Smart Meter Energy Data: Public Interest Advisory Group

A policy dialogue and work programme
led by
Centre for Sustainable Energy & Sustainability First

PIAG Phase 2 - Workshop Report 2

Regulatory assessments and system efficiency: potential benefits of smart-meter energy-consumption data

Author: Maxine Frerk, Sustainability First
with support from Judith Ward, Sustainability First & Simon Roberts, CSE

Status of this Report

This paper was prepared as an input to the work programme of the Public Interest Advisory Group on access to smart meter energy data.

The paper pulls together contributions made at a workshop with PIAG members held on 23 April 2020.

Sustainability First & CSE
PIAG

(Public Interest Advisory Group on access to smart meter data for a public-interest purpose)

Workshop Report 2
Regulatory assessments and system efficiency: potential benefits of smart-meter energy-consumption data

PIAG Workshop – 23 April 2020

Summary

This workshop looked at how access to aggregated / anonymised smart meter data could enable Ofgem to operate more effectively and enable network operators to meet their regulatory obligations more efficiently. As well as hearing from Ofgem we also invited a number of colleagues who on occasion act in an advisory capacity to Ofgem, to contribute their thoughts.

Data is and has always been important to Ofgem in monitoring the market, assessing potential interventions and evaluating their impacts. Ofgem has been active in looking at ways to make more effective use of data and to make that data open where it can, in line with the recommendations of the Energy Data Taskforce.

There are a number of areas where access to more granular smart meter data – anonymised, aggregated and via a trusted processor arrangement - would be of particular value going forward including on assessments of policy decisions on **vulnerability** and **flexibility**, on delivering system level **efficiency**, and in supporting new approaches to **evaluation** and identification of **emerging market issues**.

On **vulnerability** a particular concern is the need to understand the distributional impacts of regulatory decisions. The various advisers presenting at the workshop had all been engaged in projects trying to do this and the paucity of data was a common theme. In particular there is presently no robust source of data that combines time of use energy consumption data and socio-demographic information and advisers all had to make trade-offs or use proxies. Ofgem have recently published new guidance on how they will assess distributional impacts, including the use of updated archetypes, but they acknowledge that this does not address the need in future to take account of profiles of usage across the day. The SERL database represents a step change in this area and UCL are committed to publishing outputs to facilitate wider use of the data. However there remain limitations given the SERL sample size, in particular for regional level analysis.

On **flexibility** many of the same challenges arise in terms of a lack of granular consumption data from which to identify the scope for demand side response and to assess the impact of interventions.

On **efficiency** the ESO currently has only top down data. A clearer picture of the demand side would enable it to better fulfil its central role in the energy system – both at an operational level keeping the system in balance and in terms of long-term planning. DNOs do have the right to access granular smart meter data once their privacy plans, demonstrating how the data would be aggregated /

anonymised, are approved by Ofgem. Although three DNOs do now have their privacy plans approved and can start collecting data there are questions about whether the present Data Access and Privacy Framework that governs their access is unduly restrictive.

With the uptake of EVs and other low carbon technologies, conventional measures of peak demand cease to be a good guide to where investment will be needed which is why the DNOs need the more granular data. For the same reason, Ofgem also needs access to better data to enable it to judge the efficiency of DNO actions and to carry out benchmarking as part of the RIIO price control process.

On **evaluation** Ofgem has been making increased use of randomised control trials to evaluate policy interventions. Currently bespoke arrangements have to be put in place to deal with each trial which is inefficient and slows the process down. Readier access to smart meter data could enable evaluation to be carried out more effectively.

Finally, as well as looking at known gaps where access to smart meter data could deliver benefits there was a need to reflect on the risks if the regulator did not have access to data that those in the market it oversees did have and would be using to their advantage. The ability to spot **emerging market issues** is dependent on Ofgem having access to relevant data.

In other areas Ofgem have found that combining different data sources, using AI and third-party expertise through Hackathons has yielded fresh insights. It must be assumed that access to smart meter data could yield similar benefits that cannot be articulated at this stage. Opening up the data Ofgem uses to support its decisions would enable a wider range of third parties to contribute to the policy making process.

Of course, the benefits from wider access to smart meter data have to be weighed against the privacy concerns. This was a theme from Phase 1 of PIAG which identified ways in which the benefits could be delivered without compromising privacy through the use of a trusted processor (such as ONS) to aggregate / anonymise the data.

From this workshop it is clear that from a regulatory perspective there are significant public interest benefits in Ofgem, and others, having access to granular smart meter data.

Introduction

This report sets out the conclusions from the PIAG workshop held on 23 April 2020 to explore potential regulatory uses of more granular smart meter data.

The workshop included presentations from **Ofgem** on their use of data for evaluation and monitoring, from advisers to Ofgem who make use of demand-side data (**Frontier Economics, University of Reading, Centre for Sustainable Energy**) and from regulated companies reliant on data to comply with their regulatory obligations (**National Grid ESO and Northern Powergrid**).

The workshop was held virtually which enabled the participation of a wider set of stakeholders than usual and included inputs from Citizens Advice, UCL SERL and others.

This report is organised as follows:

- Background on the PIAG process
- Overview of this second workshop in phase 2 of PIAG
- Discussion of the issues raised under the various themes identified in the workshop:
    - **Vulnerability** and distributional impacts
    - Facilitating **flexibility**
    - Delivering system level **efficiency**
    - New **evaluation** approaches
    - Identifying **emerging market issues**

It concludes with a summary of the workshop findings against the four questions being used to frame the workshops for PIAG2:

- what demand side data is available now?
- what data will be available in future?
- what are the gaps that smart meter data could fill?
- what are the public interest benefits in filling those gaps?

Background

The first phase of the PIAG project considered whether access to smart-meter energy data could better serve GB public policy-making and therefore the public good. We explored how smart meter energy consumption data – suitably anonymised and aggregated via a trusted processor arrangement - could provide a more accurate and granular customer-side data-source than today. We developed some basic use-cases and also looked at what consumer research tells us about customer-thinking on access to their meter data. We also explored potential routes to the data given customer privacy rules plus the absence of any hub in which GB smart-meter data comes together. All the first-phase PIAG papers can be found here - https://www.smartenergydatapiag.org.uk/

The second phase of the PIAG project is designed to explore the **additionality and public-interest value** which access to smart-meter data could potentially unlock. We are assembling **evidence** on how far increased access to smart meter data might offer **additional analytical insight**[1] in service of

[1] against currently available energy data sources – or other data likely to become available (i.e. due to developments from EDTF, SERL, other pilots/new data-sets)

the public interest - **be this to government, regulators, national advisory bodies or public-interest bodies more widely.**
This workshop was the **second** of four in PIAG phase 2. In the first workshop we learned of **the essential role of the official energy consumption statistics produced by BEIS and MHCLG**, both as an input to government's own analysis and also for many other public-interest bodies who make wide use of official energy statistics and data as key inputs to their own work and analysis. The report on that workshop is available here.

This second workshop considered the potential use of smart-meter data for regulatory assessments and evaluation. Future workshops will explore the public-interest benefit which smart-meter data could bring to energy and net-zero planning of devolved, regional and local government (10 July); and, last, the public-interest benefit which smart-meter data could bring to improving analysis and insight on heat-decarbonisation (11 November tbc).

Overview of second PIAG workshop – 23 April 2020

This second workshop focussed on regulatory uses of demand side data: how better data could enable Ofgem to do its job better - monitoring the market and developing the regulatory framework – and also how it could help companies meet their regulatory obligations.

The current and future sources of demand-side data used in the regulatory space are essentially those discussed at the first workshop. Ofgem is beginning to collect and share more data on the energy market more generally – for example on patterns of switching between suppliers. However, the main demand-side data-source currently available to Ofgem remains the BEIS sub-national energy statistics (and, based on that, the NEED database which includes some basic demographic and property information). What this provides is simply a single kWh consumption figure submitted annually to BEIS for every meter by energy suppliers. This provides the main data-input on the consumption-side to the Ofgem typical domestic consumption values (TDCVs), on which much Ofgem and industry analysis depends. The ESO is similarly dependent on that data source for creating the Future Energy Scenarios which again underpin much of the industry thinking on developments in the energy system.

Other customer-side data-sources available to Ofgem include a patchwork of historic pilot data (in particular EDRP, CLNR and Low Carbon London – but also SSEN's more recent SAVE project). Going forwards the UCL SERL database will be an important addition. These sources are described more fully in the report on workshop 1.

The only additional source of demand side data discussed in this workshop was the activity data used by Jacopo Torriti based on an extensive diary survey. While this does not directly provide energy consumption data it can provide useful insights into consumer behaviour by time of day, linked to demographics.

Although not discussed in detail at the workshop a further source of data that Ofgem draw on in their work on distributional impacts (published following our workshop) is the ONS Living Costs and Food Survey which provides survey information on expenditure on energy linked to quite extensive demographic information. This allows an estimate to be made of consumption using assumptions on price.

Building on this review of existing data sources, the workshop focussed on a number of user cases aimed at showing where the gaps in understanding currently are and how access to aggregated / anonymised smart meter data could provide additional public interest benefits.

Ofgem provided an over-arching framework which other user cases fitted into well. As they set out, data (obviously) plays a fundamental role in Ofgem's analytical work, including:

- Appraising different regulatory options in advance – particularly in impact assessments.
- Monitoring the energy sector, and identifying emerging problems. (e.g. annual State of the Market Report, Wholesale and Retail market indicators published in the Data Portal).
- Evaluating their work, and enabling others to do so.

Greater access to smart meter data could produce benefits in all of these areas, grouped under four themes which are used as the structure for this paper:

- **Vulnerability:** Enabling more granular assessment of the likely impacts of different options on vulnerable consumers, and thus better regulatory decisions.
- **Flexibility:** More detailed analysis of potential effects of e.g. charging reforms on system flexibility and decarbonisation.
- **Efficiency:** Allowing better monitoring and forecasting of potential inefficiencies in the system, such as future network constraints.
- **Evaluation:** Quicker and better-informed evaluation of Ofgem's interventions, including by external experts. Also scope to explore data with stakeholders in different ways – e.g., hackathons.

There was also a strong fifth theme that emerged around the need not just to think about today's challenges but to look ahead to try to anticipate future requirements as industry players themselves have more data. The ability to identify **emerging market issues** would require Ofgem itself to have better data. Examples were also given of where Ofgem had been able to draw new insights from other sources of data, through machine learning for example. It is not always possible to identify what the benefits of sharing data will be but there is a growing recognition, through the Energy Data Taskforce, of the benefits of making data more widely available.

Theme 1 - Vulnerability

It is a clear priority for Ofgem to understand the impacts its policies have on different groups and three of the speakers had been involved in different ways in trying to help Ofgem do this – but all highlighted the challenges of trying to do so effectively in the absence of decent data.

Ofgem explained that it now has 13 Consumer Archetypes (recently refreshed by the Centre for Sustainable Energy) to help it look beyond just the "average" GB energy consumer. These Archetypes are rich in socio-economic data from ONS and Ofgem surveys. However, in the absence of smart meter data, they cannot robustly reflect consumers' consumption profiles nor their behavioural response to policy changes. Ofgem have to use proxies as a way of getting some insight on these aspects. This was acknowledged by Ofgem as creating potential issues given the increasing policy focus on customer flexibility and potential demand-side response (eg in looking at the impacts of Half Hourly Settlement).

For the future Ofgem would want to have more data on time of use patterns of different customer groups – including the ability to identify behavioural changes. This should also help them in better identifying potential problems. It could also provide the scope for more targeted interventions.

Since the workshop Ofgem have published updated information on their approach to distributional impacts (available here) alongside more detail on the Archetypes mentioned above. In that they say: *"This framework .. provides us with a strong foundation, but we recognise there will be opportunities to develop it further in time. For example, the framework cannot easily help in assessing [the] impact on consumers where it depends on their usage throughout the day. However the roll-out of smart meters will provide us with more information on consumers' consumption profiles and how they use their energy day-to-day. If we can match this with our socio-economic data, it will help us understand how different groups of consumers may respond to our policies in future."*

Frontier Economics gave the example of the analysis they had done for Ofgem looking at the Targeted Charging Review. This work attempted to assess both the static and dynamic impacts of changes to network charges, on different levels and shapes of consumption across the broad domestic and I&C groups and at different voltage levels. For domestic customers they had looked at high, medium and low users, high economy 7 users and then medium users with various low carbon technologies (solar PV/storage, heat pumps and EVs). They had taken typical consumption values from Ofgem's TDCV analysis (which draws on BEIS NEED data) and then estimated peak usage using profile data from Northern Powergrid's CLNR pilot for customers with similar overall consumption levels (or known to have the particular low carbon technologies). The modelling also included I&C customers which created further challenges as they are a heterogeneous group.

Modelling the impacts resulting from changes in network charges (after taking account of behavioural shifts) is complex and is dependent on a range of factors including suppliers' decisions on whether to pass through changes in the structure of network charges into end user tariffs and customers' response to these revised tariffs.

More generally it was acknowledged that the socio-demographic data in the pilots was relatively weak (in some cases using only postcode level data).

Frontier's view was that having a wider set of half-hourly data might have provided more regional insights. However, to make the results tractable, the analysis had focussed on a relatively small number of consumer archetypes. They were not persuaded that having access to a much larger dataset of consumption profiles would in itself have made this particular type of analysis significantly more informative (as the dataset used was sufficient to produce the number of archetypes required). However, it would have permitted a more detailed look at regional distributional issues – for example caused by different requirements for heating or the potential for solar PV in different parts of the country.

Access to more detailed consumption data that was coupled with socio-demographic information may also have allowed Ofgem in its analysis to have explored the distributional impacts in more depth.

In particular, consumer groups and Sustainability First had raised concerns through the Targeted Charging Review about the impact of the reforms on low income customers who on average have lower consumption[2] and in their opening remarks Ofgem had acknowledged that it was in part the

[2] This response from Grid Edge Policy reflects the conclusions of a roundtable hosted by Sustainability First and attended by a number of consumer bodies

debate around the TCR that had convinced them of the need to do more work on distributional impacts and to update their Archetypes.

Moreover, for the Targeted Charging Review only total consumption was needed to assess the leading options for domestic customers. Understanding time-of-use patterns would be more critical for other policy decisions such as the distributional impacts of half-hourly settlement. It is notable that in their draft impact assessment for market wide half-hourly settlement (published after the workshop) Ofgem said *"It is challenging for us to set out a quantifiable distributional impact analysis across different categories of domestic consumers without a robust evidence base".*

Use of survey diary data

Professor Jacopo Torriti (University of Reading) had used diary survey data from the UK Time Use Survey (UKTUS), based on a sample of around 11,000 participants, to try to determine energy consuming activities by time of day for different socio-demographic groups as input to Ofgem's work on half-hourly settlement[3]. This had yielded some insights but without actual consumption data the bill impacts were speculative.

His analysis took data from UKTUS and clustered households based on the probability of them undertaking particular activities at peak time. The data includes a wide range of property data and household demographics (family type, employment, income, age).

To make the link with consumption he had also looked at half-hourly electricity consumption data from the CLNR pilot by income. He had also looked at the pattern of peak power usage by Acorn segmentation using data from the Low Carbon London pilot. The graphs presented at the workshop show that for these broad demographic categories there is much more variation in usage patterns *within* groups than there is *between* groups. This suggests that to get a handle on what drives usage profiles you need to look at a much wider set of social factors, habits and other determinants of usage patterns than traditional demographics[4].

Jacopo summed up by comparing two pieces of research that he had carried out to inform different Ofgem assessments and which highlighted the trade-offs that had to be made in the absence of robust granular consumption data linked to rich socio-demographic data.

1) Assessing the distributional impacts of Time of Use tariffs as part of the Impact Assessment on Market-wide Half-Hourly Settlement Reform as discussed above: This used a non-energy dataset with rich socio-demographic characteristics using quantile analysis (a form of regression). This provided useful clustering around patterns of activities by time of day linked to demographics but lacked metered energy data.

2) Estimating core capacity (effectively peak demand) for different segments of residential consumers on behalf of Citizens Advice / SSEN as input to the Access and Forward-Looking Charges Significant Code Review[5]: This used granular smart meter data for particular customer groups from various pilots including in particular the SSEN SAVE project (which included smart meter data as well as diary recording of activities). This provided useful clustering of patterns of energy usage but with limited socio-demographic data.

[3] The paper including a full description of the methodology is available here
[4] This conclusion is also reflected in a Grid Edge Policy report which draws on earlier Sustainability First research - here
[5] Further detail available here

Finally looking to future alternative sources of data there was discussion around the Archetypes developed for Ofgem by CSE (now published) but with general recognition that whilst a valuable tool they still remained a poor substitute for a fuller database linking half-hourly consumption and socio-demographics. At present these Archetypes do not include any data on the profile of usage but, as noted above, Ofgem has said that, when such data becomes available, they will update the Archetypes.

The other key development is the UCL SERL database which will ultimately hold around 10,000 customers' half-hourly usage, regularly updated, alongside demographic data drawn from a survey completed by participants, with the potential in future to link to other government data for example. It was acknowledged that this would be a very positive step enabling a much more robust analysis of distributional impacts. UCL committed to aggregated or anonymised versions of the data being made available beyond the academic community to allow wider engagement by NGOs and others in the policy debate.

However even with the SERL data there will remain gaps – in particular the dataset is not sufficiently large to provide local /regional breakdowns (to be explored more at the PIAG workshop on 10 July 2020). It also does not have the capacity to readily track the impacts of targeted interventions / trials if the individuals are not part of the SERL dataset, although the facility will exist for academic researchers to add trial participants to the SERL database.

As the energy transition leads to a greater emphasis on policy decisions affecting the demand profile through the day, with new low carbon technologies impacting these profiles in many different ways, understanding the impacts of such trends and policy options on different socio-demographic groups will be increasingly important. The relationship also ceases to be quite as simple as it is with overall levels of consumption which tend to increase with income for example. Patterns of consumption will be driven by a wide combination of factors (eg working household out-all-day, single mother with young children, not working) that are not yet well understood.

There was a broad consensus that smart meter data would be crucial in enabling the distributional impacts of policies to be assessed going forwards. Citizens Advice cautioned about the need to maintain privacy and the need for balance was accepted by the group. The PIAG phase 1 work had sought to resolve this tension by identifying routes to data that already existed in the system, or would be available in future, and having a trusted processor (eg ONS) produce aggregated or anonymised datasets so that privacy would be protected.

Theme 2 - Flexibility

As set out in the Ofgem and BEIS Smart Systems and Flexibility Plan, a key strand of the energy transition will be the increased importance of flexibility and demand-side response. As a part of this Ofgem set out how they are looking to build their understanding of the potential for demand-side response and how that is best delivered. Some excellent work has been done by academics to feed into this but often based on relatively small-scale datasets.

For example, the work that Professor Jacopo Torriti has been doing using diary survey data (discussed above) is important in understanding activities and electricity demand through the day, which of those activities are time critical and hence the potential for flexibility. However, as noted above, it does not provide robust energy consumption measures.

There is a particular interest in how customer flexibility can be used to improve balancing with renewables, particularly at a local level, reducing the costs of electricity generation. There is also interest in how best use can be made of smart network management, demand turn-up and battery storage. Finally, there is interest in testing the range of interventions that could be used to stimulate demand-side response from tariffs to simple prompts to reduce consumption, with variants in both cases around whether initiatives are opt-in or opt-out and how they are presented. In all cases, improved consumption data is needed to help in building understanding and evaluating potential interventions.

The caveats highlighted above about the existing and upcoming data sources apply here as well. Moreover, the major pilots on low carbon technologies and demand-side response (CLNR and Low Carbon London) had a regional focus and are possibly also now somewhat historic.

Given the importance of pushing ahead on flexibility there is a need for larger-scale, more regularly-updated data to provide a more robust understanding of the potential (including at a regional level and linking to the presence of other low carbon technologies). There is also a need for an enhanced ability to carry out trials of interventions and to evaluate their impacts.

Theme 3 - Efficiency

In reflecting on how better data could improve the efficiency of the energy system Ofgem reflected on how the Future Energy Scenarios developed by National Grid ESO were top down and that more 'bottom up' modelling was desirable, to enable better forecasting of future network dynamics and greater confidence in distributed solutions. Ofgem also referred to the companies who run the system needing access to better data to efficiently manage their assets. These themes were reinforced by the company presentations. Frontier also talked about how Ofgem might be able to use more granular demand side data to enable better benchmarking of performance of the distribution networks as a part of their efficiency assessments in the RIIO price controls (as well as some of the downsides).

This aligns with the growing focus on open data through the Modernising Energy Data initiative involving BEIS, Ofgem, Innovate UK, the Energy Systems Catapult and ONS which seeks to improve on current approaches to access through greater standardisation of data-formats and visibility of market-facing data – to drive greater market transparency and efficiency.

ESO use of more granular demand-side data

The ESO had presented on its use of government consumption-side data at the first workshop and the report on that workshop provides an account of how access by the ESO to more granular demand-side data could helpfully support their role. This includes how the FES scenarios, on which much further industry analysis depends, might benefit if customer-side data-inputs could become less 'average'. Given the centrality of the ESO's role in the broad regulatory landscape the ESO was invited to recap once more on the main messages at this second workshop.

The ESO has a number of distinct roles, all of which are dependent on the availability and quality of data that it has. These include:

- Second by second real-time forecasting and grid operations: design and update of decision support tools.

- Long term forecasts (Future Energy Scenarios & System Operability Framework): used out to 2050 to inform GB system planning, anticipate operational changes, investment, regulatory development and lead debate.
- Regional investment planning (Electricity Ten Year Statement, Network Options Analysis): used in 10-year and 20-year planning horizons for network planning, build and maintenance.

In particular access to more granular demand-side data (aggregated / anonymised) could support better annual forecasting by enabling the ESO to take greater account of usage patterns in homes with e.g. EV, storage, solar or heat pumps.

Data is needed at a national level but also at a regional and sub-regional level, covering different time slices – monthly / daily / hourly. Uses would include maximum demand planning, operational and seasonal decision making.

The ESO also made the case for greater public access to aggregated / anonymised demand-side data. This would allow different organisations to do their own forecasting and analysis and would stimulate healthy debate on what are important public policy questions. Exposing data to scrutiny by academia and consultants would allow the latest analytical techniques to be employed to draw out insights and learning from the data, ultimately enabling industry to improve its efficiency.

DNO use of more granular demand-side data

The Data Access and Privacy Framework (DAPF,) put in place by BEIS to provide reassurance to customers around the smart meter rollout, allows DNOs to have access to customer demand data, without their consent, for the purpose of fulfilling their regulatory duties, including their duty to maintain an efficient and economic network. The DAPF requires distribution network companies to prepare a privacy plan setting out how they will aggregate and anonymise the data as far as possible and Ofgem has to approve this plan before the companies can access consumption data that is more granular than monthly.

Northern Powergrid described how they envisaged being able to use demand side data to help them in planning and operating their network more efficiently. In particular they believe that having access to the individual customer data is important to enable them to look at the impacts of alternative network configurations over time – rather than just having access to data aggregated in line with the current network structure (ie by feeder) as other DNOs have done. The ability to reconfigure the network to cope with changing patterns of demand with the rise of EVs, the electrification of heat and increased solar PV/ storage is important in terms of maintaining an efficient and economic network.

Northern Powergrid had considered how best to protect the privacy of this more granular data and had been in discussion with Ofgem about their draft privacy plan for some time and submitted a draft plan in December 2018. Despite some encouraging signs it was now regretfully evident that, regardless of the merits of the case, the legal framework did not permit Ofgem to approve a plan on this basis.

Northern Powergrid were therefore now looking at following a path more aligned to the approaches taken by other DNOs whose privacy plans had been approved. In particular it was noted that UKPN whose plan was approved just prior to the workshop in April 2020 were looking at holding data aggregated to feeder level (and using data masking to avoid any issues where there were small

numbers on a feeder) but would also hold summary metrics such as maximum demand and total monthly demand at an individual MPAN level.[6]

It was also open to Northern Powergrid to download data again in future and aggregate it on a different network configuration but this is clearly a less efficient solution and would not provide a full historic picture in the way that building up their own database would do (given that eg only 13 months of consumption data is held on the meter).

It was acknowledged that proceeding down the path that other DNOs had taken would enable Northern Powergrid to start to make use of smart meter data as SMETS2 volumes built up in the North. However, the question remained as to whether the DAPF was proving unduly restrictive as the challenges facing DNOs became clearer.

The PIAG 1 report had identified DNOs as a potential nation-wide source of smart meter data subject to approval of their privacy plans and coupled with a new obligation on them to take on such a role. The DNO role in developing local area energy plans will be considered further in the next workshop.

Benchmarking DNO performance

Frontier explored the potential for more granular consumption data to help Ofgem in benchmarking whether DNO spending is efficient as part of the RIIO price control process. To avoid incentives that could result in DNOs spending too much on some types of cost and too little on others, benchmarking is often best carried out at the level of total expenditure (totex). Ofgem would typically then use high level drivers of costs (e.g. overall peak demand) to enable comparisons to be made between networks. The choice of drivers is important as using an inappropriate driver can drive inefficient behaviour. In particular, Frontier highlighted the following issues:

- **Temporal issues:** If the peak period is not coincident across the network, the level of peak demand for the network overall may not reflect constraints at the local level which is what will actually determine the investment required.
- **Spatial issues:** Issues that require spending can be highly locational (and this will increase as a result of e.g. clustering of heat pumps and EVs). An overall peak demand measure does not take into account whether demand is concentrated in certain sections of the network, or more widely dispersed. Similar to the issue above, this may mean that two networks that look very similar in terms of overall peak demand can have quite different costs.
- **Use of flexibility**: DNOs can increasingly shape peak demand through contracts for flexibility. In the long-run, a regulatory regime that uses peak demand as a determinant of revenue allowances may disincentivise this. This is since, by lowering peak demand, the DNO would potentially be lowering its cost allowance in the future.

To design better regulatory cost-drivers and incentives that reflect the changing network cost drivers, Ofgem could use granular smart meter data if it were available. For example, Ofgem might consider:

- Building a more accurate cost-driver based on customer-level demand. E.g. calculate maximum demand in each small area / portion of the network, and aggregate up to produce a DNO-level metric.

[6] Since the workshop SSEN have also had their privacy plan approved based on a minimum of 5 MPANs being aggregated on a feeder

- Using detailed data to model a "synthetic peak" reflecting an estimate of what the peak *could* be after DSR which could then be used as the cost-driver to ensure that companies remain incentivised to make use of contracts for flexibility and allow Ofgem to take a more informed view on reinforcement allowances in particular.

Such approaches are not easy – they need both engineering and economist input to ensure they do not themselves create mis-incentives - but above all they need much better consumption-side data (more granular, more locational).

As discussed above there remain issues with DNOs getting access to the data to enable them to carry out such analysis themselves and to plan their networks appropriately taking account of the temporal and locational variations that are likely to become more prevalent. However, once their privacy plans are approved, they will be able to access this data and to exploit any regulatory incentives that exist. As set out in the PIAG Phase 1 report it is important that the regulator also has visibility of this same data to avoid "flying blind" as it seeks to regulate the networks.

Frontier also noted the potential for smart meter data to improve the picture around service quality. Granular (by customer and half-hour by half-hour) data could give a more accurate and more timely picture of outages, voltage disturbances etc

Monitoring behind the meter energy consumption could also help networks and Ofgem to understand where generation and other distributed energy resources are located.

Theme 4 - Evaluation

An important part of Ofgem's role involves monitoring the market and identifying and evaluating market interventions.

Evaluation and Randomised Control Trials carried out by Ofgem (eg the CMA Microbusiness Price Transparency Remedy and Cheaper Market Offers Letter Trials) have to date relied on target-group surveys, mailings, and specific Requests for Information to suppliers.

This has led to inevitable lags in data availability and placed significant demands on resources across all parties (for collating data, quality control and analysis). Although Ofgem has been working on improving this process through their Data Services Team, challenges remain.

Ofgem are of the view that improved access to data in future would provide a greater ability for academics and others to assess and challenge their work rapidly and would allow them to do more, more cost effectively in-house. Readier access to data would also allow regular feedback of evaluation evidence into decision-making processes.

Frontier gave some specific examples of how improved data could be used in assessing the impact of interventions. In particular before-and-after data could be used to assess interventions designed to reduce or shift energy consumption, including ex-post distributional impact assessments.

Theme 5 – Emerging market issues

CSE highlighted that as industry participants (suppliers, network operators) get more data they will make use of that data to gain advantage in the market or to exploit regulatory incentives. It is important that Ofgem stays one step ahead. The question is not just what benefits smart meter data

could deliver but what risks there are if Ofgem does not have access to a greater level of smart meter data. Understanding developments in the market through access to granular consumption-side data will also be important for others with a regulatory remit (CMA, Energy Ombudsman) as well as those with a wider public-policy role (BEIS, Citizens Advice).

From an Ofgem stand-point, examples would be:

- How could Ofgem check whether energy suppliers are 'treating their customers fairly' (e.g. offering them appropriate and fairly priced Time of Use tariffs)?
- Likewise for other market players like aggregators – and whether they need regulating to protect consumers from sharp practices (and how would BEIS or Ofgem know how sharp they are)?
- How would Ofgem and BEIS assess the distributional implications of new market behaviours and the potential value of options to mitigate 'unfair' outcomes?
- How will Ofgem know if DNOs/DSOs are responding appropriately and efficiently to newly revealed patterns of demand across their networks?
- How, without the data, can the public interest be understood and represented by those with responsibilities for policy design and regulation for the energy system in the face of commercial interests who do have access to the data?

Ofgem confirmed that they recognised the importance of looking at the data that would be needed to deal with potential future challenges not just those that they can identify today.

An inevitable consequence of this is that one cannot necessarily say precisely at this point what the benefits of improved access to data may look like in their totality.

Ofgem gave examples of how in their wider oversight of the market, using new data flows has yielded new insights eg on switching patterns. Visualisation of electricity supply switching (using Electralink data from the DNOs) based on "from *supplier size*" → "to *supplier size*" has allowed Ofgem to drill down to see which company an energy supplier's customers are switching to. Moreover, data is now being stored in ways which allows Ofgem to join across datasets (market share, gas switching, electric switching, complaints etc) and carry out a high-level compare-and-contrast. Building on this, machine-learning was used to identify attributes of long-term disengaged customers that could be used to predict long-term disengagement.

Visualisations of this data in a way that avoids breaching commercial sensitivity (or privacy) are made available on the Ofgem data portal where they are considered useful for consumers and industry, as part of Ofgem's commitment to open data.

A further extension of this was the Hackathon that Ofgem ran where they collated data that had never been linked together before. An open invite was issued to industry, universities and others to look at issues around meeting the Net Zero target. The winning team linked Met Office data to the Electralink embedded generation data to show that in terms of their location, solar panel installations do not correlate with the amount of sunlight in the UK[7]. Future Hackathons are planned including one on heat decarbonisation.

[7] https://www.ofgem.gov.uk/publications-and-updates/climate-change-emergency-hackathon

In summary applying machine learning has yielded new insights from combinations of data and involving others through Hackathons has brought in creativity and innovative thinking. Our assumption would be that with access to granular smart-meter data new insights would similarly be possible.

Making such data widely available would also provide an opportunity to improve transparency and accountability around Ofgem's use of evidence to support decisions. Consultations on policy and regulation are intended to be as inclusive as possible and there are a range of audiences that would be better empowered to promote the public good if Ofgem were able to be open with how it has processed and interpreted smart meter data to inform its consultation positions and decisions. Sharing data in this way would, as ever, have to be limited by security and privacy considerations.

One very practical example of the potential of granular consumption-side data to deliver unexpected benefits emerged at the end of the workshop. Various parties (including Smart DCC and Electralink) highlighted potential data sources they had that could help shed light on aspects of the current Covid-19 crisis, including on critical issues such as where customer load had reduced most or whether pre-pay customers were self-disconnecting more. BEIS confirmed that they would ensure that these offers in support of pandemic-handling were fed back to the relevant team in the department.

Conclusion

In conducting its regulatory evaluations and assessments, Ofgem, as the national energy regulator, has no 'privileged access' to granular customer-side data. This represents an inevitable challenge for the body charged with oversight of the energy markets on behalf of current and future consumers – in particular when those it regulates do have access to such data, further increasing information asymmetry. Through the workshop strong evidence was presented as to how access to half-hourly smart-meter data could enable Ofgem to carry out its regulatory functions more effectively and to enable the industry it oversees to operate more efficiently. Without such data there are risks that market issues will go undetected.

In terms of policy development a particular challenge exists around the assessment of distributional impacts which requires the linking of granular time-of-use smart-meter data and rich socio-demographic data given the factors that determine time-related patterns of usage are complex and not well understood, but of increasing importance in the context of a smart, flexible energy system.

Another gap highlighted at this workshop – as it was at the first workshop – was a lack of detailed visibility for the ESO of consumption-side data below the transmission-level. Questions were also raised as to whether the bar was set too high for distribution networks seeking to access the data to help in planning and operating an increasingly complex network.

As noted above (and throughout PIAG 1) there is a clear recognition that there is a need to balance the public interest in these wider benefits with potential privacy concerns, as Citizens Advice continue to highlight. Inevitably for data analysts the more granular consumption-side data they can access the better – but all parties recognise privacy is a key consideration.

As noted above, in PIAG 1 the models presented focussed on delivering the benefits without compromising privacy through use of a trusted processor (eg ONS) to aggregate or anonymise the data.

A core message throughout PIAG and reiterated at this workshop was the need not to let the best be the enemy of the good. As discussed, linking smart meter data with socio-demographic data is crucial for assessing distributional impacts and the SERL data will provide a valuable additional source of evidence in this area. While it will not be a large enough sample to enable regional analysis, many of the use cases requiring local area data are not expected to need the same level of socio-demographic data (to be confirmed at the next workshop). With more DNOs now having privacy plans approved there could be a route to access aggregated / anonymised data as an input to local area energy plans for example. DNO data could also provide some of the insights that the ESO has identified would be beneficial.

Given the importance to Ofgem of having access to more granular consumption data to keep up with those it regulates and to fulfil its role going forward, there is a need for an open dialogue between BEIS and Ofgem as to how this is best achieved.

Finally, data more widely has attracted significant focus through the Energy Data Taskforce and now the Modernising Energy Data initiative. Sustainability First has been emphasising the important role of data in driving wider change – in assessing distributional impacts more broadly, in measuring companies progress in de-carbonisation as part of RIIO and in establishing metrics by which one can judge companies' progress towards becoming a responsible business as part of its Fair For the Future work. Enabling wider access to aggregated / anonymised smart meter data should be seen as part of this bigger debate.

Annex: Summary against the four key questions for PIAG 2 – Data for Regulatory Purposes

What demand side data is available now?

BEIS sub-national energy statistics including NEED data

Pilots (CLNR, Low Carbon London)

Industry data flows (switching, meter installations)

Proxy data – diary surveys, energy expenditure (in the ONS Living Costs and Food Survey)

What data is expected to become available in the short to medium term?

UCL SERL data – underlying data only accessible by accredited academics but outputs to be made more widely available

HHS and DNO data – but no plans as yet for enabling wider access

What gaps does this leave that smart meter data could fill?

TOU profiles linked to demographic data – summary information likely to be provided through SERL but sample size will not allow regional analysis

Ability to reconcile top down national consumption data with SERL sample to ensure representative

Database of individual usage to track trends – linking to tariff / supplier

Database of individual usage to support trials / interventions

What are the public interest benefits in filling these gaps?

Clear that decisions are fairer (in terms of distributional impacts)

Opportunity for special interest groups to engage and advocate for fairer solutions

More efficient network investment decisions by companies

Ability for Ofgem to regulate through price controls to drive greater efficiency as distributed resources and flexibility become more important

Ability for Ofgem to identify market abuses and wider issues as time of use tariffs become more prevalent

Greater ability to carry out controlled pilots (with a pre-existing assessment tool rather than having to do bespoke data collection) – improved Ofgem ability to test interventions