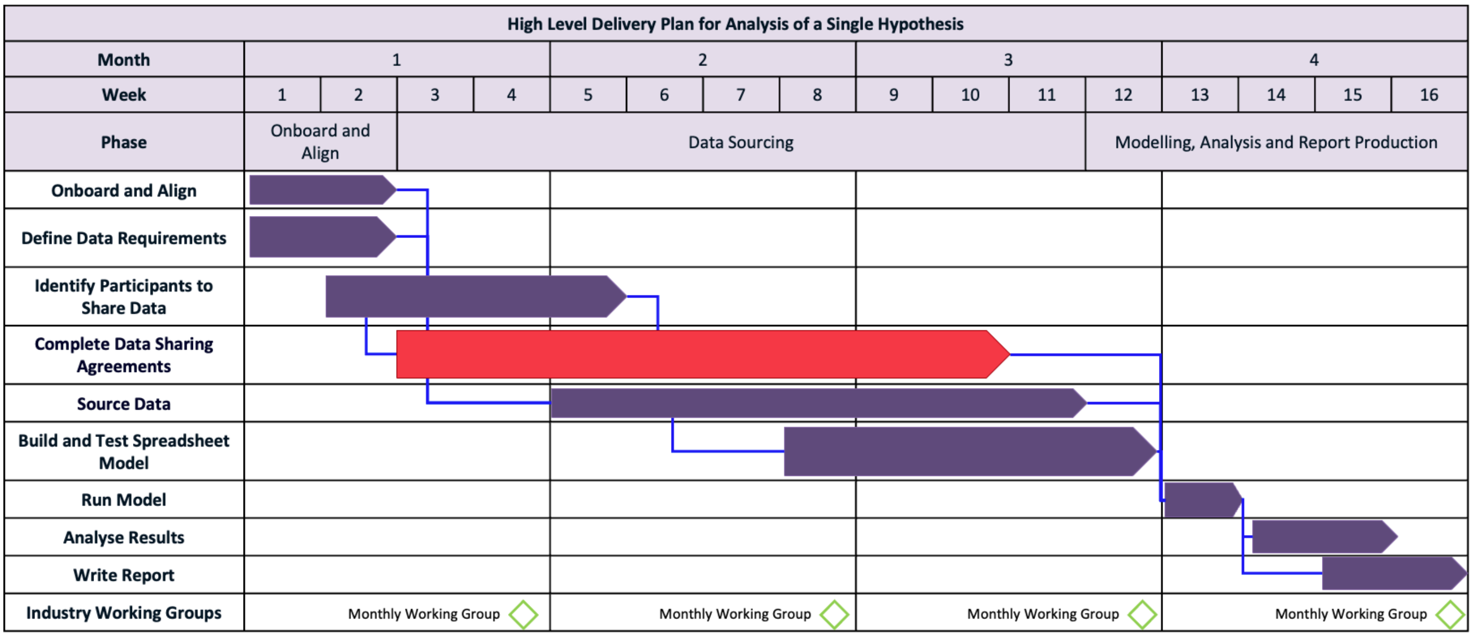


Notes:

1. This is the delivery plan for analysis of a single hypothesis. Analysis of multiple hypotheses will incur similar additional time, although there will be time efficiencies which can be taken advantage of.
2. No bespoke model development required as all required models are to be provided by Elexon.
3. 'Run Models' activity will be undertaken by Elexon. This is because, due to security reasons, Elexon cannot provide access to this model to external parties. Therefore, Elexon would need to receive data inputs and run the calculations themselves before providing outputs back to the analyst team and it is estimated, at a high-level, that this combined activity will take 8 weeks.

4. Assumed that data can be fed into the models without updates to the models being required.
5. Risk around 'Complete Data Sharing Agreements' activity which is highlighted with a red chevron. 8 weeks has been allowed for this compressed activity, but in previous projects this has taken 12 weeks and sometimes more. 8 weeks will only be possible if everything goes to plan and there are no significant legal complications on Participants' side.
6. If data is not provided by other parties in suitable format, to the correct completeness, quality, accuracy or in sufficient volume, this will delay timelines due to time required to correct. In most serious instances, may render analysis unachievable altogether if data issues are not able to be overcome.

5.2 Option 2 – Spreadsheet Style Modelling



Notes:

1. This is the delivery plan for analysis of a single hypothesis. Analysis of multiple hypotheses will incur similar additional time, although there will be time efficiencies which can be taken advantage of (although these savings will not be as significant as the equivalent savings for Option 1).
2. No bespoke complex model development required as this option relies solely on modelling through the use of Excel spreadsheets and a small number of SQL databases.
3. Risk around 'Complete Data Sharing Agreements' activity which is highlighted with a red chevron. 8 weeks has been allowed for this compressed activity, but in previous projects this has taken 12 weeks and sometimes more. 8 weeks will only be possible if everything goes to plan and there are no significant legal complications on Participants' side.
4. If data is not provided by other parties in suitable format, to the correct completeness, quality, accuracy or in sufficient volume, this will delay timelines due to time required to correct. In most serious instances, may render analysis unachievable altogether if data issues are not able to be overcome.
5. No consultation phase incorporated into final report production.

6 Costs

Note: All costs presented within this section are high-level, ballpark figures. Should the analysis go ahead, it is recommended that the analysis team undertake a more detailed exercise to define the analysis costs in more detail.

Two options are presented below:

1. **Option 1 – Use Elexon Models**
   - Higher costs;
   - The delivery plan to work to if Participants wanted to use more sophisticated models which would only be available through Elexon;
   - In theory, should produce a more accurate and reliable output than Option 2;
   - May delay analysis start as Elexon's LSS Simulator, for example, will not be available until earliest July 2024.

2. **Option 2 – Basic Spreadsheet Modelling**
   - Lower costs;
   - The delivery plan to work to if Participants wanted to prioritise time and cost savings and were willing to sacrifice, in theory, the accuracy and therefore value of the final output;
   - The modelling in this option does not include the use of any sophisticated models;
   - Modelling would rely solely on Excel spreadsheets and a small number of SQL databases;
   - Could start much sooner or as soon as CR, or another route of facilitation if analysis was to be progressed outside the MHHS Programme, is raised and approved and resource is available.

Option 2 is available and possible for all hypotheses (i.e. all hypotheses can feasibly be modelled through the basic approach and none of them are dependent on the availability of more sophisticated models). From our investigation, there would be no significant compromise on any of the hypotheses' modelling methodology by choosing Option 2.

6.1 Option 1 – Use Elexon Models

Hypothesis			Costs			
#	Description	Elexon Model	Elexon Costs[1]	Project Manager Costs[2]	Data Costs[3]	TOTAL
1	New technologies will skew Load Shapes	LSS Simulator	£60,000	£175,500	£45,000	£280,500
2	Market distortion throughout Migration	No relevant model – recommend Option 2	-	-	-	-
3	Increase in metered Export sites	SAA Test Environment	£60,000	£175,500	£22,500	£258,000
4i	Non-representative Smart / Advanced meter Load Shapes (Domestic)	LSS Simulator	£60,000	£175,500	£27,000	£262,500
4ii	Non-representative Smart / Advanced meter Load Shapes (Non-Domestic)	LSS Simulator	£60,000	£175,500	£40,500	£276,000
5	Bringing RF forward to 4 months	SAA Test Environment	£60,000	£175,500	£49,500	£285,000

Notes:

1. Elexon costs cover the receiving of data inputs from analyst team, feeding this data into the model, running of models on the analyst team's behalf and providing outputs back to the analyst team. Elexon are not able to simply provide models to the analyst team to run the analysis themselves due to security concerns with providing external access to test environments. These are high-level costs, more investigation would need to be carried out if the work went ahead to agree a more detailed cost, and its inclusion by no means signals a commitment to the presented cost.

   In the case of the LSS Simulator, whilst it is hoped that if and when the analysis starts, Elexon may be able to simply provide this model to the analyst team for use. Unfortunately this assumption cannot be currently validated as the simulator is still in development. If the simulator is not available, Elexon team would need to run the models on the analyst team's behalf in the actual LSS Test Environment. Therefore costs are the same as for Settlement Administration Agent (SAA) Test Environment.
2. Project Manager costs cover all activities apart from the running of the models and the extracting and providing of data to the analyst team. I.e. All other data sourcing work, all analysis and all report writing and approval.
3. Data Costs cover costs for Suppliers to undertake associated legal / compliance work, and then extract and provide the required data. Assumed, as a rough estimate, that 3x Suppliers provide data and each Supplier incurs a cost of £7,500 for first data request and an additional £1,500 for every subsequent request. These costs would need to be covered by Suppliers themselves and would not be reimbursed.
4. There will be time and cost efficiencies that can be taken advantage of if it is decided to investigate multiple hypotheses.
5. There are no analyst team-side legal costs included within this costing.

6.2 Option 2 – Spreadsheet Style Modelling

Hypothesis		Costs			
#	Description	Elexon Costs[1]	Analyst and PM Costs[2]	Data Costs[3]	TOTAL
1	New technologies will skew Load Shapes	£0	£113,000	£45,000	£158,000
2	Market distortion throughout Migration	£0	£113,000	£27,000	£140,000
3	Increase in metered Export sites	£0	£113,000	£22,500	£135,500
4i	Non-representative Smart / Advanced meter Load Shapes (Domestic)	£0	£113,000	£27,000	£140,000
4ii	Non-representative Smart / Advanced meter Load Shapes (Non-Domestic)	£0	£113,000	£40,500	£153,500
5	Bringing RF forward to 4 months	£0	£113,000	£49,500	£162,500

Notes:

1. Elexon costs are £0 as there are no external models to be provided or run by Elexon.
2. Analyst at 1.0 FTE, PM at 0.2 FTE. Analyst and PM costs combined covers the delivery of all activities contained within the Option 2 delivery plan. Additional PM resource included (which was

excluded from Option 1) as Analyst is already at capacity as will also have to build models (which they did not have to do in Option 1).

3. Data Costs cover costs for Suppliers to undertake associated legal / compliance work, and then extract and provide the required data. Assumed, as a rough estimate, that 3x Suppliers provide data and each Supplier incurs a cost of £7,500 for first data request and an additional £1,500 for every subsequent request. These costs would need to be covered by Suppliers themselves and would not be reimbursed.
4. There will be time and cost efficiencies that can be taken advantage of if it is decided to investigate multiple hypotheses (although these savings will not be as significant as they are for Option 1).
5. There are no analyst team-side legal costs included within this costing.

7 Options to Deliver the Analysis

7.1 Options

There are three different avenues in terms of next steps to deliver the analysis. These are presented below.

A recommendation in terms of a preferred next steps approach has not been provided as part of this report. Instead, it will be up to Participants to collectively agree the most appropriate route forward and bring this proposal back to the Programme, Elexon or other third party as appropriate.

7.1.1 Proceed with analysis under the MHHS Programme

Analysis is progressed centrally under the MHHS Programme. The work would likely be progressed through a new Working Group which Participants could attend and input into to shape the analysis and receive updates as it progresses. This would come at incremental cost which would ultimately be funded by Suppliers.

Advantages:

1. Should more than one Supplier agree to provide data, the Programme would have access to a larger, less biased dataset than if Participants progressed their own analysis internally.
2. If many Participants want to investigate further (and would do so internally if the analysis was not progressed centrally), it would be more effective and efficient overall to progress this centrally.
3. Undertaking the analysis centrally would mean that should any key insights emerge from the analysis, which may have consequences for the Programme and wider industry, this information would be made publicly available to Participants and it would be easier to coordinate centrally-led actions off the back of the work.

Disadvantages:

1. Additional central cost for Suppliers.
2. Participants would need to share personal data externally of their organisation with the Programme. This would require undertaking a Data Protection Impact Assessment (DPIA) and completion of data sharing agreements. This activity will incur significant additional time and cost.
3. The Programme is currently focused on delivery and may not have the capacity to manage this additional work.
4. Some of the hypotheses that are posed are enduring issues which will continue to impact and will require monitoring following the close of the MHHS Programme.
5. There will be many dependencies on Elexon to provide data, models and expertise.
6. Analysis less targeted to individual Supplier interests as scope would need to be centrally agreed by a majority.

7. Less targeted cost allocation. Suppliers who do not want to investigate further would shoulder additional cost.

7.1.2 Proceed with analysis under an Elexon Issues Group

Analysis is progressed centrally under an Elexon Issues Group which Participants could attend and input into to shape the analysis and receive updates as it progresses. This would come at incremental cost which would ultimately be funded by all Balancing and Settlement Code (BSC) funding parties.

Advantages:

1. Much of the required data, models and expertise is held within Elexon.
2. Some of the hypotheses that are posed are enduring issues which will continue to impact and will require monitoring following the close of the MHHS Programme.
3. SAA and LSS test environment access cannot be provided externally of Elexon. If analyst team were from Elexon, they could access these test environments directly rather than relying on a third party to run the modelling. This would potentially reduce timescales and cost.
4. Should more than one Supplier agree to provide data, Elexon would have access to a larger, less biased dataset than if Participants progressed their own analysis internally.
5. If many Participants want to investigate further (and would do so internally if the analysis was not progressed centrally), it would be more effective and efficient overall to progress this centrally.
6. Undertaking the analysis centrally would mean that should any key insights emerge from the analysis, which may have consequences for the Programme and wider industry, this information would be made publicly available to Participants and it would be easier to coordinate centrally-led actions off the back of the work.

Disadvantages:

1. Additional central cost for all BSC funding parties.
2. Participants would need to share personal data externally of their organisation with Elexon. This would require undertaking a DPIA and completion of data sharing agreements. This activity will incur significant additional time and cost.
3. Analysis less targeted to individual Supplier interests as scope would need to be centrally agreed by a majority.
4. Less targeted cost allocation. BSC funding parties who do not want to investigate further would shoulder additional cost.
5. Elexon resource constraints could lead to challenges in identifying correct resource or cause delay to analysis start.

7.1.3 Do not proceed with analysis undertaken by a Central Party

There would be no further action for any Central Party. All outputs from the CR013 work would be made available to Participants and it would then be up to Participants to progress their own analysis internally, if they so choose.

This could be approached in a number of different ways. A group of Participants may come together and seek out an independent party to take forward the analysis on their behalf, for example. Participants may then choose to socialise or not socialise key findings from their internal analyses centrally. If Participants wished to share outputs, the Programme and Elexon could facilitate publication.

Advantages:

1. No additional central cost for Suppliers or BSC funding parties.

2. Avoids complexity of Participants sharing personal data externally to their organisation which would save significant time and cost.
3. Enables more targeted analysis. Those Participants who choose to progress their own internal analysis can vary the analysis to answer the questions most pertinent to their organisation. They are not ‘bound’ to an analysis scope that would need to be centrally agreed.
4. More targeted cost allocation. Suppliers and BSC funding parties who do not want to investigate further do not need to shoulder additional cost.

Disadvantages:

1. If many Participants would choose to progress the analysis internally, it would be more effective and efficient overall to progress this centrally.
2. This option is reliant on Participants already having access to the required data or being able to seek it out.
3. The output of any internal analysis could be skewed as Participants would only be using a dataset from a single organisation, within which there may be bias.
4. Not all Participants may have access to sufficient internal resource or research funding to pursue the analysis. This approach could unintentionally favour the larger organisations who have larger pools of resource and funding to draw from.
5. It is at the discretion of individual Participants whether findings from internal analyses are socialised centrally and with other Participants. If not shared, this may result in key insights with significant consequences for the Programme and wider industry being missed.

8 Conclusion

Following agreement from the SASWG that the CR013 report is complete and ready to be brought to PSG for review, the report will be brought to the October PSG meeting for final review and approval.

In terms of next steps following PSG review, there are three different avenues to deliver the analysis as laid out in Section 7 (Options to Deliver the Analysis).

A recommendation in terms of a preferred next steps approach has not been provided as part of this report and it is not PSG’s role to issue a recommendation or decision in this respect either.

Instead, it will be up to Participants to collectively agree the most appropriate route forward and bring this proposal back to the Programme, Elexon or other third party as appropriate.

9 Appendix 1: Risks, Assumptions, Issues and Dependencies

9.1 Risks

1. There is a risk around the accuracy and representativeness of the final outputs of the analysis. There are many unknowns around how Settlement will work in MHHS and as such, any analysis will need to be based on a number of assumptions, in which there will be varying degrees of confidence. If the analysis is based on inaccurate assumptions, it will produce an inaccurate output, but it will be impossible to know whether or not this is the case until Migration starts.

   Furthermore, several of the hypotheses rely on testing through 'proxy' measures as, due to availability of data, it is not possible to test the actual hypothesis itself. The level of representativeness of these 'proxy' measures cannot be known for sure and would be hard to quantify, but will undoubtedly involve some degree of compromise on accuracy.

2. There is a risk around the accuracy of the costings and timescales that are presented in this report, and that when analysis delivery starts the actual costs and timescales diverge from those presented here. To date, the cost and time estimation has only been undertaken at a high-level and it is recommended that, should the analysis go ahead, the analysis team undertake a more detailed estimation exercise supported by the greater amount of information available at that time.

3. Personal data will need to be handled as part of the analysis. There is a risk that personal data is shared without the required data sharing agreements being in place, or that this data is mishandled (e.g. shared with wrong party).

4. There is a risk that no Suppliers agree to share the data that is required from Suppliers, perhaps due to commercial, data privacy or effort concerns, thus rendering the analysis unachievable.

5. There is a risk that the data that is provided by other parties is not provided in a suitable format, to the correct completeness, quality, accuracy, or in sufficient volume. If there are any minor issues with the data that is provided, this could incur additional time and cost as part of the analysis to correct these minor issues. If there are major issues, this could render the analysis unachievable altogether.

6. There is a risk that the analysis does not complete / findings are not made available to Programme Participants in sufficient time before the start of Migration. This could mean Participants do not have sufficient time to take any pre-Migration actions to reduce their level of risk exposure based on the findings from the analysis. A key objective of this work is helping Participants to enter Migration in a better prepared state. If the timing of the analysis does not allow for Participants to take any pre-Migration actions they deem necessary, this objective would be jeopardised.

7. There is a risk that there are important MHHS Settlement impacts that have not been scoped as part of the analysis. With so many unknowns involved in MHHS Settlement, it is possible that impacts will only emerge during Migration that it would have been impossible or very difficult to predict beforehand.

9.2 Assumptions

1. All expertise required from Industry as part of the analysis is provided by those parties and not recharged to the Programme, in accordance with their BSC obligations to support the Programme.
2. Industry (primarily Suppliers and Elexon) will provide the input data listed in Appendix 3 (Data Required and Data Sources) to make analysis of the hypotheses possible. If this data is not provided, it will not be possible to test the relevant hypotheses.
3. No bespoke complex models are required to be developed. As it is assumed, that if an Option 1 (Use Elexon Models) approach is preferred, Elexon will provide or provide access to all models. And if an Option 2 (Spreadsheet Style Modelling) approach is preferred, modelling will only take place across

Excel spreadsheets and a small number of SQL databases. If bespoke complex models were required, this would significantly increase both cost and time estimates.

4. If an Option 1 (Use Elexon Models) approach is preferred, no updates to Elexon-provided models are required to enable the models to ingest the data and run the required scenarios. If updates were required, this would need to go through the usual Elexon process of a new work request and would significantly increase both cost and time estimates.

9.3 Issues

1. The LSS simulator, which is the required calculation engine for several of these hypotheses if an Option 1 (Use Elexon Models) approach is preferred, will not be available through Elexon until July 2024.
2. Elexon's SAA and LSS test environments are the required calculation engines if an Option 1 (Use Elexon Models) approach is preferred. However, Elexon are unable to provide external access to these test environments for security reasons. This means to test these hypotheses, Elexon would need to receive the relevant data inputs and run the calculations on the analyst team's behalf.

9.4 Dependencies

1. There is a dependency on Subject Matter Experts (SMEs) across Industry to provide their knowledge and feedback to the analysis team to shape and progress the analysis. Particularly, but not limited to, Programme market analysts and Elexon market experts.
2. If the preferred modelling approach is Option 1 (Use Elexon Models), there is a dependency on Elexon to receive the data inputs from the analyst team, model various scenarios in their SAA / LSS test environment and provide the outputs back to the analyst team.

10 Appendix 2: Hypotheses, Methodology for Analysis and Outputs

The hypotheses are presented in this section in descending priority order. N.B. The prioritisation order does not necessarily inform the order in which the hypotheses are potentially analysed.

Each hypothesis is presented alongside the recommended step-by-step proposed methodology for analysis, in addition to the form the final output of the analysis should take.

The methodologies outlined below are suggestions of possible ways to investigate each of the hypotheses. They are not a directive and more information may emerge when the analysis begins which means other, different methodologies become more viable approaches.

10.1 Impact of New Technologies on Load Shapes

Hypothesis Title: New technologies commercially enabled by MHHS will cause Load Shapes to become increasingly unrepresentative of the Traditional / NHH Smart / NHH Advanced population they are applied to.

Methodology:

1. Take ‘LSS Smart meter sample consumption data’. Allocate each MPAN to the 56x Smart LSCs as defined in Appendix 1 “ISD Table: Load Shapes Categories” of the LSS Method Statement v5.0, using a combination of Market Segment Indicator, Grid Supply Point (GSP) Group ID, Domestic Premises Indicator, Measurement Quantity and Connection Type Indicator.
2. Take ‘LSS Advanced meter sample consumption data’. Allocate each MPAN to the 8x Advanced LSCs as defined in Appendix 1 “ISD Table: Load Shapes Categories” of the LSS Method Statement v5.0, in the same way just done for the Smart data.
3. For each Settlement Period level consumption data item, convert these values into both Settlement Day and Settlement Year coefficients by dividing by either the total yearly or total daily consumption for that MPAN.
4. Across each of the 64x LSCs (Smart and Advanced), calculate the average of the Settlement Day and Settlement Year coefficients, for each Settlement Period, from all MPANs that fall within that specific LSC. This will provide the average daily and yearly Load Shape for each of the LSCs.
5. Repeat Steps 1-4 using the same ‘LSS Smart meter sample consumption data’ and ‘LSS Advanced meter sample consumption data’. However, this time incorporate varying numbers of heat pump, EV and new half-hourly (HH) ToU tariff customers into the samples.
6. For information on the consumption data and how this is distributed across Settlement Periods, refer to the ‘Heat pump / EV / new HH ToU tariff customer sample consumption data’ provided by Suppliers.
7. For information on which LSC, the heat pump / EV / new HH ToU tariff MPANs fall into, refer to the Market Segment Indicator, GSP Group ID, Domestic Premises Indicator, Measurement Quantity and Connection Type Indicator for each MPAN.
8. For information on the number of these customers to introduce into the sample, refer to the forecasts of these types of customers that are to be provided by Supplier and DESNZ / Ofgem forecasts.
9. Once Steps 1-4 have been repeated for this new sample, this will produce an average daily and yearly Load Shape for the 64x LSCs for a sample which includes varying numbers of heat pumps, EV and new HH ToU tariff customers.
10. Compare the ‘new technology Load Shapes’ to the standard Load Shapes, ensuring only equivalent LSCs are compared, to understand to what extent the introduction of the new technology customers cause divergence from a standard Load Shape.

Output:

1. ‘Standard Load Shapes’ for a Settlement Day and Settlement Year for 64x LSCs.

2. ‘New Technology Load Shapes’ for a Settlement Day and Settlement Year for 64x LSCs.

10.2 Market Distortion throughout Migration

Hypothesis Title: The rules in place throughout Migration may result in market distortion, and therefore there may be opportunity for Suppliers to benefit commercially by modifying their Migration plan to benefit from the state of legacy and / or MHHS Settlement at that particular point in time.

Methodology:

1. Take ‘LSS Smart meter sample consumption data’. For each MPAN, sum the daily consumption volume between the peak hours of 16:00-19:00. Divide this peak consumption volume by the total daily consumption to produce a % of peak hours consumption for 365 days. Take an average of the % of peak hour consumption across the 365 days to produce one annual % of peak hours consumption for each MPAN.
2. Divide the total Smart sample up by the annual % of peak hours consumption so there are 10x evenly sized categories.
3. For every Settlement Period level consumption data item, convert these values into Settlement Day coefficients by dividing by the total daily consumption for that MPAN.
4. Within each of the % of peak hours consumption categories, map each MPAN to its respective Profile Class. Then compare the Elexon Profile Coefficients for that Profile Class to the coefficients calculated through the sample data.
5. Sum up the differences for every MPAN within one of the % of peak hours consumption categories and this provides the commercial opportunity, across a single day, to migrate or not migrate that % of peak hours consumption category, expressed as a coefficient.
6. To turn the coefficient into an estimated kilowatt hour (kWh) figure, multiply by the average total daily consumption for that % of peak hours consumption category.
7. Finally, this number will need to be scaled depending on the number of Smart MPANs in the % of peak hours consumption category sample and the number of Smart MPANs in the Supplier’s portfolio that is being modelled that also fall into the same % of peak hours consumption category. This will provide the commercial opportunity, across a single day, to migrate or not migrate this % of peak hours consumption category of the Smart segment of a Supplier’s portfolio.
8. Repeat Steps 4-6 for the other nine % of peak hours consumption categories. Once scaled for the number of Smart MPANs in the Supplier’s portfolio that is being modelled that also fall into the same % of peak hours consumption category, sum the 10 totals together to provide the commercial opportunity, across a single day, to migrate or not migrate the entire Smart segment of a Supplier’s portfolio.
9. There will also be variance, although admittedly not as large in the legacy vs MHHS Settlement of Traditional and Advanced meters without working communications. The two segments will be calculated separately. If trying to keep the analysis simple, recommendation would be to focus on the Smart segment.
10. Firstly, the Traditional segment. Take ‘NHH segment of portfolio consumption data’. Map all Traditional MPANs to Profile Class and LSC, using Market Segment Indicator, GSP Group ID, Domestic Premises Indicator, Measurement Quantity and Connection Type Indicator.
11. Calculate the coefficients across a Settlement Day for each of the relevant LSCs using the ‘LSS Smart meter sample consumption data’. These values represent the Load Shapes these Traditional MPANs would be settled to under MHHS.
12. Then compare these calculated coefficients vs the Elexon Profile Coefficients for the respective Profile Class. Sum up the coefficient differences and this provides the commercial opportunity, across a single day, to migrate or not migrate this segment of Traditional MPANs expressed as a coefficient.

13. Then multiply by the average total daily consumption for the Traditional MPAN sample to get a kWh estimate, and finally scale this number based on the number of Traditional MPANs in the sample and the number of Traditional MPANs in the Supplier's portfolio that is being modelled. This will provide the commercial opportunity, across a single day, to migrate or not migrate the Traditional segment of a Supplier's portfolio.
14. For the Advanced meters without working comms, repeat Steps 10-13. This will provide the commercial opportunity, across a single day, to migrate or not migrate the Advanced meters without working comms segment of a Supplier's portfolio.
15. Sum the outputs of Steps 7, 13 and 14 together to provide the total commercial opportunity, across a single day, to migrate or not migrate these segments of a Supplier's portfolio.

Output:

1. The commercial opportunity, across a single day and expressed in kWh, to migrate or not migrate the Smart segment of a Supplier's portfolio. This commercial opportunity will be split by the 10 different % of peak hours consumption categories, as the commercial benefit will vary depending on the category.
2. The commercial opportunity, across a single day and expressed in kWh, to migrate or not migrate the Traditional segment of a Supplier's portfolio.
3. The commercial opportunity, across a single day and expressed in kWh, to migrate or not migrate the Advanced meters without working comms segment of a Supplier's portfolio.

10.3 Increase in Metered Export Sites

Hypothesis Title: Increase in metered Export sites from <50k to ~1million MPANs will result in reduced GCF, especially in summer.

1. Take 'Unmeasured Export sites sample consumption data'. Filter the data by GSP Group ID and average the Settlement Period level consumption data for each GSP Group. This will produce 17,520 (365 x 48) individual data values per GSP Group. This is the export consumption, by Settlement Period, for an average unmeasured Export site in each GSP Group.
2. Take 'Count of currently unmeasured Export sites by GSP Group' from the Feed-in Tariff (FIT) Register. Multiply the count for each GSP Group by the Settlement Period level export consumption for an average unmeasured Export in the same GSP Group. This will produce the estimated total export consumption, by Settlement Period, for all currently unmeasured Export sites in each GSP Group.
3. Divide the total unmeasured export consumption by the 'GSP Group Take by GSP Group' for the relevant GSP Group. Ensure GSP Groups and Settlement Periods are matched, and all data is converted into one of either UTC or Clock Time. This will produce the estimated impact, at a Settlement Period level, on the GSP Group Take by GSP Group if all currently unmeasured Export sites were to be metered. From this, it is possible to derive the GCF which would have previously been required to 'correct' this unmeasured export volume and ensure the total measured volumes match to the GSP Group Take.
4. Finally, compare to previous 'GCF by GSP Group and Settlement Period' figures, ensuring to match up relevant GSP Groups and Settlement Periods. Subtract the GCF which was calculated at the end of Step #3 (the GCF previously required to 'correct' the unmeasured export volume but that is no longer required once the Export sites are metered) to understand what the GCF would be once all Export sites are metered. Compare the final two sets of GCF figures to understand whether GCF would increase or decrease as a result of metering all Export sites.

Output:

1. The estimated impact, at a Settlement Period level, on the GSP Group Take by GSP Group if all currently unmeasured Export sites were to be metered.

2. The GCF required, by Settlement Period and GSP Group, to match total measured volumes to the GSP Group Take once all Export sites are metered.
3. Can then subtract Output #2 from previous 'GCF by GSP Group and Settlement Period' figures and compare the resulting final two sets of GCF figures to understand whether GCF would increase or decrease as a result of metering all Export sites.

10.4 Non-representative Smart / Advanced meter Load Shapes

10.4.1 Domestic

Hypothesis Title: Smart / Advanced meter Load Shapes are not representative of the Traditional / NHH Smart / NHH Advanced population they are applied to – Domestic

1. Due to a lack of availability of HH consumption data for NHH settled meters, the closest way to model this hypothesis is through proxy by comparing the Load Shapes of Unrestricted tariffs (Profile Class 1 (PC1)) vs Time of Use (ToU) tariffs (PC2).
2. Take 'LSS Smart meter sample consumption data'. Use Domestic Premises indicator to only select the Domestic MPANs. The sample will now only consist of PC1 and PC2 meters.
3. Map each MPAN in the sample to a LSC using Market Segment Indicator, GSP Group ID, Domestic Premises Indicator, Measurement Quantity and Connection Type Indicator.
4. For each Settlement Period level consumption data item, convert these values into both Settlement Day and Settlement Year coefficients by dividing by either the total yearly or total daily consumption for that MPAN.
5. Within each LSC, average the coefficients of all MPANs within that LSC for each of the Settlement Periods. This will produce the daily and yearly Load Shape for each LSC.
6. Then, referring back to the blended PC1 and PC2 sample, average the coefficients across all LSCs for each Settlement Period. This will produce one average daily and yearly Load Shape for the blended PC1 and PC2 sample.
7. Compare this average daily and yearly Load Shape vs the equivalent Elexon Profile Coefficients for PC1 and PC2. The average daily and yearly Load Shape should resemble a middle ground of the PC1 and PC2 Profile Shapes. They won't be as 'peaky' as the PC1 shape, but they won't be as 'flat' as the PC2 shape. By comparing these different profiles, it is possible to identify how unrepresentative the Smart meter Load Shapes will be for non-Smart PC1 meters. This will lead to commercial benefit for non-Smart PC1 meters as they are settled to a 'flatter' Load Profile.

Output:

1. Average daily and yearly Load Shapes for a blended PC1 and PC2 sample. These can be compared to the PC1 and PC2 Profile Shapes to identify how unrepresentative the Smart meter Load Shapes will be for non-Smart PC1 meters. This will lead to commercial benefit for non-Smart PC1 meters as they are settled to a 'flatter' Load Profile.

10.4.2 Non-Domestic

Hypothesis Title: Smart / Advanced meter Load Shapes are not representative of the Traditional / NHH Smart / NHH Advanced population they are applied to – Non-Domestic

1. Similarly to the Domestic part of this analysis, due to a lack of availability of HH consumption data for NHH settled meters, the closest way to model this hypothesis is through proxy. However, the availability of Standard Industrial Classification (SIC) codes, which categorise premises type, for the non-Domestic sector does allow for more detailed analysis if the three following assumptions are applied:

i. Same premises types (as defined by SIC codes) use electricity in similar profiles across the day.
ii. Smart and Advanced meters with working communications use electricity in similar profiles across the day.
iii. Traditional and Advanced meters without working communications use electricity in similar profiles across the day.

2. Take 'Count and consumption data of non-Domestic SIC codes'. Separate this data into Smart and Advanced meters with working communications, and Traditional and Advanced meters without working communications.
3. For each Settlement Period level consumption data item, convert these values into both Settlement Day and Settlement Year coefficients by dividing by either the total yearly or total daily consumption for that MPAN.
4. Allocate each MPAN to the 36x non-Domestic Smart and Advanced LSCs using Market Segment Indicator, GSP Group ID, Domestic Premises Indicator, Measurement Quantity and Connection Type Indicator.
5. Use the Smart and Advanced meters with working communications dataset and from this, calculate the 36x non-Domestic daily and yearly Load Shapes by averaging the coefficient for each Settlement Period for all MPANs that fall into each LSC.
6. Return to the original Smart and Advanced meters with working communications dataset and note the distribution of SIC codes. For each SIC code, calculate the average Settlement Day and Settlement Year coefficients by averaging the coefficients for each Settlement Period for all MPANs within a SIC code.
7. Turn to the Traditional and Advanced meters without working communications dataset and note the distribution of SIC codes. Use the average coefficients by SIC code, derived from the Smart and Advanced meters with working communications dataset, and apply these to the Traditional and Advanced meters without working communications dataset using the SIC code information. This will provide an estimate of the HH pattern of usage of these meters.
8. Compare the coefficients for the Traditional and Advanced meters without working communications dataset with the corresponding daily and yearly Load Shapes, ensuring that LSCs are mapped correctly.
9. Multiply the daily or yearly difference in the two sets of coefficients by the average total daily or yearly consumption of the Traditional and Advanced meters without working communications dataset to calculate the variance in kWh per MPAN. This number can then be scaled up depending on how many MPANs are being modelled.

Output:

1. 36x non-Domestic daily and yearly Load Shapes for the MPAN sample.
2. Estimate of the HH usage of the Traditional and Advanced meters without working communications dataset, using SIC code information.
3. Output #1 can be compared with Output #2 to understand how representative or not the non-Domestic Load Shapes are of the Traditional and Advanced meters without working communications population.

10.5 Bringing RF Forward to Four Months

Hypothesis Title: Bringing RF forward to 4 months will have a negative impact on overall Settlement accuracy.

1. This will be approached separately for Traditional, Smart without working communications and Advanced without working communications.

2. Take ‘Forecast of % estimated consumption per Traditional MPAN at 4 months’. Multiply this percentage by the ‘Count of Traditional meters across UK market’ to produce an estimate for the volume of Traditional estimated consumption at 4 months across the UK market.
3. Repeat Step 2 separately for Smart without working communications and Advanced without working communications. The count of Smart and Advanced without working communications will need to be estimated by applying the ‘% Smart / Advanced meters with working communications’ data provided by Suppliers to the ‘Count of Smart / Advanced meters across UK market’ data extracted from the Electricity Enquiry Service (EES).
4. Take ‘Estimated Annual Consumptions (EACs) vs Annualised Advances (AAs)’. Split this sample into Traditional, Smart and Advanced meters.
5. For each MPAN, calculate the ratio of most recent AA pro-rata’d for 1x day and corresponding EAC (which relates to same period as most recent AA) pro rata’d for 1x day.
6. Calculate the average of this ratio across the three separate Traditional, Smart and Advanced samples. This provides a view of how accurate previous estimates have been vs actual reads.
7. Multiply the volume of Traditional / Smart without working communications / Advanced estimated consumption at 4 months by the EAC vs AA ratio for the respective meter type. This will provide the volume of likely error in the estimated consumption.
8. Summing up these final three numbers will produce the estimated daily Settlement accuracy error, in kWh, across all Traditional, Smart without working communications and Advanced without working communications meters.
9. Divide this number by the total daily electricity settled by Elexon to express the Settlement accuracy as a percentage.
10. Compare this percentage to publicly available Elexon figures on Settlement accuracy when RF is at 16 months to understand whether Settlement accuracy would increase or decrease when RF is brought forward to 4 months.

Output:

1. Estimated daily Settlement accuracy error, in kWh, across all Traditional, Smart without working communications and Advanced without working communications meters.
2. Settlement accuracy estimate, represented as a percentage, for when RF is brought forward to 4 months.
3. Output #2 can then be compared to publicly available Elexon figures on Settlement accuracy when RF is at 16 months to understand whether Settlement accuracy would increase or decrease when RF is brought forward to 4 months.

11 Appendix 3: Data Required and Data Sources

There is an assumption that Industry (primarily Suppliers and Elexon) will provide the input data listed below in this section to make analysis of the hypotheses possible. If this data is not provided, it will not be possible to test the relevant hypotheses.

Furthermore, one Supplier alone providing data would make analysis of the relevant hypothesis possible, however the more Suppliers who provide data, the larger and more balanced the dataset the analyst team will have to work with. This, in turn, will improve the reliability and accuracy of the final output of the analysis.

All sample sizes are for guidance only. They have been selected to provide samples that are sufficiently large to have comfort in the averages, but there is limited science behind them beyond this. It is possible that the requirement to provide 12 months of data significantly reduces the available sample size for specific segments of customers e.g. EV customers where there will be larger sample sizes available for shorter, more recent periods. If this is the case, the analyst team will make a decision on whether or not to reduce the period of data requested based on the sample size variance.

All data that is provided should date from within the last 12 months.

11.1 Suppliers

The table below presents all potential data that could be requested from Suppliers for the purpose of analysing the hypotheses. The table also shows the mapping between the requested data and the hypothesis / hypotheses that they are related to. The numbering of the hypotheses refers to the hypotheses in the order that they are presented in Section 4 (Scope) of this document.

All Supplier-provided data that is required for a given hypothesis should ideally be provided by the same Supplier(s), focusing on the same MPAN sample, and covering the same time period.

#	Required Data – Suppliers	Hypothesis					
		1	2	3	4i	4ii	5
11.1.1	LSS Smart meter sample consumption data	X	X		X	X	X
11.1.2	LSS Advanced meter sample consumption data	X				X	
11.1.3	Heat pump customer sample consumption data	X					
11.1.4	EV customer sample consumption data	X					
11.1.5	New HH ToU tariff customer sample consumption data	X					
11.1.6	Forecasted number of customers on new HH ToU tariffs	X					
11.1.7	NHH segment of portfolio consumption data		X				
11.1.8	Unmeasured Export sites sample consumption data			X			
11.1.9	% Smart meters providing HH data				X	X	
11.1.10	% Advanced meters providing HH data					X	
11.1.11	Count and consumption data of non-Domestic SIC codes					X	
11.1.12	Forecast of % estimated consumption per Traditional MPAN at 4 months						X
11.1.13	Forecast of % estimated consumption per Smart MPAN without working communications at 4 months						X
11.1.14	Forecast of % estimated consumption per Advanced MPAN without working communications at 4 months						X
11.1.15	EACs vs AAs						X
11.1.16	% Smart meters with working communications						X
11.1.17	% Advanced meters with working communications						X

11.1.1 LSS Smart meter sample consumption data

Minimum required sample size: 10,000 HH Smart MPANs, where sample is representative of the Supplier's full Smart portfolio (i.e. sample is free from bias). Require a minimum of 500 MPANs for Domestic and non-Domestic segments in above sample.

1. Consumption data by Settlement Period for 12 months (i.e. 17,520 (365x48) individual consumption values for each MPAN). All data to be provided in UTC time.
2. Anonymised MPAN (require a random ID for each customer)
3. Market Segment Indicator
4. GSP Group ID
5. Domestic Premises Indicator flag
6. Measurement Quantity – Only Active Import and Active Export, no Reactive Import / Export to be included
7. Profile Class
8. Connection Type Indicator
9. All data needs to pass validation rules as applied by SDS

Analysis team will then turn this data into 'Proxy' MHHS Load Shapes by allocating to the 56 different Smart LSCs as defined in Appendix 1 of the LSS Method Statement v5.0. Suppliers do not need to undertake this exercise themselves.

11.1.2 LSS Advanced meter sample consumption data

Minimum required sample size: 2,000 HH Advanced MPANs, where sample is representative of the Supplier's full Advanced portfolio (i.e. sample is free from bias). Require a minimum of 300 MPANs for Domestic and non-Domestic segments in above sample.

1. Consumption data by Settlement Period for 12 months (i.e. 17,520 (365x48) individual consumption values for each MPAN). All data to be provided in UTC time.
2. Anonymised MPAN (require a random ID for each customer)
3. Market Segment Indicator
4. Domestic Premises Indicator flag
5. Measurement Quantity – Only Active Import and Active Export, no Reactive Import / Export to be included
6. Connection Type Indicator
7. All data needs to pass validation rules as applied by ADS

Analysis team will then turn this data into 'Proxy' MHHS Load Shapes by allocating to the 8 different Advanced LSCs as defined in Appendix 1 of the LSS Method Statement v5.0. Suppliers do not need to undertake this exercise themselves.

11.1.3 Heat pump customer sample consumption data

Minimum required sample size: 2,000 heat pump customer MPANs, where sample is representative of heat pump customers across the rest of the Supplier's portfolio (i.e. sample is free from bias). Require a minimum of 300 MPANs for Domestic and non-Domestic segments in above sample.

1. Consumption data by Settlement Period for 12 months (i.e. 17,520 (365x48) individual consumption values for each MPAN). All data to be provided in UTC time.
2. Anonymised MPAN (require a random ID for each customer)
3. Market Segment Indicator
4. GSP Group ID (desirable)
5. Domestic Premises Indicator flag
6. Measurement Quantity – Only Active Import and Active Export, no Reactive Import / Export to be included
7. All data needs to pass validation rules as applied by SDS / ADS

11.1.4 EV customer sample consumption data

Minimum required sample size: 2,000 EV customer MPANs, where sample is representative of EV customers across the rest of the Supplier's portfolio (i.e. sample is free from bias). Require a minimum of 300 MPANs for Domestic and non-Domestic segments in above sample.

1. Consumption data by Settlement Period for 12 months (i.e. 17,520 (365x48) individual consumption values for each MPAN). All data to be provided in UTC time.
2. Anonymised MPAN (require a random ID for each customer)
3. Market Segment Indicator
4. GSP Group ID (desirable)
5. Domestic Premises Indicator flag
6. Measurement Quantity – Only Active Import and Active Export, no Reactive Import / Export to be included
7. All data needs to pass validation rules as applied by SDS / ADS

11.1.5 New HH ToU tariff customer sample consumption data

Minimum required sample size: 2,000 new HH ToU tariff customer MPANs, where sample is representative of new HH ToU tariff customers across the rest of the Supplier's portfolio (i.e. sample is free from bias). Require a minimum of 300 MPANs for Domestic and non-Domestic segments in above sample.

1. Consumption data by Settlement Period for 12 months (i.e. 17,520 (365x48) individual consumption values for each MPAN). All data to be provided in UTC time.
2. Anonymised MPAN (require a random ID for each customer)
3. Market Segment Indicator
4. GSP Group ID (desirable)
5. Domestic Premises Indicator flag
6. Measurement Quantity – Only Active Import and Active Export, no Reactive Import / Export to be included
7. All data needs to pass validation rules as applied by SDS / ADS

11.1.6 Forecasted number of customers on new HH ToU tariffs

1. Forecasted number of customers on new HH ToU tariffs across the next 10 years in quarterly intervals

2. Forecast split by Domestic and non-Domestic
3. Forecast split by GSP Group (desirable)

11.1.7 NHH segment of portfolio consumption data

Minimum required sample size: 10,000 NHH MPANs, where sample is representative of the Supplier's full NHH portfolio (i.e. sample is free from bias).

1. Annual average of the estimated daily consumption data for each MPAN. All data to be provided in UTC time.
2. Anonymised MPAN (require a random ID for each customer)
3. Market Segment Indicator
4. GSP Group ID
5. Domestic Premises Indicator flag
6. Measurement Quantity – Only Active Import and Active Export, no Reactive Import / Export to be included
7. Profile Class

11.1.8 Unmeasured Export sites sample consumption data

Minimum required sample size: 2,000 Export MPANs, where sample is representative of unmeasured Export consumption across the rest of the Supplier's portfolio (i.e. sample is free from bias), and contains as even a distribution of GSP Group IDs as possible.

1. Consumption data by Settlement Period for 12 months (i.e. 17,520 (365x48) individual consumption values for each MPAN). All data to be provided in UTC time.
2. GSP Group ID

11.1.9 % Smart meters providing HH data

Minimum required sample size: 10,000 Smart MPANs for Domestic average, where sample is representative of the Supplier's full Domestic Smart portfolio (i.e. sample is free from bias).

500 Smart MPANs for non-Domestic average, where sample is representative of the Supplier's full non-Domestic Smart portfolio (i.e. sample is free from bias).

1. 1x Average % value for Domestic Smart meters providing HH data, and count of how many MPANs in sample.
2. 1x Average % value for non-Domestic Smart meters providing HH data, and count of how many MPANs in sample.

11.1.10 % Advanced meters providing HH data

Minimum required sample size: 2,000 Advanced MPANs for non-Domestic average, where sample is representative of the Supplier's full non-Domestic Advanced portfolio (i.e. sample is free from bias).

500 Advanced MPANs for Domestic average, where sample is representative of the Supplier's full Domestic Advanced portfolio (i.e. sample is free from bias).

1. 1x Average % value for Domestic Advanced meters providing HH data, and count of how many MPANs in sample.

2. 1x Average % value for non-Domestic Advanced meters providing HH data, and count of how many MPANs in sample.

11.1.11 Count and consumption data of non-Domestic SIC codes

Minimum required sample size: 10,000 non-Domestic MPANs, where sample is representative of the Supplier's full non-Domestic portfolio (i.e. sample is free from bias). All sub-section level non-Domestic SIC codes should be represented in the sample.

1. Consumption data by Settlement Period for 12 months where Settlement Period level data is available. If not available, see #2. I.e. 17,520 (365x48) individual consumption values for each MPAN where Settlement Period level data is available. All data to be provided in UTC time.
2. Where data not available at a Settlement Period level, provide monthly average of the estimated daily consumption data for 12 months of the year (i.e. 12 (12x1) individual consumption values for each MPAN where Settlement Period level data is not available).
3. Sub-section level non-Domestic SIC codes
4. Anonymised MPAN (require a random ID for each customer)
5. Market Segment Indicator
6. Meter Type

11.1.12 Forecast of % estimated consumption per Traditional MPAN at 4 months

Minimum required sample size: 1x Average percentage where value has been calculated from a minimum sample size of 2,000 Traditional MPANs which are representative of the Supplier's full Traditional portfolio (i.e. sample is free from bias).

1. Forecast for the average % of estimated consumption (as opposed to actual consumption) per Traditional MPAN at the end of the new 4 month RF window.

11.1.13 Forecast of % estimated consumption per Smart MPAN without working communications at 4 months

Minimum required sample size: 1x Average percentage where value has been calculated from a minimum sample size of 1,000 Smart without working communications MPANs which are representative of the Supplier's full Smart without working communications portfolio (i.e. sample is free from bias).

1. Forecast for the average % of estimated consumption (as opposed to actual consumption) per Smart MPAN without working communications at the end of the new 4 month RF window.

11.1.14 Forecast of % estimated consumption per Advanced MPAN without working communications at 4 months

Minimum required sample size: 1x Average percentage where value has been calculated from a minimum sample size of 1,000 Advanced without working communications MPANs which are representative of the Supplier's full Advanced without working communications portfolio (i.e. sample is free from bias).

1. Forecast for the average % of estimated consumption (as opposed to actual consumption) per Advanced MPAN without working communications at the end of the new 4 month RF window.

11.1.15 EACs vs AAs

Minimum required sample size: 5,000 NHH MPANs, where sample is representative of the Supplier's full NHH portfolio (i.e. sample is free from bias).

1. Most recent AA consumption pro-rata'd for 1x average day. To reduce volatility, AA total length must cover a period greater than 2 months. If most recent AA is not greater than 2 months, take next most recent AA which covers a period greater than 2 months.
2. Corresponding EAC (which related to the same period as the most recent AA) consumption pro-rata'd for 1x average day
3. Anonymised MPAN (require a random ID for each customer)
4. Market Segment Indicator
5. Profile Class
6. Connection Type Indicator
7. Domestic Premises Indicator flag

11.1.16 % Smart meters with working communications

Minimum required sample size: 10,000 Smart MPANs for Domestic average, where sample is representative of the Supplier's full Domestic Smart portfolio (i.e. sample is free from bias).

500 Smart MPANs for non-Domestic average, where sample is representative of the Supplier's full non-Domestic Smart portfolio (i.e. sample is free from bias).

1. 1x Average % value for Domestic Smart meters with working communications, and count of how many MPANs in sample.
2. 1x Average % value for non-Domestic Smart meters with working communications, and count of how many MPANs in sample.

11.1.17 % Advanced meters with working communications

Minimum required sample size: 2,000 Advanced MPANs for non-Domestic average, where sample is representative of the Supplier's full non-Domestic Advanced portfolio (i.e. sample is free from bias).

500 Advanced MPANs for Domestic average, where sample is representative of the Supplier's full Domestic Advanced portfolio (i.e. sample is free from bias).

1. 1x Average % value for Domestic Advanced meters with working communications, and count of how many MPANs in sample.
2. 1x Average % value for non-Domestic Advanced meters with working communications, and count of how many MPANs in sample.

11.2 Elexon

The table below presents all potential data that could be requested from Elexon for the purpose of analysing the hypotheses. The table also shows the mapping between the requested data and the hypothesis / hypotheses that they are related to. The numbering of the hypotheses refers to the hypotheses in the order that they are presented in Section 4 (Scope) of this document.

All Elexon-provided data that is required for a given hypothesis should ideally be provided focusing on the same MPAN sample and covering the same time period.

All data that is provided should date from within the last 12 months.

#	Required Data – Elexon	Hypothesis					
		1	2	3	4i	4ii	5
11.2.1	Current Load Profiles		X		X	X	
11.2.2	Proxy for HH consumption data for NHH settled meters (PrA data)		X		X	X	
11.2.3	GSP Group Take by GSP Group			X			X
11.2.4	GCF by GSP Group and Settlement Period			X			
11.2.5	Total daily electricity settled by Elexon						X
11.2.6	Daily Settlement accuracy when RF is at 16 months						X

11.2.1 Current Load Profiles

1. Current Load Profiles for a Settlement Day for Profile Classes 1-4

   Presented as Profile Coefficients for each Settlement Period. Data will be provided by Elexon in Clock Time (not UTC) and will need to be converted to align with the consumption data (provided in UTC) by the analyst team.

2. Current Load Profiles for a Settlement Year for Profile Classes 1-4

   Presented as Profile Coefficients for each Settlement Period. Data will be provided by Elexon in Clock Time (not UTC) and will need to be converted to align with the consumption data (provided in UTC) by the analyst team.

11.2.2 Proxy for HH consumption data for NHH settled meters (PrA data)

Minimum required sample size: As large an MPAN sample size as possible – all available PrA data

1. Consumption data by Settlement Period for 12 months (i.e. 17,520 (365x48) individual consumption values for each MPAN). All data to be provided in UTC time.
2. Anonymised MPAN (require a random ID for each customer)
3. Market Segment Indicator
4. GSP Group ID
5. Domestic Premises Indicator flag
6. Measurement Quantity
7. Profile Class

11.2.3 GSP Group Take by GSP Group

1. Group Take by Settlement Period for 12 months of the year (i.e. 17,520 (365x48) individual consumption values for each GSP Group) for each of the 14 GSP Groups. Data will be provided by Elexon in Clock Time (not UTC) and will need to be converted to align with the consumption data (provided in UTC) by the analyst team.

2. GSP Group ID

11.2.4 GCF by GSP Group and Settlement Period

1. GCF by GSP Group and Settlement Period for the last 12 months

11.2.5 Total daily electricity settled by Elexon

1. Total daily electricity settled by Elexon for the last 12 months

11.2.6 Daily Settlement accuracy when RF is at 16 months

1. Daily Settlement accuracy, for the past 12 months, when RF is at 16 months. This should be presented as a percentage of total daily electricity settled.

11.3 Other Sources

The table below presents all potential data that could be requested from Programme Participants that are neither Suppliers nor Elexon for the purpose of analysing the hypotheses. The table also shows the mapping between the requested data and the hypothesis / hypotheses that they are related to. The numbering of the hypotheses refers to the hypotheses in the order that they are presented in Section 4 (Scope) of this document.

All data that is required for a given hypothesis should be provided by the same party (ideally, may not always be possible), focus on the same MPAN sample and cover the same time period.

All data that is provided should date from within the last 12 months.

#	Required Data – Other Sources	Proposed Source	Hypothesis					
			1	2	3	4i	4ii	5
11.3.1	Forecasted % penetration of Smart meters	DESNZ	X			X	X	X
11.3.2	Forecasted number of customers with heat pumps	DESNZ / Ofgem Forecasts	X					
11.3.3	Forecasted number of customers with EVs and home charging vs public charging %		X					
11.3.4	Count of currently unmeasured Export sites by GSP Group	FiT Register (Ofgem)			X			
11.3.5	% split of Domestic customers on Unrestricted tariffs (PC1) vs ToU tariffs (PC2)	EES				X		
11.3.6	% split of non-Domestic customers on Unrestricted tariffs (PC3) vs ToU tariffs (PC4)						X	
11.3.7	Count of Traditional meters across UK market							X
11.3.8	Count of Smart meters across UK market							X
11.3.9	Count of Advanced meters across UK market							X
11.3.10	New TBC guidelines regarding if there will be any Supplier incentivisation to collect reads within 4 months	MHHS Programme (PAWG)						X

11.3.1 Forecasted % penetration of Smart meters – DESNZ

1. Forecasted % penetration of Smart meters across the next 10 years in quarterly intervals

2. Forecast split by Smart, Advanced and Traditional meter segments
3. Forecast split by Domestic and non-Domestic
4. Forecast split by GSP Group (desirable)

11.3.2 Forecasted number of customers with heat pumps – DESNZ / Ofgem Forecasts

1. Forecasted number of customers with heat pumps across the next 10 years in quarterly intervals
2. Forecast split by Domestic and non-Domestic
3. Forecast split by GSP Group (desirable)

11.3.3 Forecasted number of customers with EVs and home charging vs public charging % – DESNZ / Ofgem Forecasts

1. Forecasted number of customers with EVs across the next 10 years in quarterly intervals
2. Forecast split by Domestic and non-Domestic
3. Forecast split by GSP Group (desirable)
4. Forecasted % split, in the number of customer terms, between home charging and public charging across the next 10 years in quarterly intervals
5. Forecasted % split, in kWh terms, between home charging and public charging across the next 10 years in quarterly intervals

---

11.3.4 Count of currently unmeasured Export sites by GSP Group – FiT Register (Ofgem)

1. Most recent count of currently unmeasured Export sites by GSP Group
2. GSP Group ID

11.3.5 Count of Traditional meters across UK market – EES

1. Count of Traditional meters across UK market

11.3.6 Count of Smart meters across UK market – EES

1. Count of Smart meters across UK market
2. Required data item 11.1.16 '% Smart meters with working communications' can then be applied to this number to estimate the count of Smart meters without working communications across UK market.

11.3.7 Count of Advanced meters across UK market – EES

1. Count of Advanced meters across UK market
2. Required data item 11.1.17 '% Advanced meters with working communications' can then be applied to this number to estimate the count of Advanced meters without working communications across UK market.

11.3.8 % split of Domestic customers on Unrestricted tariffs (PC1) vs ToU tariffs (PC2) – EES

1. Most recent % split of Domestic customers on Unrestricted tariffs (PC1) vs ToU tariffs (PC2)

11.3.9 % split of non-Domestic customers on Unrestricted tariffs (PC3) vs ToU tariffs (PC4) – EES

1. Most recent % split of non-Domestic customers on Unrestricted tariffs (PC3) vs ToU tariffs (PC4)

11.3.10 New TBC guidelines regarding if there will be any Supplier incentivisation to collect reads within 4 months – MHHS Programme (PAWG)

1. Confirmation from the Performance Assurance Working Group (PAWG) what the latest direction of travel is on this topic (noting that nothing has currently been agreed).
2. Confirmation from the Performance Assurance Working Group (PAWG) what Supplier incentivisation has been agreed on, if any, once confirmed and approved.