Discussion paper for electricity settlement expert group – options for data estimation

Author: Ofgem
Audience: Electricity settlement expert group and other interested stakeholders
Date of meeting: 10 July 2014

1. Purpose of the paper

1.01 As part of the Smarter Markets Programme, Ofgem has convened an expert group to support its work to examine how consumers can be settled against their half-hourly (HH) consumption data. This paper sets out Ofgem's initial views on the options for data estimation for discussion at the expert group meeting on 10 July 2014.

1.02 We are seeking views from the expert group on the following questions:

- Are there other options that should be considered?
- Are there links to other market arrangements that should be considered?
- Do you have any comments on our initial assessment of the options?

2. Structure of the paper

2.01 This paper contains the following sections:

- Section 3 – explains why we are exploring options for data estimation
- Section 4 – describes the options we have developed
- Section 5 – explores interactions with other areas of the project and key dependencies
- Section 6 – presents our initial assessment of the options against the evaluation criteria
- Section 7 – explains our next steps
- Annex – describes current processes for estimation.

3. Description of the issue

3.1 Why will estimation be necessary?

3.01 Smart and advanced metering, capable of recording HH consumption data, is currently being rolled out across GB. Ofgem considers that consumers are likely to benefit from being settled against this data. However, even at the completion of roll-out, there will be situations where HH data is not available for one or more settlement runs, thus requiring some form of estimation.

3.02 This could occur in the following situations for sites where a smart/advanced meter is installed:

- There is a delay in receiving data from the meter, for example due to communications disruptions.[1]
- There are errors in the data, for example owing to data corruption.
- The meter had been incorrectly installed or configured and did not correctly record the HH data.

3.03 It would also be the case at sites where no smart/advanced meter is installed and sites where there is no WAN coverage.[2] The former will continue to use traditional non-half-hourly (NHH) meters and therefore no HH data will be available at any point.

3.04 Included within the scope of this project is settlement of export. Smart meters will be able to record HH export data. For export, and for the same reasons as those listed above, there will be times when the actual export data is missing for settlement.

3.2 Why is estimation in scope for the settlement project?

3.05 The advent of smart metering is likely to require one or several new processes around estimation methods for sites that are settled HH. Although there is a current process for estimating data from HH sites – BSCP502 – some elements of it may make it unsuitable for application to millions of sites without some form of modification. It was designed for larger sites with different meter types and technologies. For example, unlike some larger sites, domestic sites will not have check meters, which can enable the actual data to be recovered if the main meter fails.

3.06 As part of the settlement project, we are exploring the options for delivering an efficient process for settling domestic and smaller non-domestic customers using their HH data. This is important because the costs of this process are ultimately borne by consumers. There are direct costs associated with estimation, for example data processessing requirements. There are also indirect cost implications: for example, a sufficiently accurate estimation process has the potential to reduce the need for expensive manual retreival of actual data prior to a settlement run.

3.07 Members of the expert group meeting have highlighted the need to consider how the minority of consumers with traditional meters will be managed under a HH regime. We consider it important that the interests of these consumers are protected, including by identifiying an appropriate solution for estimating their consumption. The solution will affect the accuracy of settlement for consumers with traditional meters, with potential implications for how suppliers charge them for energy. There will also be costs associated with estimation that could be allocated in a variety of ways.

[1] At the expert group meeting on 10 July, the DCC will give a presentation on its performance standards. This will help in understanding the likelihood of communications disruptions driving estimation.

[2] In line with the assumptions underpinning the settlement project, we are assuming that only a small number of sites do not have a smart meter installed by the end of 2020. Similarly, we are assuming that only a small number of sites will lack connectivity. See, Ofgem, Analytical framework for electricity settlement project, p.8. (https://www.ofgem.gov.uk/ofgem-publications/88229/13analyticalframework.pdf).

3.3 Objective

3.08 Keeping costs low and accuracy high is important for ensuring that the full benefits of HH settlement are realised.[3] As such, the high-level objective of this focus area is to identify a process or set of processes for estimation that will enable sufficiently accurate settlement, whilst remaining cost-effective.

4. Options

4.1 Introduction

4.01 Domestic sites without smart meters will continue to be settled with some form of profiling (the current profiling process is explained in the annex, 8.3). There will be no HH data for these sites, except through sampling. We have listed the options for these sites below (section 4.2).

4.02 The majority of sites, those with smart or advanced meters, could have their HH data estimated by a range of different methods. These options can be classed as either 1) using profiling only; or 2) using site-specific estimation techniques, with profiling as a last resort. We go on to develop options for these sites (section 4.3).

4.03 The below list of options is the result of our intial analysis. It is not an exhaustive list of all theoretically possible options and we remain open to considering other ideas.

4.2 Options for profiling customers on traditional meters

Option 1) Maintain current profiles

4.04 This option would maintain the current processes and requirements for updating profiles. There is a requirement for profiles to be updated annually.[4] This requires a HH meter to be installed in a sample of customers' homes. Around 2000 sites for Profile Classes 1-4 are currently sampled in this way. This process is detailed further in the annex.

4.05 This option would retain up-to-date profiles of the consumption patterns of the residual population of consumers without smart meters. This may be particularly important before the completion of the smart meter roll-out whilst these numbers will still be significant.

4.06 As current sample participants transition to smart meters during the roll-out, their consumption behaviour may change such that they no longer represent the remaining customers on traditional meters. As such, they may have to move out of the sample. A sample of traditionally metered customers of sufficient size would have to be maintained.

Option 2) Freeze current profiles

4.07 This option would remove the requirement to keep Profile Classes 1-4 updated, in other words to freeze the profiles in their current configurations. The Profile Administrator (PrA) would thus stop all load research and sampling. The only remaining process would be to apply the sunset and temperature corrections to the profile cooefficients.

Option 3) Profiling with profiles generated from smart data

4.08 The advent of smart meters will provide a large new dataset that could be of use for creating profiles. There will be up-to-date data available for millions of sites, and no or limited additional costs of using it. This option would see data from smart meters used to generate profiles for sites with traditional meters.

[3] For Ofgem's view of the benefits of using HH data for settlement, see Ofgem, 'Electricity Settlement Reform launch statement', pp.8-16. (https://www.ofgem.gov.uk/ofgem-publications/87053/electricitysettlementlaunchstatement.pdf)
[4] ELEXON, BSCP516, 3.2. (http://www.elexon.co.uk/wp-content/uploads/2011/10/bscp516_v7.0.pdf)

4.09 The availability of smart data has the potential to make affordable a large number of combinations of options for profiling. Variables that could be changed include:

- **Dynamic versus static profiling.** With dynamic profiling a profile is created for each settlement day using HH data from a sample for that same day. The profile would be ready in time for the first settlement run. Static profiling is the method whereby standing profiles are calculated at a point in time using time series data from a sample. They can then be applied to future time periods with no further sampling. They may be updated with new data but this is not essential.

- **The size of the sample.** This could range from a small sample of a few thousand to potentially the whole population of meters.

- **The number and type of profiles.** There could be profiles for each region (GSP group) or more Profile Classes to reflect customers' tariff types, for example.

- **Profiling for volumes.** Profiling could generate estimated consumption volumes as well as load shapes. At present an individualised EAC/AA[5] is applied to profiles to estimate total consumption but new profiling methods could be used to cut out this additional step by estimating volumes directly from the sample data.

4.10 There is clearly scope for making profiling significantly more sophisticated. However, there may be an accuracy-cost trade-off as larger samples or more Profile Classes, would require more processing. Since the data itself would be readily available, these costs are likely to be modest and limited to the systems development and analyst costs.

4.11 Using smart data to generate profiles for customers with traditional meters clearly has complications. The two groups of customers may exhibit different behaviours. Smart meters are likely to change consumption patterns over time, both simply through the availability of up-to-date consumption data and through price signals (such as time-of-use tariffs). Therefore the sample selected for generating these profiles would have to be comprised of customers with similar characteristics to those of traditional meters. For example, it may not be appropriate to include customers on certain time-of-use tariffs.

4.3 Options for estimating HH data for customers on smart/advanced meters

Option 4) Profiling with profiles generated from smart data

4.12 As per Option 3, smart data opens up a range of possibilities for profiling. This option would see such data used to create profiles to estimate all missing HH data for sites with smart/advanced meters.

Option 5) Freeze current profiles

4.13 As per Option 2, this option would retain the current profiles but no longer keep them updated. They would be used to estimate all missing HH data for sites with smart/advanced meters.

4.14 The option of maintaining current profiles using current sampling techniques does not seem to be a credible option here. If we wanted the current profiles to be kept up-to-date for the population of smart meters, it would be cheaper to use the smart meter data (as per Option 4) than secondary meters.

Option 6) Site-specific estimation using historical site data

4.15 Smart meters will over time enable suppliers to accumulate historical data on a particular site's HH consumption patterns. This data could be used to estimate HH consumption when the actual data is missing. Our understanding from discussions with stakeholders is that this form of estimation can be highly accurate.

[5] Estimated Annual Consumption/Annualised Advance. These are the current means of estimating volumes, based on historical volumes for that particular site rather than using data from a sample, as suggested here.

4.16 The process would use the most relevant data available to generate the estimate. For example, the most accurate estimate for a given settlement period might be an average of the same period on the same day in the preceding weeks. Longer-term data could be of use if more recent data is unavailable (for example if the meter stopped recording accurately for a period of several months). For example, when estimating data in July, if data is lacking for both June and July, the best solution may be to use data from July in the previous year.

4.17 This approach is used today for HH sites: it is described in BSCP502[6] (explained in the annex, 8.2). The current procedure would likely require a degree of tailoring to make it suitable for application to millions of sites. The intention would be to design a process that runs automatically (and deals with exceptions automatically) and still provides accurate estimates given the available data.

4.18 In some cases, sufficient historical data will be unavailable for such estimation. This would be the case for newly energised sites, sites with a newly installed smart meter and sites where the supplier/Supplier Agent lacks access to historical HH data, for example following a change of supplier. It may also apply in cases where the customer changes tariff type or the tenancy of a property changes, in which case historical consumption data may no longer be considered representative of new consumption patterns. In these cases, it would be necessary to rely on one or several 'profiles of last resort'. The profiles could be generated in different ways:

Sub-option 6a) Profile(s) of last resort generated from smart meter data

4.19 Just as in Option 4, smart meter data would be used to create profiles of last resort, with a wide range of profiling options.

Sub-option 6b) Frozen current profiles used as profiles as last resort

4.20 Estimation would revert to the current profiles. These profiles would be frozen, as in Option 5.

4.21 For all of the above options for smart sites, sites will continue to require Profile Classes. For Options 5 and 6b this could be straightforward, since sites already have assignment to one of the current Profile Classes. However, work would be required to allocate sites to any *new* set of Profile Classes (as per options 4 and 6a).

4.4 Note on export estimation

4.22 There are currently processes in place for estimating both HH and NHH metered export. As set out in the annex (8.2), BSCP502 applies to HH metered export: Half-hourly Data Collectors (HHDCs) use zero values for sites with import/export netting. ELEXON also has a process for profiling NHH-metered micro-generation export.[7]

4.23 As such the above options are – in principle – equally applicable to export as import. The optimal solution may, however, be different for the two. An important difference is that under BSCP502 there is no resort to a profile if actual export data is lacking (for sites with netting) – a zero value is used instead. Under Option 6 therefore, if the future process reflects current HH arrangements, there would be no recourse to a profile of last resort.

5. Interactions and dependencies

5.01 In assessing the options it is important to take account of interactions and dependencies.

5.02 First, the number of customers on traditional meters. The greater this number, the greater their potential impact on the accuracy of settlement as a whole. As presented at the first expert group meeting in June, we are using the same assumptions as DECC in its impact assessment for the roll-out of smart metering:

[6] ELEXON, BSCP502, 4.2. (http://www.elexon.co.uk/csd/bscp502-half-hourly-data-collection-for-sva-metering-systems-registered-in-smrs/)
[7] ELEXON, BSC section L, 1.5. (http://www.elexon.co.uk/wp-content/uploads/2014/06/Section_L_v16.0.pdf)

namely that over 99 per cent of domestic sites will have a smart meter installed by the end of 2020.[8] However, prior to the completion of the roll-out this percentage will, of course, be lower. There is therefore an important linkage with our separate workstream on transition, which will consider the appropriate timing on moving to using HH data.

5.03 Second is the settlement timetable. This a separate focus area for the project, but it interacts with estimation. An accurate estimation process makes earlier settlement runs more feasible even if some actual data is missing; and the earlier settlement runs take place, the more reliance there may be on estimation.

5.04 Third, a separate strand of this project is looking at responsibility for the the Data Processing (DP) and Data Aggregation (DA) functions. The outcome of this work on estimation may affect the nature of the DP function. Conversely, the outcome of the work on DP/DA may influence the evaluation of options for estimation. In the more detailed assessment phase we may need to incorporate a range of possible DP/DA scenarios in our analysis (this will be important for calculating the costs of various options, for example). At this point it suffices to be aware of the linkage.

6. Evaluation of options

6.01 Table 1, below, presents our intial qualitiative evaluation of the options against the common evaluation criteria for the electricity settlement project. We would welcome comments on this evaluation from the expert group.

[8] Ofgem, Analytical framework for electricity settlement project, p.8. (https://www.ofgem.gov.uk/ofgem-publications/88229/13analyticalframework.pdf)

Table 1 – evaluation of options against evaluation criteria

	Options for profiling customers on traditional meters			**Options for estimating HH data for customers on smart/advanced meters**			
Options/sub options	**1) Maintain current profiles**	**2) Freeze current profiles**	**3) Profiling with profiles generated from smart data**	**4) Profiling with profiles generated from smart data**	**5) Freeze current profiles**	**6) Site-specific estimation using historical site data**	
						6a) Profile(s) of last resort generated from smart meter data	**6b) Frozen current profiles used as profiles as last resort**
Accuracy	This option will maintain the current level of accuracy of profiling.	Future changes to load patterns would not be captured in the profile. As such this option may reduce the accuracy of profiling in comparison to Option 1.	This option has the potential to significantly improve the accuracy of profiling over the current system, for example by generating more profiles. The uncertainty is if customers in the smart population can be found who are sufficiently similar to the traditional population.	This option has the potential to significantly improve the accuracy of profiling over the current system, for example by generating more profiles. More accurate than Option 5 but less so than Option 6.	Future changes to load patterns would not be captured in the profile. This is the least accurate solution for these meters.	From conversations with stakeholders, we understand that site-specific estimation can be highly accurate. The smart-generated profiles make this option more accurate than 6b.	From conversations with stakeholders, we understand that site-specific estimation can be highly accurate. Using frozen profiles make this option less accurate than 6a.
Speed	Unlikely to impact the speed of settlement – minor processing required to generate estimates.	Unlikely to impact the speed of settlement – minor processing required to generate estimates.	The dynamic profiling options would require data processing to generate estimates prior to settlement. This is unlikely to impact the speed of settlement.	The dynamic profiling options would require data processing to generate estimates prior to settlement. This is unlikely to impact the speed of settlement.	Unlikely to impact the speed of settlement – minor processing required to generate estimates.	This would require data processing to generate estimates prior to settlement. This is unlikely to impact the speed of settlement.	This would require data processing to generate estimates prior to settlement. This is unlikely to impact the speed of settlement.
Coverage	Sites with traditional meters. Used to estimate all consumption.	Sites with traditional meters. Used to estimate all consumption.	Sites with traditional meters. Used to estimate all consumption.	Sites with advanced/smart meters when actual data is missing.	Sites with advanced/smart meters when actual data is missing.	Sites with advanced/smart meters when actual data is missing. Profile of last resort used when sufficient historical data is also lacking.	Sites with advanced/smart meters when actual data is missing. Profile of last resort used when sufficient historical data is also lacking.

	Options for profiling customers on traditional meters			Options for estimating HH data for customers on smart/advanced meters			
Options/sub options	**1) Maintain current profiles**	**2) Freeze current profiles**	**3) Profiling with profiles generated from smart data**	**4) Profiling with profiles generated from smart data**	**5) Freeze current profiles**	**6) Site-specific estimation using historical site data**	
						6a) Profile(s) of last resort generated from smart meter data	**6b) Frozen current profiles used as profiles as last resort**
Simplicity	Relatively complex, involving several steps and different parties: consumers, suppliers, the PrA and a third party contractor.	Simpler than Option 1, as it eliminates the need for sampling.	Simpler than Option 1 because no need for secondary metering.	More complex than Option 5 but simpler than Option 6, since there is no site-specific calculation and therefore fewer steps in the process.	The simplest option for smart/advanced meters.	More complex than Options 4 and 5, as it requires both site-specific calculations and process around profiling.	More complex than Options 4 and 5, as it requires both site-specific calculations and process around profiling.
Costs – capital	Nil – the process is already in place.	Nil – simplifies existing process.	Capital costs would relate to developing new piece of software to calculate profiles (central costs).	Capital costs would relate to developing new software to calculate profiles (central costs).	Nil – introduces no new process.	Capital costs would relate to developing new software to calculate site-specific estimates and profiles (costs to suppliers/Agents).	Capital costs would relate to developing new software to calculate site-specific estimates and profiles (costs to suppliers/Agents).
Costs – operational	Costs of running the sample and conducting data analysis.	Eliminates sampling costs, which make up the majority of profiling costs. Requires minimal data analysis.	Minimal manual data analysis. Likely to be cheaper than current profiling methods (Option 1) since eliminates physical sampling.	Minimal manual data analysis. For context, likely to be cheaper than current profiling methods.	Minimal manual data analysis.	Running costs of automated systems for site-specific estimation. Minimal manual data analysis for profiling.	Running costs of automated systems for site-specific estimation. Minimal manual data analysis for profiling.
Flexibility	This option would enable estimation to adapt to changes in load patterns. However it would still lag the actual changes as the profiles are updated annually. It would also be possible to move to Option 2 at a later date, if the population on traditional meters became too small to merit updating the profiles.	Changes to demand response could diminish the accuracy of this method.	If profiles were regularly reviewed, this method could prove flexible and adaptive to changes in load patterns.	If profiles were regularly reviewed, this method could prove flexible and adaptive to changes in load patterns.	Changes to load patterns could diminish the accuracy of this method.	Since site-specific estimation is based on recent, relevant data, this would naturally adapt to changes in load patterns. The profile of last resort would also adapt to changes, as per Option 4.	Since site-specific estimation is based on recent, relevant data, this would naturally adapt to changes in load patterns. Less flexible than option 6a as profile of last resort would be fixed.

	Options for profiling customers on traditional meters			Options for estimating HH data for customers on smart/advanced meters			
Options/sub options	**1) Maintain current profiles**	**2) Freeze current profiles**	**3) Profiling with profiles generated from smart data**	**4) Profiling with profiles generated from smart data**	**5) Freeze current profiles**	**6) Site-specific estimation using historical site data**	
						6a) Profile(s) of last resort generated from smart meter data	**6b) Frozen current profiles used as profiles as last resort**
Integration	Well integrated with existing arrangements as it simply maintains a current process.	Well integrated with existing arrangements as it simplifies a current process.	Would require new process to be developed.	Would require new process to be developed.	Well integrated with existing arrangements as it simplifies a current process.	Site specific-estimation would require new process to be developed, but similar to existing BSCP502. Profiling would require new process to be developed as per Option 4.	Site specific-estimation would require new process to be developed, but similar to existing BSCP502. Profiling relies on a current process.
Implementa-tion	The process is already in place – the only change may be adjusting the sample to reflect traditional meters only.	Freezing profiles could be undertaken quickly and easily.	Would require design and testing prior to operation. It would also require sufficient smart meter data to be available. This may delay the earliest date at which it could be operationalised.	Would require design and testing prior to operation. It would also require sufficient smart meter data to be available. This may delay the earliest date at which it could be operationalised.	Freezing profiles could be undertaken quickly and easily.	Site specific-estimation would require new process to be developed. Profiling as per Option 4.	Site specific-estimation would require new process to be developed. Freezing profiles could be undertaken quickly and easily.
Consumer impact	These options do not have direct effects on consumers. The indirect effects on consumers of these options will be delivered primarily via their impact on the costs and accuracy of settlement (see above), which will in turn feed into prices and the potential for tariff innovation.			These options do not have direct effects on consumers. The indirect effects on consumers of these options will be delivered primarily via their impact on the costs and accuracy of settlement (see above), which will in turn feed into prices and the potential for tariff innovation.			

7. Next steps

7.01 At the expert group meeting on 10 July we will present a summary of this paper to set the scene for an initial discussion on the options, focused on the questions set out in section 1.02. Drawing on this initial discussion, we will further refine the options and our evaluation of them in preparation for a second detailed discussion at the third expert group meeting on 31 July. We ask that expert group members reflect on this paper and the discussion at the 10 July to develop their thoughts ahead of the 31 July meeting.

8. Annex – Background on current estimation processes

8.01 There are currently processes in place for estimation for settlement, one for sites that are settled HH (BSCP502) and one for sites that are settled NHH (profiling). This annex gives a brief overview of how they work.

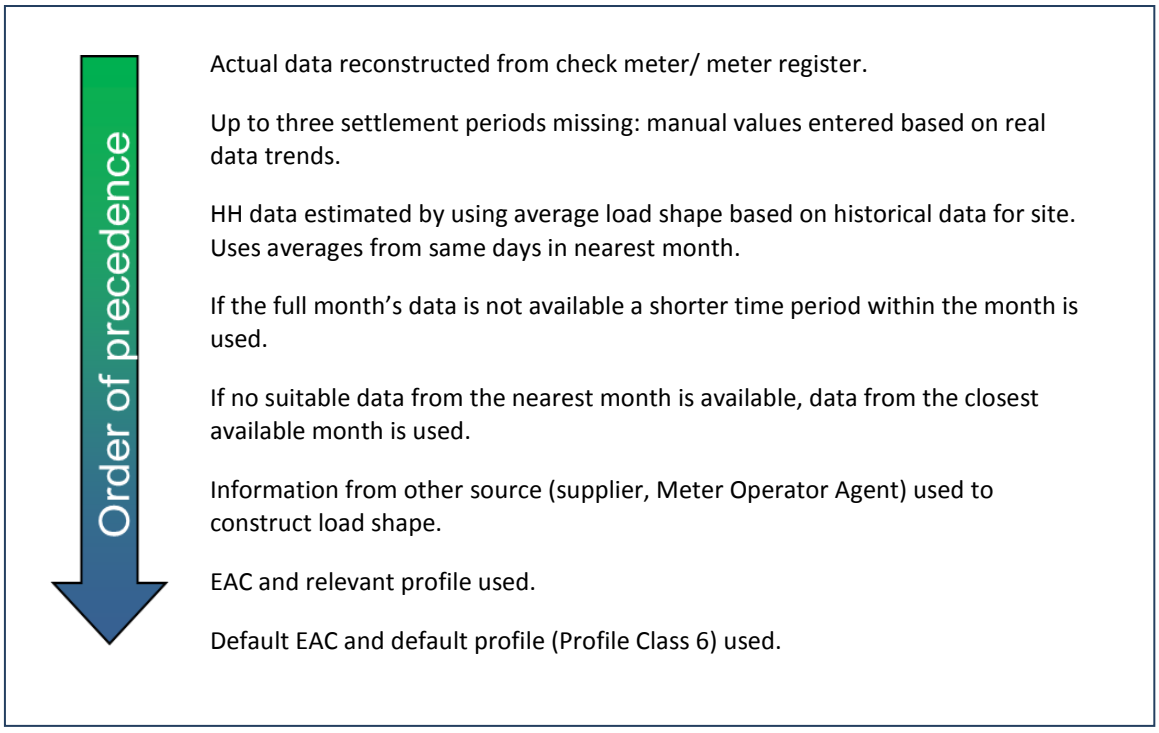
8.2 HH estimation (BSCP502)

8.02 BSCP502 sets out the data estimation methods that HHDCs are to follow when actual data is missing or erroneous. It follows an order of precedence, first reconstructing actual data, and then estimating based on data of declining relevance. These are the 'standard methods'.

8.03 However, HHDCs are also instructed to use 'local information' when available and use appropriate historical data if this is considered to give a more accurate estimate. The example given in the procedure is that when estimating a school's consumption during August it would be more appropriate to use the previous August's data, rather than July's.

8.04 Figure 1 summarises the methods to follow for the estimation process for **import**.

Figure 1 – standard methods in the estimation process for import



8.05 For **export**, at sites that net import and export, the values will be set to zero pending actual data.

8.06 For **export**, at sites with a dedicated export channel, the first step, as for import, is to attempt to reconstruct the actual data. Should this fail, the values are set to zero unless operational data or additional information from another source (supplier, MOA) can be used to construct the profile.

8.3 Profiling

8.07 Profiles determine how a site's electricity consumption should be allocated to each settlement period.[9] All NHH sites are allocated to one of eight Profile Classes according to their characteristics. In scope for this project are Profile Classes 1-2 (domestic) and 3-4 (smaller non-domestic).

8.08 Profiling works by calculating the average consumption shape for each Profile Class across the year. These calculations, adjusted for each day's actual temperature and sunset times, can then be used to allocate a fraction of annual consumption to each HH settlement period. This fraction, multiplied by the annualised consumption (or an estimate of it) provides a volume for each half hour.

8.09 The profiles are created by installing HH meters at a random sample of sites in order to get a representation of the whole Profile Class. The profiles are recalculated on an annual basis using the latest information from the sample.

8.10 Profiles are created and maintained by ELEXON. The actual sampling is undertaken by the Data Gatherer and Sample Manager, a contracted-out role.

[9] For a full explanation of profiling, see ELEXON's guide: http://www.elexon.co.uk/wp-content/uploads/2013/11/load_profiles_v2.0_cgi.pdf.