British Gas comments on the Proposed Allocation of Unidentified Gas Statement (AUGS) for 2020/21

22 January 2020

We appreciate this opportunity to provide feedback on the work being undertaken by the AUGE.

2020_1:

1. Theft methodology: targeting and lead generation bias

The theft methodology is the most critical element of the AUG statement given that it is used to apportion the balancing factor, by far the largest segment of unidentified gas.

Based on input from the Theft Issues Group (a group to consider theft issues under the SPAA and DCUSA) the methodology is now in a state of flux pending further analysis. We would have expected the alignment of understanding between the AUGE and the Theft Issues Group to be prioritised, given the impasse that arose for the 2019/20 process. We understand that the AUGE is developing a method to counter any bias in the volume data provided by suppliers to TRAS (on which qualified outliers are generated) and look forward to more clarity on this work.

In terms of specific feedback for this year, we expect the bias introduced through differences in the address matching rates between residential and commercial sites to be corrected. The population of unbiased leads (i.e. not generated by supplier activity) is underrepresented for non-residential sites. The way these leads are generated relies on the successful matching of address data. British Gas has provided to the AUGE the reported success rate for address matching, which should warrant an appropriate uplift to the allocation of balancing factor to commercial sites.

Finally, there are insights from the relative performance of different suppliers under the Gas Theft Detection Scheme that should be informing the theft methodology, given that suppliers representing nearly a third of the market have not reported a single theft to the scheme.

Recommendations

- Outline how the theft methodology is likely to change to address newly identified sources of bias and indicate the impact on the allocation of the balancing factor for 2021/22.
- Apply an uplift in the theft analysis to account for the differential in address matching success between residential and commercial sites.
- The AUGE should obtain and analyse the anonymised results of the Gas Theft Detection Incentive Scheme.

Response:

The AUGE agrees that the theft methodology is critical to the overall UIG factor calculation, and as such it was given top priority in the 2020/21 analysis alongside the Volume Conversion analysis. We submitted our data specification in April 2019, immediately following the close of the previous AUGE year, and this document included a request for full extended data including TRAS Outliers, ETTOS leads and address matching rates. Unfortunately, despite this early engagement and constant focus on the area, it was not possible to gain agreement with TIG to release the full data for inclusion in the UIG factor analysis. The reasons for this and the agreed route forward, following TIG's recommendation, have been published elsewhere.

We appreciate British Gas raising our awareness of the issue of address matching, which they first did in April 2019. We fully intend to address this area and include a consideration of this

effect in the theft method if the results of our investigation support this: we have already developed a technique to include the effects of address matching success rates in the calculations, but this does not currently form part of the methodology due to a lack of the necessary data to support either analysis or implementation. In order to allow our investigation to take place, domestic and non-domestic address matching results formed part of the data request first submitted in April 2019. Following our meeting with TIG in December 2019, we have resubmitted this data request and we are awaiting the outcome of this. Depending on the outcome of this process, if appropriate data is received it may be possible to carry out an analysis of address matching rates and, if appropriate, apply a correction for this in time for inclusion in the final 2020/21 factors. If this is not possible, then the analysis would be carried out in time for the first draft of the 2021/22 factors. Alternatively, if any industry party has access to such data and is permitted to share it with the AUGE, we would be grateful to receive it. An internet search has found no public domain data in this area.

It has been confirmed by Electralink that data from the Gas Theft Detection Scheme is a subset of the TRAS Outcome File data that we already have. As part of our work we have analysed anonymised Shipper data for both TRAS and CMS. Our conclusions from this work agree with those of British Gas, i.e. that a small number of Shippers are responsible for the vast majority of confirmed thefts, whilst a significant minority have no confirmed thefts at all associated with them. The key to our theft analysis, however, is the identification of the pattern of bias caused by Shipper/Supplier detection activity and the removal of this effect. Individual Shipper/Supplier strategies create this pattern, but it is the pattern itself that must be quantified and corrected for. This is precisely what our theft method has been designed to do, and therefore it is not necessary to go back to the cause of the effect (i.e. differences between Shippers/Suppliers): quantification and removal of the effect itself is the key. This is what our method does, and so it isn't necessary to go back to Supplier-level information in order to achieve this.

As with all areas of the UIG analysis we encourage contributions from the industry regarding the theft method. The method is described in detail in the AUGS, and the calculations themselves can be followed in spreadsheets provided on UK Link Secure Docs. If any industry party has suggestions for further sources of bias that we have not currently taken account of these will be taken into consideration and included in the method if they are found to be significant.

2020_2:

2. Theft methodology: assumptions related to the impact of smart meter installation

It would be useful to provide transparent assumptions on the changes to the rate of theft once smart meters are installed. There are several factors that when fully considered should reduce the rates of undetected theft expected from smart sites, for example:

- sites where theft is occurring are less likely to willingly adopt a smart meter; and
- theft at smart metered sites is likely to be more rapidly detected.

Recommendation

- *Please provide a summary of the explicit assumptions and derived values related to the difference in rates of theft detection between smart and non-smart sites.*

Response:

There are no direct assumptions made regarding the relative level or magnitude of theft at Smart Meter sites: the approach is fully data-driven and so all conclusions in this area are based on information from the TRAS Outcome files. The meter type associated with each theft is derived using the Meter Serial Number, which takes distinct forms for traditional meters, Smart Meters and Electronic Token Meters (ETMs). This is done by the CDSP prior to the data being supplied to the AUGE in order to maintain anonymity.

The theft population is then split into components for traditional meters, Smart Meters and ETMs. Valve Tampers from ETMs, which are used as the best indicator of fiscal theft (i.e. financial theft that does not affect the meter read) are removed from the dataset because fiscal thefts do not produce UIG. At this stage the dataset is reduced to only those records that arise from an unbiased source; the selection of which sources are considered to be unbiased is the only assumption made throughout the analysis.

From this point onwards, all aspects of the analysis are data-driven: this covers unbiased leads, the relationship between investigations and confirmed thefts, and the average magnitude of theft for each site. For all types of meter, including Smart Meters, these are calculated from TRAS records from the relevant meter type. The differences in such things as Investigation to Detection Ratio and Average kWh per Theft can be found in the calculation spreadsheets on UK Link Secure Docs.

In summary, the following principles therefore apply. These are presented below using PC3 01B and PC4 01B as examples:

- Theft rates for PC3 01B and PC4 01B are calculated using the most recent TRAS Outcome File data available.
- Any site in PC3 01B is assumed to have the same rate of theft, calculated in this way. The same applies to PC4 01B.
- If a site moves from PC4 01B to PC3 01B, its theft rate is assumed to change to the one associated with its new category.
- If sites with no theft are transferred to PC3, this will result in the same amount of theft being seen in the TRAS Outcome Files for this Product Class, but from a larger population. Hence the theft rate will drop for subsequent AUGE years.
- If sites with a high proportion of theft are transferred to PC3, this will result in an increase in the theft to population ratio being seen in the TRAS Outcome Files for this Product Class for subsequent years. Hence, whilst any site moving between Product Classes will be assigned the prevailing UIG rate for its new class, as class compositions change this will be reflected in UIG factors that evolve over time as they are recalculated year on year.
- As more Smart Meters are installed in PC4, it is likely that the theft rate will go down for subsequent AUGE years. This is due to the fact that Smart Meters currently have a lower theft rate than traditional meters: as long as this continues to be the case, increasing the population of Smart Meters in PC4 will reduce the theft rate.

In order to account for the changes in meter populations between the time period covered by the TRAS Outcome Files and the forecast year, these population changes are first estimated. This includes an assessment of Product Class populations as well as Smart Meter populations (based on installation rates from the quarterly BEIS Smart Meter reports). This allows all relevant populations to be forecast for the 2020/21 gas year, which in turn allows the unbiased theft figures described above to be adjusted for the effects of those changes.

2020_3:

3. Theft methodology: PC3 analysis

The AUG Statement describes the process to uplift the rates of observed theft from smart sites due to the relative recency of most smart installations.

An initial review of the AUGE's theft analysis relating to PC3 highlights some points of concern:

- the starting point for PC3 theft is based on a sample of five observations, suggesting a high degree of unreliability, which is then amplified significantly to reflect the mass migration to PC3;
- more than half of the observations of theft in PC3 EUC1 to date are from credit meters, which we believe should not be in PC3 in the first place; and
- average thefts in the analysis are higher for existing PC3 EUC1 supply points (23,311 kWh) than for PC4 EUC1 smart metered supply points (14,255 kWh) – given the mass migration is comprised of smart sites from PC4, we would expect an adjustment so the average theft volume assumptions for 2020/21 are more reflective of the PC4 volumes.

Recommendation

- *Consider revising the assumptions in the theft analysis related to PC3 EUC1.*

Response:

Whilst it is true that there have only been 5 confirmed thefts from PC3 since TRAS go-live, this data is not used in the calculation of theft rates. It is provided in the calculation spreadsheet available on UK Link Secure Docs for comparison purposes only, and is greyed out to indicate that it is not a direct part of the calculation. The root data set for PC3 (and indeed all PCs and meter types) is Unbiased Leads: there are 63 of these in total for PC3. Whilst this leads to a total of only 11 unbiased detections for PC3, this is a result of a series of calculations based on the initial 63 data points and therefore this small number of detections is not the basis for calculations itself.

It is also recognised that 3 of the 5 confirmed thefts in PC3 are listed as being on credit meters. These have been retained in the analysis for a number of reasons:

- There is no requirement for a site to have a Smart Meter to be in PC3. The only requirement of this Product Class is to submit batched daily readings. According to the asset data supplied to the AUGE by the CDSP, there are a number of sites without Smart Meters in PC3.
- A large number of sites in PC3 are not fulfilling their meter read obligations and, subject to the approval of Mod 664, may end up being moved back to PC4. We do not have information about what types of meter these are, but it is highly likely that any credit meters moved to PC3 will be part of this. Again, this illustrates the potential for credit meters to be in PC3.
- The meter type calculation relies on the Meter Serial Number as held by the CDSP, and hence also relies on this information being updated when the meter changes. The Product Class associated with each site is likely to be accurate because it is a fundamental part of the CDSP process. Meter type may be inaccurate if the Meter Serial Number has not been updated when the meter was exchanged. As such it would not be valid to remove these records from PC3.

The values of kWh per Theft are based on Confirmed Theft data from the TRAS Outcome files and it is recognised that some of these are based on small numbers of thefts. This will be reassessed following the supply of the new full TRAS dataset (as per the resubmitted data request) and updated as appropriate.

2020_4:

4. Other factors influencing theft assumptions related to PC3 migration

Where a supplier is active in revenue protection, there is the possibility that risk profiling insights could be used to support decisions about which supply points to migrate to PC3, as a "curation" effect.

While strategic decisions of this nature are unlikely to be uniform across suppliers, and will by nature be confidential, we nevertheless expect that future TRAS data will display a disparity in theft detection rates between otherwise identical smart-enabled sites in PC3 and PC4.

In addition, the increased access to data necessitated by the PC3 settlement processes introduces greater opportunities to detect anomalous data, which may well increase theft detection rates on smart meters in PC3, reducing undetected theft and feeding back into the ability to curate theft prone sites out of PC3.

Recommendation

- *We would be supportive of an effort to derive an initial estimate of the size of any curation effect, for example through a confidential poll asking migrating suppliers about their risk profiling intentions.*

Response:

The theft calculation method does not assume uniform rates of theft for Smart Meters across Product Classes – it recognises the effect described above and works on a Product Class by Product Class basis. There are therefore different values for PC3, PC4 Smart and PC4 Traditional (and indeed all other categories) for Unbiased Leads per Meter Point, Investigation to Detection Ratio, kWh per Theft, etc. The advantage of this approach is that it captures the effect described above. The disadvantage at this point is that data is limited for some market sectors due to their relatively low populations and the relatively small number of leads, investigations and confirmed thefts from them. The numbers will increase as additional TRAS Outcome data is sent to the AUGE, which will result in increasing sample sizes and hence confidence in the output. In the meantime, should any further information be gathered, for example via the type of survey suggested, this would be extremely useful in expanding the data available and increasing the confidence in the output as quickly as possible. The AUGE would therefore also be supportive of such an approach.

2020_5:

5. Parameter smoothing

We note that smoothing is applied to the factors in the table along the EUC bands, and query why this is not also applied to some degree across Product Classes.

There is a strong practical argument that there should be less undetected theft from supply points as they move from PC4 to PC3 and start providing daily data to the CDSP, assuming the rules in the Product Class are being followed in terms of read submission (with any deviation from the rules being a matter for the Performance Assurance Committee). Counter to this the draft factors for EUC bands 5-7 are higher in PC3 than they are in PC4, and there is also an interesting anomaly in the draft AUG Statement where PC2 EUC 8 is lower than PC1. These discrepancies are difficult to practically justify.

Recommendation

- *The AUGE should ensure allocation factors don't increase as sites move from Product Class 4 to 3, from Product Class 3 to 2, and from Product Class 2 to 1.*

Response:

It should be noted that the purpose of parameter smoothing is not to ensure that the factors either increase or decrease monotonically across the range of EUC bands, but to remove any excessive scatter that is the result of common cause variation in the data. There is no practical reason why sites in EUC 06B should always have a lower factor than those in 05B, for example, if the data consistently shows that this is the case. Such an effect is quite possible within the current calculation framework, but the smoothing ensures that the transition from EUC to EUC is consistent (be it up or down) rather than highly variable.

The logic that factors can move in either direction also applies across Product Classes. Whilst undetected theft is the largest component of UIG, the data shows that theft from AMR sites is extremely rare. This means that differences between Product Classes for higher EUCs are due to other causes: in particular, volume conversion is also significant and this applies in a pattern unrelated to Product Class. The AUGE's analysis has shown that the proportion of sites in higher EUC bands using standard correction factors is higher in PC3 than PC4, which is the reason for the observed pattern in the factors.

In the case of undetected theft, there are competing reasons why it may go up or down between PC4 and PC3:

- PC3 sites return a higher granularity of data than those in PC4 (assuming that network code obligations are fulfilled). This allows greater scope for identification of theft through suspicious consumption patterns, either via TRAS or the Supplier's own analysis. Whilst this will potentially reduce undetected theft in PC3 sites, this type of theft detection has a low conversion rate and so its impact may be limited.
- PC3 sites, assuming they have Smart Meters, have a much lower rate of physical interaction with the meter (such as through Meter Reading Agents). This lowers the scope for reporting physical signs of theft at the meter, which is a lead type with a very high conversion rate. This has the potential to increase undetected theft in PC3.

The current data-driven approach to the theft calculation means that no assumptions are made about the relative size of these competing effects: the output is based on the data, which in turn is based on the reality of the number of leads and confirmed thefts coming from these sources.

As such we do not believe it would be appropriate to smooth the factors between Product Classes, but to allow step changes where the data shows that these exist.

2020_6:

6. Stability between draft and final versions

We welcome comments made by the AUGE on 10 January 2020 that there is no intention for significant changes to the factors in the final version of the statement.

We are supportive of updates that address "clear weaknesses" such as those outlined above, but significant updates other than corrections in response to industry representations should be avoided.

We should note that we are reviewing our rights to appeal under circumstances where:

- there is a failure to take into account consultation responses; or
- there are significant changes made to the final table that are not directly addressing a weakness identified in a consultation response.

Recommendation

- *Any significant changes to the factors for the final version should be limited to where a "clear weakness" identified through consultation has been addressed.*

Response:

The AUGE recognises that it is important for the UIG factors to be as stable as possible between the first draft and the final version, where this is appropriate. It is our aim to make the draft factors as accurate as possible given the data available at the time, and therefore as close to the final version as possible. It is recognised, however, that there was a significant change between the two versions in the 2019/20 analysis. This was for two reasons:

1. The adoption of the new theft method, which was approved by the majority of the industry via the consultation process. It was not possible to implement this in time for the first draft of the AUGS due to the length of time it took to obtain authorisation to use TRAS data.
2. The start of the mass migration to PC3.

For the 2020/21 factors there will be no major change in the methodology before the final AUG Statement, and no major change in current population trends is envisaged before this time. Mod 664 may see the population of PC3 reduce as sites are returned to PC4, but as this modification is still in the Workgroup stage any potential effects of it will only be seen in future AUGE years. Therefore, the only changes will be as a result of the availability of additional data, along with any data corrections that may arise. It should be noted that the AUGE Framework states that the latest available data should be used in factor calculations, and so as the Framework stands, it is necessary for new data to be included, which may result in minor changes to the factors.