Integrating Renewables - The Future of Network Charging

Working Paper

19 August 2018

Maxine Frerk – Visiting Fellow

Oxford Martin School – Integrating Renewables Programme

Oxford University

Citation: Frerk M (2018) Integrating Renewables – The Future of Network Charging. Working Paper Oxford Martin School

Acknowledgments: I am indebted to the Oxford Martin School Integrate programme for its support and for the extended discussions we have had on network charging. I would particularly thank Nick Eyre and John Rhys for their thorough comments on an earlier draft of the paper.

Integrating Renewables – The Future of Network Charging

Summary

In a world where it is accepted that use of electricity networks will change radically to accommodate more distributed generation, it naturally follows that the basis of charging for use of those networks must change too. At a high level the purpose of network charging is to enable the network company to recover its costs, as determined by the regulator, and to send signals to users and to the network about efficient use of those assets and the need for future investment.

There is broad consensus among academics and practitioners around the principles that should be applied in setting charges. While cost reflectivity is important there are other criteria such as simplicity, predictability and fairness.

What has been less debated is the basis for making tradeoffs with regulators able to fit whatever ideological approach they want to take into this broad framework of principles. However with the move to a more flexible energy network the debate on network charging has risen in prominence across all jurisdictions – and the need to consider these tradeoffs has become more pressing.

Another aspect that has been less debated is how the shift to a more flexible network with more distributed energy resources changes the basis for cost reflectivity. Historically network charging has been seen as a problem of how to charge where there are low marginal costs in the short run and a resulting focus on long run marginal costs (LRMC) which commentators often claim will deal with both short run and long run efficiency. Looking ahead, with more operational options open to both networks and users, more thought needs to be given to the short run costs and the network models used to determine LRMC need to be revisited to take full account of the effect of more renewables on the system.

Where costs can be linked to underlying drivers then it is generally acknowledged that the critical tradeoff to be made is on the level of granularity (temporal and spatial) to be used. More granular charges will be more cost reflective but there are practical challenges around implementation and the ability of end consumers to understand and respond to such charges (at least absent automation). There may also be issues of fairness in a world where for good policy reasons charges have always been socialised across for example rural and urban areas.

Understanding how to make those trade-offs requires evidence on how consumers will respond to such charges in practice, recognising for domestic customers at least that there are a range of behavioural factors that need to be taken account of in tariff design. It is reasonable to expect that the appropriate level of granularity will increase over time and hence part of the challenge is actually to identify the best transition pathway. In determining how much effort to put into developing highly granular tariffs it is also important to recognise where these signals may anyway be distorted by policy costs, for example, and to what extent these signals will actually be passed on by end retailers where there is unbundling of provision.

Where costs cannot be linked to underlying drivers, there will be what is termed a “residual” element that is needed to allow full cost recovery by the networks. The principle of cost reflectivity points to these costs being recovered in a way which minimises distortions in the underlying cost signals. This is the idea behind Ramsey pricing where such costs are allocated with the inverse of the elasticity. However again thought needs to be given to implementation practicalities and issues of fairness. More fundamentally it needs to be acknowledged that in principle the “residual” element

can be used to achieve wider public policy goals such as to encourage energy efficiency, provide support for renewables or help for the fuel poor.

The same principles that apply in relation to recovery of the residual apply in relation to recovery of policy costs (such as renewable subsidies). A valuable spin-off of the debate on network charging should therefore be to bring some rigour into the debate around how these costs are recovered. Historically the arguments for recovering policy costs through taxation has been cast in terms of the regressive nature of recovering such costs through bills. However there is also a strong economic argument that recovery of these costs through bills distorts the price signals and can lead to economically inefficient outcomes.

This is seen in practice with the growing concerns about grid defection as, in particular, small scale solar and storage in combination allow customers to go off grid. This is a real and live issue in places like Hawaii and Australia. The concern is that where customers go off grid either fully – or simply reduce imported energy as happens in GB – the level of residual costs does not reduce and hence other customers are left to pick up an increasing share. It is this pattern that has been referred to as the “death spiral” of the grid – with rising charges encouraging yet further defection. This problem can be caused by both how network residual charges are recovered but also by the inclusion of policy costs in the bill.

Recognising this challenge, there has been growing consideration of the range of different bases that could be used for charging from the flat charge per Kwh used today to time of use usage charges (static or dynamic), standing charges (again either flat or linked to capacity – which may be peak coincident capacity or the user’s own maximum demand). As noted above there are also questions about the level of locational granularity to apply – and the potential for charging on the basis of locally matched supply and demand (using virtual MPANs). There is a developing consensus that a shift to more capacity based charges (eg linked to “fuse size”) can be simpler, fairer and less distortive -but the transition path will be important where customers are not used to such charges.

More fundamentally the question has also been raised (by MIT) as to whether the use of property taxes as a basis for charging would actually be more aligned with Ramsey principles. Such a solution would obviously benefit those in fuel poverty and, as noted above, should apply both to network charges and policy costs. It merits proper debate.

More broadly, although the focus of debate has been on the structure of charges there is a need to widen this to include the basis for connection charging and the scenarios in which paying for the provision of ancillary services may be more effective than trying to send very location specific signals through network charges for example. Ofgem have now acknowledged this in their Regulatory Strategy but the debate is still at an early stage.

In trying to identify how important cost reflectivity and “non distortion” are as principles compared to other criteria, the critical question is what distortion are you worried about and how realistic a problem is it.

One aspect of “non distortion” that is important in the network charging context is to ensure that the choice of whether to connect at transmission or distribution level reflects the true relative costs of the two options. This requires more joined up thinking across transmission and distribution including on the approach to connections which is currently very different for the two systems in GB.

Equally, as we look to decarbonise heat, avoiding distortive price signals between gas and electricity will become more important but is not yet on policy makers' radar as the focus is essentially on electricity.

Looking internationally a range of different approaches are being pursued which reflect the different geographies, policy priorities and market structures. Nonetheless at such a time of transition there is much to be learned from continuing to track international experience.

Finally there is an important question about how to bring a consumer perspective into what is a highly technical debate. In its RIIO approach to price controls, Ofgem requires network companies to engage with consumers in developing their business plans, but this does not include the structure of charges. If radical changes are to be made to charges there will be winners and losers. Consumers – and in particular vulnerable consumers - need a stronger voice in this debate.

Introduction and Context

Much has been written on how networks will need to change to accommodate the shift to more distributed renewable generation. No longer will distribution networks simply be carrying electricity out from a central power source on the transmission network to passive consumers at the ends of the network. Increasingly distribution networks will have to cope with two-way flows, varying through the day and significant new loads in the form of electrical vehicles and heat pumps. The DNOs will have to actively manage their networks to cope with constraints in real time, taking on what is in effect a system operator role.

At the same time the demand side is going to have to become more engaged. Rather than electricity always being there to meet demand, demand will start to play a role in helping keep the system in balance as generation itself becomes more intermittent and non-dispatchable.

With such radical changes in the use of networks it is inevitable that network charges will need to change to support that transition. Debates about network charging are taking place across the globe and although this paper focuses on the GB position it does touch on thinking in other jurisdictions as well.

A particular motivation for reviewing network charging – the "burning platform" if you like – is the fact that increasingly customers in some parts of the world customers are starting to go "off grid" in order to avoid these charges (and other policy costs / taxes that are recovered through the bill). More common currently is what might be considered as partial grid defection – or load defection - where consumers with their own generation pay less in such charges. In both cases this leaves the remaining customers paying an ever-increasing share in what has been termed the "death spiral of the grid". This impending crisis has forced regulators and others to go back to first principles on network charging which should help deliver a structure that supports the transition to a low carbon energy system in a way that promotes competition and is fair, in particular not disadvantaging those on low incomes.

There is a broad consensus on the principles that should underpin network charging but no real debate on how the tradeoffs should be made between competing criteria. The purpose of this paper is to provide a framework for how these tradeoffs should be considered.

At a practical level, there are then a range of different structures that can be adopted for network charging. The paper sets out the front runners for a system of network charging in GB.

In particular this paper:

- Sets out the background to network charging generally and in GB specifically (including the distinction between cost reflective and residual charges);
- Sets out the generally established principles for network charging;
- Explores the issues around cost reflective pricing from an economic and engineering perspective;
- Explores how tradeoffs have been made in the past by reference to a number of case studies– including looking at the risks around grid defection;
- Draws on these to provide a framework for thinking about the trade-offs between the principles;
- Identifies the range of potential tariff structures and which would seem to be the front runners;

- Notes other related considerations including connection charging and ancillary services, the need for a system level view including gas, and the need for a stronger consumer voice in the debate; and
- Draws out some conclusions and recommendations.

The background to network charging

In most jurisdictions a distinction is drawn between the transmission network (the high voltage network linking different parts of the country) and the distribution network linking from that to end users (at a lower voltage). Historically generation connected to the transmission network but, with the growth of renewables, an increasing amount of generation is connected at the distribution level (now over 25%) or even resides behind the meter, as with rooftop PV.

Transmission and distribution networks are long term monopoly assets and as such are almost always regulated. Typically the regulator will set both the allowed revenues (intended to recover the networks costs and allowed return on capital) and the way in which those revenues should be recovered (ie what the charges are).

Historically in GB (and as is still the case in some US states) the network carried the volume risk. Its charges were set and it was dependent on the volumes materialising in line with projections if it were to recover its full allowed revenues. Aside from the risk to the network this also creates an incentive for the integrated utility to drive up volumes, counter to ambitions for energy efficiency. In the US many states have now go down the path of "decoupling" so that the allowed revenues for the regulated utility can still be recovered in full with a mechanism for compensation if the regulated charges do not deliver the allowed revenues (or for redistributing in the event of over recovery). This is now also the arrangement in GB for networks.

In GB the regulator sets the allowed revenues through a price control process known as RIIO (Revenue = Innovation, Incentives and Outcomes). It also has a role in approving the charging methodologies.

A particular consideration when looking across jurisdictions is the extent of vertical integration. Where a utility remains fully integrated as in some US states there may not be a separate network charge. The rate making process looks at the rate in the round including wholesale energy costs. Some of the same principles will apply but the priorities will be different. In cases where there is retail competition then the network charges can either be charged separately or bundled in with the retailer's overall charge. In the latter case, as applies in GB, there is no guarantee that any change in network charges will flow through to consumer bills. While in theory a failure to properly reflect the network costs in prices would create market opportunities and that retailer would lose out. In practice – other than for the largest users – this is not a constraint.

Focussing then on a regime like GB where there is uncoupling and a separate network charge, there are two essential roles for network charges. The first is to send signals to users about the costs of using the network so that the decisions they take are economic and efficient (ie they use it where the costs are less than the value they derive) and the network operator builds and runs it in a way that again is economic and efficient (ie they take actions where the costs are less than the value delivered). However given that these are essentially long term assets the cost reflective prices required to send appropriate signals may well not result in the network recovering the full amount of revenues to which it is entitled. There is then what is termed a "residual" charge aimed at ensuring that full cost recovery.

These same points apply in relation to both transmission and distribution networks. In GB the charging arrangements for distribution and transmission are currently considered quite separately but one of the challenges that people recognise needs to be addressed is to align them more to avoid any artificial distortions in decisions on which level to connect at.

The table below gives an indication of the scale of the charges in GB:

2016/17 Charges (£bn)	Transmission	Distribution
Connection	0.2	0.2
Use of system - Forward looking	0.5	4.0
- Residual / cost recovery	2.1	1.4
Total charges	2.8	5.6

Source: Charging Futures (2017) – Ofgem presentation

What this highlights is that:

- Distribution charges in total are roughly double transmission;
- Within transmission around 80% of use of system charges are residual whereas in distribution they account for only 26%.

This is significant given that over recent years Ofgem's policy focus has been almost exclusively on transmission charges (with Project TransmiT in particular). Ofgem's decision in its latest review to focus initially on the residual element of charges reflects that perspective – and the fact that residual charges on transmission have been growing strongly.

However the forward looking element of distribution costs accounts for around half of total network charges. With distribution networks changing rapidly the cost drivers will be changing too and a serious examination of distribution forward looking costs is a priority.

Since this paper was completed Ofgem has published proposals for a full review of distribution forward looking costs (Ofgem 2018). An early draft of this paper was shared with Ofgem and its latest consultation reflects a number of the points made here. However this paper has not been updated to reflect these latest developments.

Principles for network charging

Given that network charging has become a hot topic for regulators across the globe there are numerous reports which present sets of principles that should be applied for network charging.

Much of this thinking can trace its origins back to the original work by Dr Bonbright of Columbia University (Bonbright 1961) which the American regulator's group cites (NARUC 2016). This set out principles as being revenue requirements, the fair apportionment of production costs among consumers and optimal efficiency[1].

In Ofgem's document launching its Targeted Charging Review (Ofgem 2017d), looking explicitly at the allocation of the residual element of costs, it sets out the principles that it intends adopting as follows:

- reducing distortions to underlying cost signals;
- fairness and
- proportionality and practical consideration.

There are a wide number of academic and international studies, plus reports by stakeholders that set out similar lists of principles, with examples in the tables in Annex 1.

In every case the broad headings that Ofgem picks up are covered – with economic considerations and practicality featuring in every instance – and fairness as a consideration in most, with a small number also identifying wider issues as well. These broad headings are considered in turn below.

Cost reflectivity / economic considerations

In broad terms all commentators recognise the importance of sending cost reflective signals to encourage the efficient use of and future investment in the networks. There is a slightly different emphasis revealed in the tables between the European and American approaches. Americans articulate the objective more broadly as economic efficiency, whereas the Europeans talk about cost-reflectivity which is arguably only a means to that end.

Either way the consensus is that this is achieved by setting prices at the long run marginal cost although some commentators do explore this in more depth (and this paper considers it further below).

However, this has to be balanced against a principle that is articulated explicitly by many commentators (and is implicit for others such as Ofgem) of ensuring recovery of regulated costs.

Delivering against these two principles takes you automatically to a third (which is what Ofgem is focussed on in its Targeted Charging Review) that the residual element of costs, which is needed to ensure full cost recovery but where there is no cost driver, should be allocated so as to minimise distortions to the underlying cost reflective price signals.

One specific dimension of "minimising distortions" that is pulled out by some commentators is around avoiding distortions to competition (where users of the network may be competing with each other to provide services).

[1] https://pubs.naruc.org/pub.cfm?id=5388D962-2354-D714-51A8-F5FD79C756F5

Proportionality and practical considerations

While cost reflective prices are seen generally as the theoretical ideal there is again consensus that consideration needs to be given to practicalities.

The range of factors listed in terms of practicality varies a bit (and they are often individually listed as principles rather than being grouped under a broad practicality heading). But broadly they cover:

- implementation costs / feasibility (including availability of smart metering);
- transparency / simplicity / ease of understanding (as pre-requisites for customer response);
- predictability (both to help consumers in planning but also as a pre-requisite for customer response);
- avoiding unnecessary price volatility.

Fairness and other considerations

Fairness is not listed by all commentators and is a difficult concept that everyone may read different things into. Ofgem focuses in particular on the implications for vulnerable customers, as does the GB consumer body Citizens Advice (2016). Others see it as a more general concept.

Professor David Newberry (Newbery et al, 2005) notes that other policy considerations could be relevant if the regulator so chose, citing the example of promoting energy efficiency (though some may see that as another dimension of economic efficiency rather than a separate policy consideration).

Professor Michael Pollitt looks in particular at the impact on richer and poorer customers in the context of a rapid uptake of distributed energy resources (Pollitt 2016).

Industry bodies are more likely to raise other considerations (such as harmonisation with Europe). In contrast regulators typically keep to the narrow economic and practical considerations.

Reflections

In some cases (in particular in the US where companies are still vertically integrated and / or unbundled) the principles are addressed at rate design more broadly not just network charging.

However, in summary there is a strong consensus – across jurisdictions and across interest groups - about the principles that should be adopted in setting network charges (or in rate design more broadly).

There is also a general acknowledgment that some tradeoffs are involved – for example in the level of granularity to which cost reflectivity should be taken - but with only limited discussion on how those tradeoffs should be made.

The main points made by commentators in considering tradeoffs are often around transition paths, in particular:

- MIT (2016) makes a point around the trajectory ie they recommend that you progressively increase the locational granularity of economic signals.
- CEPA (2017) talk about the transitional arrangements that may need to be put in place where significant changes are proposed.

> MIT also talks more generally about tradeoffs and highlights for example that simplifications can be made if there is a concern about equity or other issues with peak prices that are highly granular with respect to location.

The aim of this paper is to try to draw out what might be criteria for use in balancing between these principles – and trying to go beyond a simple statement that CBA should determine the tradeoffs, to look at the factors that would be likely to drive any CBA and hence provide a framework for making such tradeoffs. The considerations that are highlighted include the extent of spare capacity; the nature of choices open to users; the visibility of charges to end users and evidence of the impact on vulnerable customers.

The starting point – cost reflectivity

The core concept around pricing of network assets is that cost reflectivity is important both to allocate what is a fixed resource efficiently (ie so that those who place highest value on the use of it get access to it) and to send signals to the network as to where future investment would be justified.

Developing a view on the approach to cost reflectivity requires both an economic perspective (what sorts of costs should be used) and an engineering perspective (to understand what that means in practical terms for networks).

Most recent work on network charging for a network with greater distributed resources has tended to focus on the problems around the allocation of the residual and to take the current "cost reflective" models as read.

However, the changing demands on the network and the fact that increasingly we expect distribution networks to use flexible "smart" solutions to address network constraints means that the cost structure is likely to look very different going forward and hence a rethink around what constitutes cost reflective pricing is also urgently needed.

The economic viewpoint

The starting point for thinking about network charging is the concept of marginal cost pricing reflecting the longstanding economic principle that this will maximise economic efficiency. However for network industries with high fixed costs this raises a number of issues in terms of the timeframe to use (long run or short run), the lumpiness of investment and how to allow the network operator to recover their full costs. These are considered in turn below.

Short run versus long run

In thinking about the appropriate timescales for the cost signals there is a need to consider the balance of long and short run costs on the networks and the potential long and short term responses from users. From the network's perspective the long run costs are related to investment decisions while short run operational costs include losses and constraint management. From a user's perspective the long-term decisions are essentially about where to locate whereas short term decisions are about levels of usage.

Goran Strbac (Strbac et al, 2005) brings this out in stating that the purpose of cost reflective charging is to send signals to users to ensure that in the short term the system is operated efficiently and that, in the long term, it follows the path of least cost development.

The early work on regulated utilities such as railroads and toll bridges (Hotelling 1937) argued for pricing based in effect on short run marginal costs in order not to discourage economically efficient usage, with any shortfall made good through taxation. The exception would be if there were a capacity shortage where market clearing prices should be used. Over time the debate shifted to a focus on long run marginal costs, with good economic justification but also the political expediency of helping bridge the funding gap (Bonbright 1957).

Network charging in GB is focussed on network investment requirements (with short term balancing and constraint management costs recovered through a separate Balancing Services Use of System - BSUOS- charge). In general, the consensus among economists has been that the appropriate cost signal is the long run marginal cost of incremental usage given that historically costs have been largely fixed in the short run and there have been limited opportunities to influence short term behaviour through price signals.

However with increasing levels of intermittent generation, and flexibility playing an increasingly important role, there will be more short term operational options available and hence a greater need for within day price signals at the distribution level.

In its consultation on the Targeted Charging Review (Ofgem 2017a), Ofgem talks about the need to look at “forward looking” charges in terms of the cost reflective element. As such they do not really explore what the appropriate timeframe for costs is. The example they cite is that if someone can pay less to locate a generator on a part of the network where costs imposed are lower, then both gain – suggesting an implicit focus on the long term. In the subsequent Ofgem paper on Network Access and Cost Reflective Charging (Ofgem 2017c) there is a more explicit acknowledgment of increasing short run costs as part of the move to Distribution System Operators (DSOs).

This interplay of the short and long run price signals is complex and risks distorting signals in its own right if not thought through. Strbac talks about some of the tradeoffs that need to be made including the choice of the timeframe for estimating future demand and generation profiles and the balance between long and short-term signalling.

In particular he says, in a historical context, “Once a decision is made to focus on long term signals in pricing it may not be easy to simultaneously resolve short term efficiency issues. The view is taken that in the distribution system long term investment signalling is more important than short term signalling.”

An example of the sort of problem that arises is if long run costs are recovered in a way that impacts on customers’ short run decisions. Sending price signals which reduce use of the network in the short run is arguably not efficient if it does not result in any short run network cost saving.

In particular, if there is excess capacity LRMC can lead to under-utilisation and inefficiency. This is explored more fully in the case study on gas capacity below.

In its report on the Utility of the Future, MIT (MIT 2016) advocate the use of locational marginal pricing to send short term price signals about local grid congestion, and peak coincident network capacity charges to send long term signals about the need for grid investment. They acknowledge that if you do both there is a need to coordinate to ensure efficient short run and long run decisions.

The idea of locational marginal pricing is more common in the US where the existence of vertically integrated utilities means that the charges reflect the generation operating costs (where the short run marginal costs have historically been significant), while taking account of network constraints. This would be a radical shift for GB (where there is an effective wholesale market) and Ofgem have rejected it as a possible solution in their update on the Targeted Charging Review, arguing it is too complex especially in the context of distribution networks.

However, at a high level, the MIT model suggests a potential route through the challenge of managing log run and short run price signals, which is an important one for regulators to address. If short run signals were to be linked to usage and long run signals linked to capacity (in effect the right to use the network) then this would reflect the timescale for decisions by both the networks and the users, and help resolve potential conflicts.

Lumpiness of investment

Investment in networks tends to be lumpy and indivisible so demand from one customer could trigger the need for major reinforcement. If true long run marginal cost charging was used that user would face very high costs. It is for this reason that typically some sort of average LRIC[2] is used so that all users who purchase the service in the relevant period pay a common average price, which is not actually the marginal price of their usage.

Typically, the LRIC costs of providing an incremental level of service over a year are divided by the (actual or forecast) demand for that service so that what is being calculated is actually the average LRIC not the marginal cost. CEPA in their work for Ofgem (CEPA 2017) describe the current GB arrangements as “forward-looking average incremental cost”.

While this is a practical approach for network charging there is a sense that on the distribution network it can lead to diluted signals. When faced with the prospect of needing to carry out reinforcement in a particular area some DNOs are now tendering for demand side response (or other solutions) that would allow them to avoid that significant investment. Conventional network charging is not able to provide the signals to deliver that demand response both because it is not location specific enough and because an average incremental cost is a muted signal when the decision facing the DNO is one of lumpy investment. The issue of contracts versus charging is discussed further below.

The other area where this comes in and which Newbery (Newbery et al 2005) picks up is connection charging and the move to shallow connection charges on distribution which helped overcome the “first comer” problem where someone connecting would trigger reinforcement costs which others would then be able to benefit from. This again reflects the “lumpy” nature of grid investment. The idea is that the connecting customer pays their site-specific costs and then a share of the common reinforcement costs. Ofgem are proposing to look again at the question of the connection boundary in their work on network access. Addressing the “first comer” problem remains an important consideration.

Residual costs and the principle of minimising distortion

The third issue from an economic perspective is how to handle the residual element of costs for which there is no cost driver but which the companies need to recover.

Again, going back to very early thinking on utility regulation shows these problems are not new. A hundred years ago the engineer William Raymond (Raymond 1918) articulated the problem in terms that are equally relevant today:

“A fundamental principle would seem to be that the basis of charge must be the cost of service… In some cases the charge based on cost may require some modification because the cost of the service to a given possible customer may be more than he will pay… a certain possible user of large quantities of electric power may be able to supply himself with other power at a cost less than the complete cost of furnishing him power from the public plant. But, as before, he may be willing to pay the full operating costs and something towards the fixed charge, thus lowering the average charge to all consumers to something less than it would be were the user of large quantities not served.”

Raymond then argues for the utility charging “what the traffic will bear” in such cases but doing so within the constraint of overall returns being fair and noting that determining what the traffic will bear required judgment and should only be done where it was of advantage to the general community.

[2] Long Run Incremental Cost

These days the standard economic argument is that the aim should be to minimise distortions to the underlying cost signals through use of Ramsey pricing (or more strictly Ramsey-Boiteux pricing)[3]. Ramsey pricing is based on allocating fixed costs in inverse proportion to the demand elasticity (ie fixed costs are loaded onto those services / customers which are least price responsive – which equates to least distortion in terms of economic efficiency).

In effect Ramsey pricing is looking to minimise effects on consumption. Whether this should apply in an energy context where demand reduction is a clear policy goal is a moot point. From a pure economic perspective this would not be an issue if the externalities associated with consumption were fully reflected in the price. In a situation where they are not then it is not clear that the Ramsey objective is appropriate. Ofgem acknowledge this implicitly where they make the point that some distortions are more harmful / beneficial to consumers than others.

Newbery (2005) talks about Ramsey pricing for fixed and common costs and notes that how sensitive an issue this is will depend on the shortfall. On this basis the natural sequence would be to look first at revisiting the cost-reflective element of charges, which is not the approach Ofgem have adopted.

One of the criticisms often levied at Ramsey pricing is on practicality. Acknowledging this Decker (Decker 2014) lists a wider set of options for recovering the residual including average cost pricing (often called "second best" pricing and used in telecomms) using a common per unit mark-up; and two part charges (commonly used in gas) comprising a capacity charge to recover fixed costs and a commodity charge to recover variable costs.

Pollitt (2016) talks about the idea that Ramsey-Boiteux pricing could be combined with social weights which would reduce the share of the residual born by low income customers. He notes that historically richer customers were more price inelastic and hence under Ramsey rules should anyway be paying a higher share. Clearly this has not proved practical but the idea is important. Pollitt does however flag that in the future with DERs such customers might become more elastic.

Newbery (Newbery et al 2005) also explicitly recognises the potential for wider considerations to play in, saying that the aim should be to "minimise distortion subject to other objectives that Ofgem may wish to take account of such as equity or extra encouragement to particular forms of generation".

This is redolent of the argument that Professor Helm makes in his Cost of Energy review (Helm 2017) in relation to the recovery of legacy renewable support costs. His core message is that the allocation of these costs is a political decision – there is no right economic answer. This same argument applies to the residual element of network costs and it is therefore surprising that BEIS have not been more actively engaged in this aspect of network charging. There is little point in the regulator seeking to level the playing field if government is intent on tilting it in a particular direction through other subsidies and charges.

While the concept of Ramsey pricing has historically been used as an excuse for loading residual costs onto domestic customers rather than commercial, the MIT paper (MIT 2016) talks about allocating these costs on the basis of property size or property tax. This serves as a proxy for wealth (with greater wealth linked to lower elasticity) and also has the benefit of being linked indirectly to consumption and being seen as equitable. They note the alternative of recovering these residual

[3] The concept was first developed by Ramsey in relation to taxation (Ramsey 1927) but then extended by Boiteux to cover monopolistic utility pricing (Boiteux 1971)

costs through taxes in the same way as other public goods such as roads are paid for. Their top recommendation is to move away from usage based charging for policy and network residual costs to an annual sum dependent on some measure of wealth or possibly some measure of capacity like fuse size.

The use of some sort of wealth proxy has not been contemplated in the GB discussions on network charging. It has practical challenges but should not be dismissed as part of more radical reform. The same arguments apply to recovery of policy costs. More fundamentally there is no reason why the residual costs of networks should not be recovered through taxation in the way that other public services are – this is a policy choice for government as Dieter Helm notes in relation to other sunk costs in the system.

From an economic point of view the principle of minimising distortions follows on naturally from the concept of cost reflectivity. However, given that it is impossible to avoid all distortions there is more debate to be had around the trade-offs in this area. In their consultation on the Targeted Charging Review (Ofgem 2017a) Ofgem acknowledged that some distortions were seen as more desirable than others and fairness concerns, in particular in relation to vulnerable customers, were an important consideration.

However one could go further and as part of the public debate on network charging there needs to be recognition that recovery of the residual could be used to help support wider public policy goals. That may be beyond the remit of the regulator but it should be on the table.

The engineering viewpoint

In so far as the focus of network charging is on future network investment there is a close link with network planning, which is driven by the need to satisfy network engineering design standards and other regulatory incentives on eg interruptions and losses. Ensuring that these engineering design standards properly reflect the changing nature of networks should therefore be a precursor to any assessment of the cost reflective element of charges.

Equally a broad appreciation of key cost drivers from an engineering point of view is important in determining what might make most sense in terms of the structure of charges. There are some elements of costs that have long been acknowledged as important in terms of network charges – such as losses - but where time constraints and practicality considerations meant they were not properly taken account of previously. With the increase in distributed generation on the network, the importance of some of these factors has increased and they need to be revisited urgently.

Engineering design standards

When looking at the incremental costs caused by additional loads on the system it is network engineering design standards that will ultimately determine the investments that companies have to make. The first step in reviewing network charges should therefore be to review these standards. In the GB this work has been started, led by industry, but its strategic significance is under-estimated.

On the distribution network the relevant engineering standard is known as Engineering Recommendation P2. The standard currently sets out in prescriptive terms the level of redundancy the networks have to build into their networks to ensure that in certain outage situations customers continue to receive a supply. The detail of P2 is included in the Distribution Code (a technical code maintained by industry with Ofgem having some oversight). The requirement to comply with the standard is also reflected in the distribution network licence.

The current version, ER P2/6, was published in July 2006. However, ER P2/6 is only a relatively minor update to its predecessor ER P2/5; which was published in 1978. It is therefore overdue for review.

In recent years industry have been carrying out a review of P2/6 (see the Distribution Code Review Panel web page and industry Phase 1 report[4]). This has drawn on work by Imperial, NERA and DNV GL looking at the implications of the widespread deployment of non-network technologies and the changing role of the customer, looking to move away from the model of security based on a set level of redundancy in the system to a more risk based approach recognising the risks in different assets and the range of solutions available. The phase 1 report recommended pursuing two potential options – one a revised deterministic set of standards, the other a lighter touch set of standards but with an obligation to conduct CBAs around future investment. In January 2018 industry consulted on a revised version of P2 which set out the security standards to be applied but was not prescriptive on how they should be met (although some of that detail remains in guidance)[5].

This work is important and has a strong read across into network charging. If the conclusion is that it is no longer practical to set deterministic standards for network design then that would seem to rule out having a common cost model. A more flexible approach to network charging would seem to be the obvious consequence. The industry work on P2/6 should be given more public profile and its significance for network charging should be properly recognised.

Horses for courses

Understanding from an engineering perspective what drives particular categories of costs ought to be key as regulators try to think about the structure of network charges.

According to Strbac (Strbac et al 2005), from an engineering perspective, the design of individual network circuits (lines and transformers) is determined by considering whichever one is relevant out of:

- maximum load and secure generation output (for demand driven design)
- minimum load and maximum generation conditions (for generation driven design).

In line with economic theory he focuses on the marginal cost impact of each user – but highlights that this will vary depending whether the particular plant is demand dominated or generation dominated, noting that seasonality will tend to have an effect with plant potentially being load dominated in winter and generation dominated in summer.

While this distinction was anticipated back in 2008 it was not built into the model for the common charging methodology which takes no account of how distributed energy resources might be used to support security, which is a critical weakness in the current charging regime.

Moreover, previously when all regions had broadly the same network challenges there was a strong case for a common charging methodology to simplify and clarify arrangements for users. With the growth of distributed energy happening differentially across the network the cost drivers may be quite different in different areas. As such – while the principles may be common – it may not be appropriate to have a single model of distribution costs.

[4] https://www.enwl.co.uk/innovation/smaller-projects/network-innovation-allowance-projects/enwl003---review-of-engineering-recommendation-p26/

[5] DCRP/18/03/PC - Revision to Engineering Recommendation P2 - Security of Supply

Losses

Both Newbery and Strbac talk about the importance of losses as an element of network costs but ultimately, they were not built into the common charging methodology model with any level of granularity.

DNOs publish Line Loss Factors which estimate the losses between the Grid Supply Point (GSP) and customers for different voltage levels and classes of customer. These factors are then used in settlement to increase the volumes as measured at the meter to provide a notional figure for the amount the supplier has to purchase at the GSP. While an attempt was made at one point to provide a financial incentive on DNOs to reduce losses there were significant problems with measurement and the current RIIO incentive is purely reputational. Revisiting how the cost of losses is recovered would be a major task but one that will be of increasing importance with the growth of DERs.

Newbery talks about the fact that losses vary in time and place (eg they are highest at the end of a circuit). Increased DG will reduce losses and marginal losses will be greater than average given that losses vary with the square of energy. He recognises that it is too complex to vary costs fully by location but maintains "DG can reduce system losses and should be credited with that." This point has been well understood for many years (eg Evans 1993).

Strbac confirms the impact of losses on network design can be quite significant and would point to the optimal utilisation of distribution circuits being quite low. He also argues that given that the magnitude of electricity losses in distribution networks is significant (and in particular in low voltage networks) there may be a case for reflecting these in cost allocation.

The P2/6 review mentioned above has referenced losses noting that increased redundancy will help reduce losses. However they do not plan to reflect losses in the updated standard, seeing this instead as something to be addressed elsewhere in the regulatory framework.

One of the benefits of local energy solutions where local demand and supply are matched is that electricity has to travel less far which reduces losses. At present local energy projects cannot be rewarded for that genuine efficiency which they deliver. With the growth of DG and a rising interest in local energy it is vital that this element of costs is properly modelled and reflected in the charging regime going forward.

Other costs

Other costs may be hard to attribute to a particular cost driver but it is clearly worth putting effort into attempting to understand certain areas of cost better to try to identify at least a proxy driver. Not doing so will lead to the costs being treated as a residual and allocated on whatever generic basis is decided for such costs.

For example, one category of other costs which Strbac discusses is circuit breakers and other switching equipment to deal with fault currents (ie any abnormal currents that might damage the system). Most DG contributes to higher fault levels – and the need for investment to address this can be one of the constraints on DG connections. Accurate allocation of these costs requires fault level analysis but finding some reasonable proxy would help in attributing costs to where they should fall.

Another driver of network investment which Strbac discusses is voltage considerations where DNOs have a requirement to keep voltage within certain parameters. DNOs will carry out voltage control activity but voltage considerations may drive design of long distribution feeders and result in higher capacity conductors being used than would otherwise be needed. However, Strbac argued that it was still right in this case to treat maximum load as the cost driver (as maximum voltage drop will occur during maximum load).

The interruptions incentive in the price control has driven additional capital and operating expenditure including automation to restore supply more quickly, under-grounding of circuits etc. While not load related this is part of the service provided to customers and could be allocated in line with the perceived value (which may or may not vary with actual usage).

These examples highlight the importance of bringing an engineering perspective to consideration of network charging structures.

More generally it is important to try to understand what is driving potentially large shifts in costs. For example, the residual element of GB transmission charges has doubled in the last 5 years and is projected to continue growing[6]. This has resulted in a large increase in the residual charge and the focus on the distortive impacts of "embedded benefits" (discussed further below). However the first step ought to be to try to understand why transmission costs have increased which may help in determining how to allocate what is currently counted as a residual.

Cost modelling

To determine future investment costs (and hence LRIC) the DNOs use a model which looks at the network cost of adding a 500MW demand at each voltage level. These costs are then allocated across voltage levels and customer groups to give maximum demand and /or unit level charges. Simplifying assumptions are made – eg rural and urban charges are the same, despite cost differences, for social / political reasons. The current model does not take account of multi-directional flows.

The thinking behind such a model is set out in Strbac who looks at pricing based on the concept of a reference network where a model is constructed reflecting the current loads and topology but capacity is optimised to minimise investment and operational costs associated with constraints. For a distribution network, he argues that any model needs to be run under different loading conditions, reflecting the presence of DG on the network.

He talks about the ideal model being one where you are able to optimise and cost expansion plans to find the least cost way of meeting a specified increment of demand or generation at any node (for different sizes of increment). Predictions of future demand and generation would be needed and the model would need to be calibrated to the particular DNO network given they face different constraints.

It is clear that the current model falls a long way short of this ideal and that the challenges are only increasing as the discussion on P2/6 above highlights.

At the minute the task of updating the common charging methodology model sits with industry. Clearly they have the detailed understanding of their networks but as should be clear from the discussion above there are some important design factors – whether to move from a common

6 https://www.ofgem.gov.uk/system/files/docs/2017/06/impact_assessment_and_decision_on_industry_cmp264265.pdf
(See figure 1)

model, how far if at all to unpick some of the socialisation of costs, the level of geographical granularity to include etc - which need a strong policy steer from the regulator.

Potential impacts of a move to DER

From the discussion above it is clear that a move to more distributed energy resources on the system will lead to a very different structure for cost reflective charges. These need to be worked through urgently, ideally ahead of any decision on the residual charging arrangements given proportionality and practical considerations.

One question underpinning some of these points is how flexibility is best encouraged onto the system. Implicit in the network charging debate is an assumption that this will be done by having appropriate price signals – potentially set on a dynamic basis to reflect short term network constraints. However networks may not be confident that such signals will be sufficient to drive a response. If networks consider that they need a "firm" response in order to protect the security of the system they may well contract for flexible resources that can be called at times of local network stress as discussed below. This means in effect that there will in future be much more significant short run marginal costs associated with peak supply at certain times. The network charges that are levied should reflect those short run costs but that may in turn trigger further changes in demand.

Historically, with limited opportunity for customers to respond to price signals, the networks did not have to model likely response as part of setting tariffs. In future, modelling the demand side elasticity will be a factor that DNOs will have to take into account in setting tariffs if they are not to suppress demand unduly.

Similarly the limited opportunity to respond to price signals historically meant that there was little benefit in setting different charges for geographical areas below a DNO region. Although costs varied significantly by area the practical and equity considerations outweighed any efficiency benefit from further geographical disaggregation. With the increase in DERs the balance shifts.

In summary there are a number of ways in which increased DER will impact on network charging. While the aim remains to maximise economic use of the network the practical implications will be different:

- the need for short term signals is of increasing importance
- cost drivers may vary geographically depending whether an area is demand or generation constrained
- new drivers of cost such as reactive power will be of increasing importance
- DNOs will need to think about the impact of customer's response to their tariffs
- there is a stronger case for more granular tariffs.

Trade-offs

As noted above while there has been a strong consensus about the principles that should inform network charging there has been little discussion of how the trade-offs should be made, although some authors do explore particular aspects.

The overview of US regulatory approaches (NARUC 2016) says "Rate design .. is often said to be more art than science", noting that while there is often agreement on the goals and principles, parties will value and weight these goals and principles differently.

This reflects Bonbright's line (Bonbright 1957) that "this attempt to make rates perform multiple and partly conflicting roles calls for wise compromise, and the key to wise compromise can seldom be found in any simple formula or in any simple measure of economic optima".

With a move to distributed energy some of these tensions are only set to increase and regulators are typically looking at quite fundamental reviews of how they approach network charging. As such the question of how to make the tradeoffs between different principles becomes important. One aim of this paper is to try to provide a framework for thinking about these tradeoffs.

As with many such problems, if it is possible to identify a solution that "squares the circle" and that does not require tradeoffs to be made then clearly that would be ideal. However none of the papers seems to have identified such a clearcut winner.

For example, the MIT paper notes the idea of residual costs being allocated linked to property tax. This would have the advantage of minimising distortions in relation to energy use while also being socially progressive. However there are likely to be practical implementation challenges (given energy companies do not have access to this information).

The same paper talks about finding an "efficient level of granularity" given implementation costs and complexity. It argues that you can get a significant proportion of the benefit if you remove residual and policy costs from volumetric charges and then move to an hourly or sub hourly basis for wholesale costs.

On the basis that tradeoffs will have to be made there needs to be a framework for determining in what circumstances a particular principle should be given a higher weight. It is likely to be impractical (or spurious) to do a formal cost benefit assessment of the different options but identifying the factors and the evidence needed is important. Absent such a framework, regulators are likely to base these tradeoffs on their personal preferences and prejudices eg placing a heavy weight on cost reflectivity above other considerations.

In GB the fact that proposals for changes to the charging methodology are under industry self-governance tends to mean more weight is given to the economic and technical arguments. The industry codes include specific objectives that industry have to judge any proposal against. These are cost reflectivity, promotion of competition and efficiency. While Ofgem has a final veto on any changes – and can bring in at that stage wider considerations around the consumer interest – there is an inherent bias in the types of proposals that get brought forward.

These industry driven changes can be quite significant. For example, a recent modification(DCP228) changed the way the residual charge was recovered within the distribution charge for half hourly settled customers[7]. These customers face a time of use charge which varies by time band (known as "red-amber-green" periods). The modification involved changing the recovery of the residual from a

[7] https://www.ofgem.gov.uk/system/files/docs/2016/09/dcp228_decison_letter.pdf

% uplift on the cost reflective charge to a flat rate uplift across the day on the basis that this was less distorting of the underlying cost signals. The impact was to reduce the "peakiness" of the charges in all DNO areas with the most extreme being Western Power Distribution's area where the peak ("red") rate fell from 18p/kwh to 6p/kwh. This clearly had damaging implications for demand side response. However there was little or no debate on the strategic implications of this change.

Case studies to illustrate tradeoffs

While there hasn't been an attempt previously to articulate the factors that should be considered in making tradeoffs between the principles, it is possible to draw some conclusions from looking at particular regulatory decisions and the issues considered as part of those decisions. These are considered as a series of case studies below.

Spare capacity – case study looking at the gas sector

Chris Decker has explored some of the issues around networks in decline (looking in particular at gas transportation networks) and the implications that has for the tradeoffs that have to be made between different objectives (Decker 2014, Decker 2016). In particular where there are no capacity constraints and demand is declining then there will be less emphasis on sending cost signals to influence short term usage and the focus becomes on cost recovery in a way that is perceived as "fair". In gas this has tended to lead to a shift towards more of the costs being recovered through capacity charges and much lower commodity (or unit based) charges.

As noted above, the issue of whether prices should be based on long run or short run marginal costs is a key consideration. If there are capacity constraints then LRMC will send the right signals about efficient usage. Whereas if there is excess capacity LRMC will lead to under-utilisation and inefficiency.

Recovery of historic costs may lead to some loss of static efficiency as prices will be higher than marginal costs but may lead to dynamic efficiency gains by enhancing the credibility of the regulatory regime and making others more confident to invest.

Decker identifies a number of tradeoffs including:

- Between static efficiency and cost recovery;
- Between allowing costs to reflect prevailing conditions and price volatility (with implications for investment by users);
- Around fairness and non-discrimination considerations and allocative efficiency when looking at how to recover fixed costs.

Exploring the same issues, the 2002 Brattle report on the European gas market argued that

"with growth or congestion, capacity is scarce and tariffs face the primary challenge of ensuring efficient allocation. The relevant cost concept is prospective, related to scarcity value and the marginal cost of construction"

"with no growth or congestion the primary role of the price mechanism is to allocate the fixed costs of previous investments among system users. The relevant cost concept is retrospective, relating to the allocation of costs already incurred (average cost). It emphasises cost allocation methodologies designed to correspond to intuitive notions of fairness."

Before setting off down the path of a major network charging review, the regulator needs to be clear what the projections are for capacity and hence what the primary purpose of a charging regime should be.

Government Policy – case study on Project TransmiT

The growth of transmission connected renewables in GB raised a number of questions around the transmission charging structure. Transmission charges include an element of locational pricing and most large-scale wind generation is based in Scotland far from centres of demand. On a cost reflectivity basis such generation should face higher charges but this was seen as running counter to government policy to increase renewable generation.

To help in thinking about how to balance economic efficiency against the facilitation of carbon reduction Ofgem commissioned four academic papers and a peer review of them (Eakins 2011).

In his paper Professor David Newbery (Newbery D.M. 2011) makes the point that there is a choice if you want to support renewables. Either you can take account of that goal in the way that you set network charges or you can pay higher subsidies through some other route in order to offset the impact of the higher network charges. Whether such subsidies are more or less distorting than addressing the policy through network charges is a difficult question but the key one.

This is particularly relevant in the case of renewable generation which in practice may have limited options over where it is to locate given heavy planning constraints.

In his comments Professor Keith Bell (Bell 2011) positions the problem differently as being to minimise network costs subject to meeting the renewable targets. If changes to network charging arrangements would make that target more costly to achieve then that should be taken into account in considering the case for change.

This clearly makes sense where, as was the case with renewables, the targets are legally binding (through EU legislation). In general the need to meet legal obligations can be taken as a given. Clearly where the issue is one of broader government policy then a judgment would need to be formed on whether or not government would take steps to mitigate the impact of changes to charging.

In general however this highlights the importance of putting network charging in the context of wider government policy. There is little point in the regulator working to establish a level playing field if government policy requires that it be tilted – apart perhaps from greater transparency about the reasons for particular decisions.

Ultimately Ofgem's decision found something of a middle course with greater account being taken of the different cost implications of intermittent generation justifying a different level of charges.

User choices – Case study on the risks of grid defection

As discussed briefly above, the risks of grid defection were first given a public airing in 2013 (eg GTM 2013) in various articles which highlighted the risk of a "death spiral" on the networks where, as more people came off the network, the fixed costs would have to be borne by an ever smaller number of customers whose bills would rise as a result, driving yet more of them to defect. These risks have become more pronounced as the costs of storage have fallen in recent years opening up the market for combined storage and solar solutions which allow customers in some parts of the world to defect from the grid. Currently this is only a serious issue at the domestic level in places like Hawaii and Australia and, although the range of countries likely to be impacted will increase as costs fall, it seems unlikely to be a major issue for GB for many years (if ever).

What is a much more immediate issue is where customers do not completely defect from the grid but as a result of having their own generation are able to limit the amount of electricity they import from the grid and hence the share that they pay of network and policy costs. Given that domestic customers who have behind the meter generation will typically be the better off in society this creates distributional issues (in the same way as happens with full grid defection).

This is a particular problem in the US where the model of support for domestic scale solar has tended to be the use of net metering. Under net metering any exports are deducted from the level of imports to provide a "net" consumption figure which is used as the basis of charging. This contrasts with the GB model where customers are paid an export rate for any energy exported but are still charged – and would pay the network charges for – their gross imports.

However even in the GB model there is still a significant distortion. A recent paper for NEA (NEA 2017) identified that customers with a Feed-In Tariff would typically be paying £60 less in network and policy costs than a customer without. However this is less extreme than a full net metering model and indeed California has recently moved to a revised net metering model where network charges are levied on the gross imports.

Pollitt (2014) presents a case study from Queensland Australia which demonstrates significant transfers between richer and poorer customers with customers with PV paying over $300 less in network charges.

Thus there is a concern about partial grid defection – or what RMI calls "load defection" (RMI 2015) - and the implications for network charges, in almost all jurisdictions. However the scale and urgency of the issue will depend in part on the design of the renewable support arrangements as well as on the local climate. It is a bigger issue in the US with net metering than here but it is still an issue for GB consumers today.

The other related risk is around private wire networks – where industrial and commercial developments or local authorities are increasingly looking at the potential for private wire solutions to link their various sites with generation and storage. While this is encouraged in GB by the structure of FIT payments (and the value of export versus import) an additional consideration will be the desire to avoid network charges and the associated policy costs – which is not efficient given that these costs will still need to be recovered from other customers.

However microgrids in the US are becoming more prevalent as part of a drive to improve the resilience of the system following disasters such as Hurricane Sandy. There clearly are benefits for both the customer and for the network operator in having these systems with the capability to operate on an islanded basis. The important thing is that the decision to develop them should be based on a proper reflection of the costs and benefits.

The question is then what this means for network charging. The key point is that what one wants to avoid is customers making decisions on the basis of incorrect price signals, in a situation where it is clear that they do have real and practical choices that they can make.

Going back to economic principles, if the network charges were genuinely cost reflective then customers would only go off grid if the value to them of being on grid was less than the cost. In some places – perhaps in rural areas at the end of a long distribution cable it might actually make sense for the customer to go off grid if the costs of maintaining that link were significant (and certainly if a new link were needed).

However if the customer is making the trade-off based on looking at a total cost including a high "residual" charge, including policy costs, they may well decide to go off grid when it is not economic from a system point of view and other customers will simply end up paying more.

The same issue arises in relation to load defection where a customer is making an economic assessment of whether to install solar PV and is doing so on the basis of the charges they will avoid. If these are not actually avoidable costs at a system level then the customer will not reach a proper “economic” decision (though it may be one that the state is still happy for them to make given the desire to increase renewable penetration).

One challenge in terms of network charge design is that the least distortive approach for dealing with the residual costs differs depending whether the risk is grid or load defection. Where the risk is load defection, as in GB and many other jurisdictions, the pressure is to move to some sort of fixed or commodity charge for recovering the residual so that the customer continues to contribute to recovery of those fixed and common costs even if they are generating their own energy. This can then be positioned as the price they have to pay to access the network (either to import or to export electricity). If full grid defection isn’t an option then this will ensure that all customers continue to contribute what could be seen as a fair share of the costs and any distributional concerns are overcome.

However if full grid defection is an option then seeking to charge the full residual through a capacity charge could be what would tip the customer over into defecting. It may be that a lesser charge (perhaps linked to actual usage), reflecting in effect what the market will bear, is the best way to avoid uneconomic grid defection. It is in this scenario that the tensions between the principles become harder to manage. A principle of minimising distortion might say that you should not charge customers with solar/storage any element of the residual (given they have a choice that others don’t have). However this cuts across the principle of fairness which would want to see such customers making at least some contribution, particularly where they are typically wealthier.

For private wire solutions it may mean there is a need for more flexibility in how DNOs charge which is what would happen in a competitive market. If faced with a competitive tender then competition economics would suggest networks should be allowed to price down to their marginal cost. This is redolent of the comments by Raymond referenced above when the issues were about encouraging customers to connect to the grid in the first place and he argued for “charging what the traffic will bear”. Regulators may feel they face a tension between encouraging competition and having stranded costs that need to be picked up by the remaining customer base. The answer is for regulators to ensure that competition is only encouraged where it is economic from a system perspective, which it is unlikely to be where there are already existing assets which would be duplicated.

Further thought also needs to be given to the benefits to the distribution network of more local matching of supply and demand as can be facilitated on local energy projects. Private wire solutions are only cost effective where generation and demand are in close proximity. The current arrangements for network charging take no account of this proximity and hence can send a distorted signal to encourage private wire in such cases.

What these examples highlight is that the principles on their own are of limited value. In order to work out how to trade-off between them the regulator needs to have the practical evidence about the nature and the scale of the risk that distortions might create and ultimately to model how customers might respond in practice to different scenarios. This takes regulators much further into the territory of consumer behaviour modelling than they would traditionally go in setting network charges (whereas “cost reflectivity” is much easier to handle). However, in the context of grid defection and the need to balance fairness (or distributional impacts) with minimising distortions such evidence is needed.

User choices – case study on connections

One of the factors prompting Ofgem to look at the issues around allocation of residual costs was a concern that distributed generation was benefitting from the arrangements for network charging at the transmission level (and what are known as “embedded benefits”) (Ofgem 2017b). The residual element of transmission charges has historically been set based on the supplier’s net demand during the Triad period (the three peak demand half hours in the year). The fact that dispatchable distributed generation could be used by suppliers to reduce their net demand at potential Triad times created a major source of value for such generators known as “embedded benefits”.

Ofgem was concerned that this could lead to a distortion in the choice for companies in connecting at the distribution or transmission level (and that distribution connected generation had an unfair advantage in the capacity market because of its lower network costs once embedded benefits were accounted for).

One of the messages from a number of respondents (the author included) to Ofgem’s consultation on the Targeted Charging Review (Ofgem 2017a) stressed the need to look more widely at the basis for charging and to include within that the connection charging arrangements.

Ofgem has now published a further document on Network Access (Ofgem 2017c) which includes a fuller discussion of the access arrangements and notes that transmission connected generation is guaranteed a connection and is paid when it is constrained off whereas at distribution level generation may have to pay for reinforcement or accept a “non-firm” connection, with no payment when constrained off. These other factors are likely to be material to a generator in determining where to connect (along with non-energy issues such as planning).

There is no evidence that Ofgem was thinking about these other factors in deciding how far and how fast it needed to move to remove the distortion in costs that it had identified on “embedded benefits” (the removal of which had major distributional impacts and risked investor confidence in the stability of the regime).

The system of connection charging and constraint payments needs to be viewed as an equally important part of the network charging arrangements. Distribution connection charges are seen as a separate pot of money outside the price control – but higher level reinforcement costs are a part of the revenue that network charges need to recover.

Where constraint payments have to be made (as on transmission) this provides a clear economic signal to the system operator about the need for reinforcement. In contrast on the distribution network where DNOs currently face no marginal cost for constraining off generation (and no constraints on when they can do so) there is no equivalent signal – this remans a “free” option at the point of use which then distorts the DNO’s choices between different alternative approaches to managing a particular constraint.

Visibility of charging signals – case study on retail markets

In GB, for most consumers, the supplier-hub model means that the supplier sets charges and recovers the revenues from end consumers. As such whether or not any particular structure and level of charges is passed on to consumers is up to suppliers.

Currently, there is no regular data collected on the extent to which non-domestic consumers are seeing the time of use structure of network charges reflected in their tariffs. Pollitt (2016) cites the

CLNR study[8] which showed less than 5% of half hourly settled customers had the red-amber-green network charging structure reflected in their bill. Hence even where there is a time of use element in network charges as there is for half-hourly settled customers, there is typically no price signal being passed on to end users[9].

The argument from suppliers is that customers value the simplicity of a simpler flat rate tariff (or at best a day/night rate) and do not want to be exposed to the volatility of wholesale prices. While this may well be true – and remains an important consideration – one might expect in a competitive retail market that there would be players able to identify consumers whose pattern of usage means they would be better off on a more cost reflective tariff and either targeting such customers based on their lower cost to serve or finding more compelling ways to present time of use tariffs to such customers. In addition increased automation should allow customers to programme equipment to respond to price signals, avoiding the need for them to engage directly. Over time therefore the assumption must be that consumers will have more sophisticated tariffs in which network charges do flow through.

However in the meantime there seems little point in moving to even more sophisticated network charging arrangements given that they are unlikely to be passed on to end consumers. This is a clear example of what MIT talk about in terms of having a clear transition pathway – with the direction being to more granular charging – but without moving too far ahead of where the customer is at at any point in time.

Again regulators may be wary of trying to second guess the pace of change in tariffing and want to pursue a more cost reflective model straight away to give the signals to suppliers to develop new business models. However when faced with other pressures (such as distributional impacts or high implementation costs) this sort of reasoning argues for a slower pace of change. Gathering evidence on the extent to which charges are currently passed on would provide important context to help in making such tradeoffs.

Visibility of charging signals – case study on "Triad"

One of the features of the GB transmission charging regime is that charges for larger customers are based entirely around their use in the "Triad" period, defined as the three highest annual demand events (half hours) over the winter. This applies to both the cost reflective element and the allocation of residual costs. One notable feature of this arrangement is that it is ex post. This has the benefit of meaning that charges are loaded onto those customers who did in practice contribute to peak demand (and hence it focuses on what should be the real driver of costs). By focussing costs in this way there is a very strong incentive on consumers to try to minimise demand in this period with savings cited of £30-50k per MW (National Grid 2017). However, in terms of a practical price signal it is problematic as users do not know in advance when the Triad periods will be.

As a result, a whole industry has developed off the back of these arrangements with suppliers and aggregators offering "Triad warning" services where they attempt to anticipate when a Triad period might arise so that their end customer is able to turn down usage in the hope of avoiding these charges. Inevitably this results in a "musical chairs" type process whereby the times that are flagged as potential Triads see everyone shifting load so that that half hour is no longer the peak and the

[8] Customer Led Network Revolution report L247
[9] https://www.ofgem.gov.uk/sites/default/files/docs/2014/10/bsc_mod_p272_-_supporting_analysis_1.pdf

game starts again. Moreover, as noted above, because the Triad is based on net demand, turning up embedded generation is also an effective approach (and arguably has lead to embedded generation being over rewarded). In all likelihood the Triad days will simply be those that suppliers or aggregators were unable to anticipate. On average a company will respond to a number of Triad calls in order to try to avoid the 3 peak half hours and the National Grid Power Responsive report (National Grid 2018) has flagged that in more recent years it has become harder to predict to the extent that in 2016/7 there were 48 observed Triad avoidance days compared to 15 ten years ago. And given that ultimately the full costs still have to be recovered this simply loads costs onto those who cannot predict Triads and respond.

The consensus seems to be that Triad has been very effective as a demand response stimulus because of the high level of charges that can be avoided. It has also been argued that Triad has had positive benefits in terms of improving UK capacity margins (Regen SW2017).

However the purpose of network charging is not meant to be about manging wholesale capacity margins (although this may well have been a legitimate wider consideration at the time when Triads were introduced and there was no capacity market).

In the absence of a clear signal of what charges will be in a particular period it must be assumed that users are making economically inefficient decisions on occasions – cutting back on usage when there was no economic need for them to do so. In particular by allocating residual costs on this basis a stronger signal is being sent than is justified by any long term network cost consideration.

For smaller users the challenges of anticipating and dealing with a Triad type charging approach would be unmanageable. For such users – who will not have the resources to devote to tracking market movements – a simple price structure where prices are signalled clearly in advance (even if only 24 hours before as in some dynamic pricing models) is imperative.

However, even for larger users – and noting that Triad participation has grown significantly in recent years – a simpler model (or at least a move away from Triad for the residual costs) might be more likely to yield outcomes that are economically efficient in terms of usage of the network.

In general the actual peak capacity will only be known after the event – but charging on an ex post basis provides less clear price signals to customers. When only a small number of customers were engaging (and there were fewer options for how flexibility might be provided) it was easier to anticipate when the Triad might be and hence a weighting towards cost reflectivity over practicality was reasonable. Looking to the future that balance needs to shift. Models such as critical peak pricing at the retail level (or peak time rebates) would allow the customer to be given a firm price signal perhaps 24 hours in advance (or potentially shorter) where there are constraints on the system.

Vulnerable customers -case study on regional variations

Within GB, distribution charges are averaged across a network company area but there can be significant variations between different regions. This has led to pressure from disadvantaged regions (and in particular Scotland) arguing that such differences are "unfair". Responding to this public pressure Ofgem recently produced a report looking at regional variations to inform the debate (Ofgem 2015).

This showed that distribution network charges did vary significantly by region – although this was in part offset by transmission charges that tended to work in the opposite direction.

In deciding whether this justified any action Ofgem looked at the extent to which vulnerable customers were disproportionately represented in areas that had higher charges and concluded they were not. On this basis, from the regulator's perspective, the priority was to retain the current regional charging structure which reflected the underlying network costs. Clearly there has always been averaging among customers within an area, notwithstanding differences in costs between rural and urban areas for example. However, at a time when the pressures are to drive to more granular charging Ofgem took the view that "fairness" considerations did not point to more averaging across regions.

While it is not clear what steps Ofgem would have taken had it found that vulnerable customers were particularly disadvantaged (or if it would simply have passed that decision on to government), it was clearly an important consideration. If pressed by government to take action one route through would have been for Ofgem to look at alternative ways of allocating the residual costs to deliver on these wider policy goals around affordability and the regional dimension (which would have involved a process for re-allocating costs between regions as set out in the report).

Despite the difficulties in defining "fairness" it remains an important consideration when significant changes are being considered to the nature of charging. In its Targeted Charging Review consultation (Ofgem 2017a) Ofgem have articulated a general principle that "all users who are connected to the licensed networks should make some contribution to common costs". Beyond that it is clear that Ofgem's interests in what constitutes fairness is focussed primarily around the impact on vulnerable customers.

A framework for considering tradeoffs

Pulling together themes from these case studies the factors that it is proposed should be considered in any framework for considering tradeoffs are as follows:

- Spare capacity: What level of spare capacity is there on the network? With a network in decline or where capacity is not constrained then cost reflectivity to drive allocative efficiency becomes less important.
- Wider government policy: Is the policy enshrined in legislation? What mitigating actions would government take? How would the cost of delivering that policy be impacted by choices around the allocation of the residual or the structure of charges?
- User choices: What are the real choices facing users and how likely are they to pursue them? What other factors might shape those choices beyond network charges? Understanding the risks that arise if charges are not cost reflective shapes how much weight should be placed on that principle and what cost comparisons are most relevant.
- Visibility of charges: Will customers be able to see and respond to the price signals in practice? If the underlying cost signals won't be passed on or will be swamped by other elements of charges then a strong emphasis on cost reflectivity is less justified.
- Vulnerable customers: Who are the winners and losers and in particular how would any proposal impact on vulnerable customers? While distributional issues will arise in changes to network charges, if these impacts will hit vulnerable customers in particular then it will be hard to gain political and public support for any change.

It is acknowledged that these factors are somewhat subjective and that users will develop new business models to exploit distortions in pricing so one cannot simply rely on historical observations of behaviour. This argument tends to lead regulators to focus on cost reflectivity to avoid having to second guess the market. However faced with difficult trade-offs, including significant distributional issues which can be politically difficult, having a strong evidence base (and not simply a "theoretical argument") is important.

Tradeoffs can also change over time. For example, as smart meters are rolled out that will create more opportunities. There is also a need to think ahead to a time when there will be more electric vehicles (EVs) and heat pumps With EVs, sending a TOU signal to encourage charging off peak will become more important. It will also be important that reinforcement costs to cope with these higher loads are not automatically socialised into a higher standing charge given the demographics of those likely to take up EVs in the short term. In contrast the case for socialising the reinforcement to support the uptake of heat pumps is a question that should be explicitly considered as part of the strategy for heat de-carbonisation. Considering these future trends is important and care is needed to avoid the danger of overly focussing on the immediate problem of behind the meter generation.

Potential Tariff Structures

Again there is a strong consensus among the commentators referred to earlier on the range of different structures that could be adopted for network charging although the way that they are categorised does vary and there are some detailed differences.

The table below attempts to list all the variants that are mentioned in the literature grouped under the broad headings that Ofgem uses in its Targeted Charging Review consultation. This highlights that beyond a crude discussion of capacity versus usage based charging there are levels of detail that need to be considered in settling on any particular tariff structure (or combination thereof).

Table 1: potential tariff structures

Category	Variants	Comments
Flat KWh usage based	Can cover import and export separately (or net off)	Most common today Tends to benefit low use customers Could be based on "gross usage" if information available on generation behind the meter
Time of Use KWh usage based	Static TOU – can be on broad periods or more granular Dynamic TOU – including critical peak pricing or peak time rebate Seasonal pricing Can cover import and export separately (or net off)	Needs smart meter
Rising (or decreasing) block tariffs		More relevant to end use pricing – rising block encourages efficiency (and benefits low users)
Standing charge (per meter)	Can be simple per customer or linked to eg property tax band / size	Some historical resistance in GB – impacts low users who are more likely to be vulnerable
Capacity charge	Can be based on actual or booked capacity (cf fuse size) Can be based on system("coincident") peak or user's own peak Granularity of what counts as peak can vary (eg Triad or "red" period) Can be based on anticipated system peak or ex post Can cover import and export separately or net	If using booked capacity need arrangements if go over (lose supply or face penal prices)
Charges for other network functionality	Ancillary services Charge for reactive power	

Quality of service	Payment for being interruptible / interrupted Charges linked to level of resilience provided	

Historically peak usage tended to be well correlated with overall usage and hence, in the absence of the ability to measure peak usage at smaller sites, usage levels were seen as a reasonable proxy and basis for charging. With the ability for users to be self-generators this correlation between peak and total usage starts to break down and it becomes more important to ensure that the basis for charging does properly reflect the underlying drivers.

Most commentators have focussed on the usage / standing charge / capacity framework for thinking about the structure of charges, based on concerns about potential grid defection.

However, coming from a different angle, the Rocky Mountain Institute (RMI 2014) talks about the what, when and where of charging:

- Attributes – energy, capacity, ancillary services;
- Temporal – time differentiated market signals
- Locational – unique, potentially site specific

There has also been some interest from distributed energy providers in how network charging could be used to encourage “local energy”. Currently there is no incentive for customers to try to match local supply and demand, although balancing the system at a local level will reduce the distance that energy has to travel and hence reduce losses. As noted earlier the network charging arrangements currently provide an incentive for people to look at putting generation behind the meter or going as far as building private wire networks. The focus to date has been on reviewing on how to mitigate that distortion – but with limited attention paid to solutions such as “virtual private wires” (as an alternative to building duplicate networks) or options for providing a signal to encourage such balancing. Distance is clearly a cost driver – private wires solutions are only viable where sites are n close proximity – but is not reflected in the charging structure.

Pollitt (2016) talks about introducing a charge for maximum kW export for domestic and small business PV prosumers as a way of ensuring these, typically richer, customers pay a fair share of the residual. This has clear merit given that dealing with reverse power flows caused by domestic scale PV is a driver of costs for the distribution network and customers get a value from being able to export onto the grid which they should then pay for.

Another, different, way into some of the issues arising is to think about how best to segment the customer base for charging purposes. Network charging is not intended to reflect the costs imposed by individual customers (except perhaps for the very largest customers where in GB the EDCM methodology does look at site specific charges). Instead what most charging methodologies do is first allocate costs between different classes of customer (residential, commercial etc) looking at the real cost drivers and then allocate costs across customers within that class to get to unit costs. In GB the cost modelling allocates cost between meter classes (so distinguishing single rate from multi rate low users, for example). However there is not a separate class for customers with their own generation which may be appropriate going forward.

Equally within GB we know for Grid Supply Point (GSP) areas whether they are importing or exporting GSPs. As discussed above this could be expected to have implications for distribution network requirements and hence arguably the structure of charges in that area. Indeed there is a wider question about how to aggregate the customer base spatially – the GSP represents a sizeable region and some of the interests in local energy are aimed at addressing issues on individual feeders.

The Regulatory Assistance Project (RAP 2018) are very resistant to fixed charges which they see as discouraging energy efficiency and self-provision. They make the point that in most other commercial sectors (petrol, hotels etc) fixed costs are simply recovered through usage charges.

While energy efficiency remains important (and in particular while externalities are not fully reflected through the cost of carbon) the changing energy mix may point to a greater emphasis being placed on reducing peak consumption – which is typically most carbon intensive – than on reducing consumption overall. As such capacity charges do not necessarily run counter to energy efficiency goals.

While further analysis is needed to examine the impacts of the different options, some provisional conclusions would be that:

- In general a greater emphasis should be put on capacity based charges rather than usage given the risks around load defection – although in jurisdictions where full grid defection is a serious risk – including in the context of private wire solutions in GB - this could actually drive customers off grid.
- The use of capacity charges based on property value as a proxy should be given serious consideration as a way of dealing with the residual while avoiding adverse distributional impacts.
- If using energy import capacity as the basis for charging then, from an ease of understanding and predictability perspective, the charges should ideally be linked to the customer's own peak capacity (rather than capacity at a system peak which can only be known ex post) – although thought needs to be given as to whether this could drive some customers to shift load from their own peak to the system peak which would exacerbate the problem and whether it risks discouraging usage that is economically efficient.
- A charge should also be levied for peak export capacity, including for domestic customers.
- The level of granularity of charges – including charges for other aspects of the energy service – can be handled by large users but the overall regime should be kept relatively simple for domestic customers.
- Locational signals are provided most effectively through connection charges ie at the time the customer is choosing where to locate. That said cost reflective charges are likely to vary significantly between geographical areas.
- At a practical level to keep the tariff structures simple, the recovery of residual costs should be done through charges that mirror what is considered appropriate for cost reflective charges (eg it would add unnecessary complexity to have both a system peak capacity charge and an individual peak capacity charge). This points to looking first at how the growth of DER impacts on the cost drivers for the network whereas most of the debate seems to be focussing initially on the residual.
- Reflecting the increased importance of short run costs and actions by system users it is important to reflect on both long run and short run cost signals. Having some form of fixed charge to recover essentially the long run costs and usage charges to reflect short run costs would make sense.

Other considerations

Charges versus contracts

Most of the focus in the debate around charging is on the charges for use of the network. In GB the provision of ancillary balancing services is delivered in general through bilateral contracts, often based on auctions as a good way of determining an efficient price. For example, National Grid recently tendered for Enhanced Frequency Response for resources that could respond in under 1 second to help it keep the system in balance. The focus to date has been on the wholesale balancing market but the same model could apply to network constraints.

Looking ahead we can expect DSOs to be tendering for such services at a local level. Indeed some DNOs are already using auctions to buy in flexibility services to help them manage particular constraints (eg SSE Constraint Managed Zones[10] and the recent UKPN tender[11]). The initial experience of DNOs is that this is more effective as a way of securing the flexibility they need than relying on price signals which as discussed above are both hard to tailor to the granular location where the DNO faces a constraint and risk being blunted as a signal both by the use of average marginal costs and by how suppliers choose to pass the charges on. A contractual arrangement can also provide more reliable response than relying on customers' response to price signals.

Cross vector considerations

Historically gas and electricity have not been seen as substitutes for end users – or certainly not once the customer has a heating system in place. Going forward with the need to decarbonise heat there will need to be more emphasis on efficient choices being made between alternative heating solutions. While there are other significant distortions in the price at the minute (with no carbon price on gas and policy costs largely loaded onto electricity), it is important that these are not exacerbated by how network charging is carried out. With an existing gas network (at least once the iron mains replacement programme is complete) the marginal network costs are low for gas compared to other solutions which either require district heating networks to be built or significant reinforcement if we rely more on electrification.

While it is expected that any policy analysis by government would focus on the marginal costs, individual customer choices (in so far as cost is a factor) will reflect average costs including the need to recover historic costs.

There are also distributional concerns if those who are better off (and able to afford the up front costs of alternative solutions) move away from the gas grid and the less well off customers are left having to fund the residual costs. This is the same challenge as the solar/battery death spiral that is much talked about in electricity but which also needs to be addressed as part of thinking on the future of heat. Understanding the risks and the tradeoffs customers might make now involves looking across gas and electricity to ensure that distortions here are minimised.

The consumer voice in the debate

Finally there is an important question about how to bring a consumer perspective into what is a highly technical debate. In its RIIO approach to price controls, Ofgem requires network companies to engage with consumers in developing their business plans. The fact that this process focuses on the overall revenue allowances with the structure of charges being debated separately in industry

[10] http://news.ssen.co.uk/news/all-articles/2016/12/ssen-opens-constraint-managed-zone/
[11] URL: http://www.ukpowernetworks.co.uk/internet/en/have-your-say/listening-to-our-connections-customers/flexibility-services.HTML

working groups, would seem to be missing a trick. In the US, rate cases cover the structure of charges, not just the level, which means they get proper public scrutiny. At present this is missing from the GB governance arrangements, other than through the inevitably limited involvement of Citizen's Advice. A Future Charging Forum has been established but this is inevitably dominated by industry.

Ofgem include within their principles the issue of "fairness" and the impact on vulnerable customers in particular. However, it is unclear what specific steps they are taking to ensure that perspective is brought into the debate.

If radical changes are to be made to charges it will require significant effort to build support for the proposals. Historically there has been significant opposition to the use of standing charges from consumer groups which flared up when, as part of its Retail Market Reforms, Ofgem outlawed two - part tariffs. These tariffs were confusing to customers and were in effect a standing charge in disguise. However, the re-introduction of standing charges led to complaints from low (and zero) users, in particular on gas, and pressure for suppliers to waive those charges for certain groups of vulnerable customers[12]. Network charging is about how an essentially fixed cake is divided up -there will be winners and losers. For the outcome of any fundamental charging review to have legitimacy, the consumer voice needs to be heard as part of the deliberation.

[12] https://www.ofgem.gov.uk/sites/default/files/docs/2015/05/open_letter_-_treatment_of_low_and_zero_consumers_of_gas_0.pdf

Conclusions and recommendations

The question of how to deal with network charges in a world with much higher levels of distributed energy resources is one that is being considered in jurisdictions across the globe. The right answer will depend on a range of factors including the wider industry structure, the sorts of DER resources that are most prevalent (dependent on geographical and economic factors) and the broader policy goals that need to be addressed.

In the GB context, where Ofgem has now initiated a major review of network charging, there are some important points coming out of the analysis above which Ofgem and government should heed:

- The level and drivers of the cost reflective elements of charges are determined by the engineering standards that the network companies are required to comply with. Because these are technical standards it tends to be left to industry to lead on this work with no strategic or wider debate. Given the fundamental importance of these standards in the changing energy landscape Ofgem should ensure that the work to update for example engineering standard P2/6 is given a proper airing.
- The cost reflective element of network charges must take account of whether the network is generation or demand constrained in particular areas. This will almost certainly mean a different structure of charges in different areas and may require in particular the inclusion of seasonal differences (ie a winter and summer rate) where the balance of supply and demand is very different.
- There is a need to be clear about the appropriate timeframe for looking at marginal costs. If short run costs were reflected in volumetric charges and long run costs linked to capacity charges (in effect a charge for the right to use the network) then this would reflect the timescale for decisions by both the networks and the users, and help resolve potential conflicts.
- That said, for sending price signals to deal with highly localised – or time specific – constraints, contracts (for ancillary services) are likely to be more effective than trying to capture the effects through network charges where the signals are muted and may not be passed on by suppliers.
- In considering the residual element of charges for domestic customers, where the biggest issue is load defection (by those better able to afford microgeneration and potentially storage), the principal of fairness points to recovery of these charges primarily through a greater emphasis on a fixed charge rather than volumetric as now. The concept of charging for networks as more like an insurance policy for those who self-generate is the right one.
- More fundamentally the allocation of the residual element of charges should be recognised as being essentially a political decision – not one to be determined on purely economic grounds. The idea of linking the level of fixed charges to something like rateable value – or indeed recovering the charges through local taxes – should be properly explored. Now is the time to consider radical options and this should be on the table. Such a solution would help address both fairness issues and provide a proxy for capacity without the inefficiencies caused by placing too high a charge on capacity use (when diversity in domestic usage means an individual's peak usage is not critical). It is also arguably in line with Ramsey principles. In many public services the costs are recovered through taxation rather than usage charges and a fresh look at how this might apply to energy (including in the context of a declining gas grid) would be timely.
- Rather than simply focussing on import capacity charges should be levied on export capacity even at a domestic level given that dealing with such exports creates potential costs for the

DNO and given the distributional arguments. Customers benefit from the ability to be able to export onto the grid and should pay for that benefit.

- The same arguments apply to the recovery of social and environmental policy costs which similarly have the potential to distort underlying cost signals and encourage inefficient grid defection (in whole or in part).
- For non-domestic customers there needs to be more flexibility for DNOs to charge what the market will bear (subject to remaining within the overall revenue cap). Specifically, where faced with the prospect of a company considering an alternative private wire solution the DNO should be able to price down to their long-run marginal cost as part of a virtual private wire offering using their existing network. While this will still leave other customers picking up the residual element of costs it avoids the existing assets being stranded (with those costs – which are fixed in the short to medium term - being born by remaining customers). Some sort of risk sharing arrangement with the network company could help retain the incentive for them to recover as much as they could through the virtual private wire arrangement.
- Non-domestic customers can be expected to deal with more sophisticated charging structures including capacity, time of use volumetric charges and charges for reactive power. However even for non-domestic customers a charging structure that allows them to be clear what the charges will be for use at a particular time is vital if they are to make efficient decisions about their usage. The current Triad arrangements which are based on charging for usage at what turns out ex post to be the peak times fails that test and simply rewards those who are best at forecasting. The Triad arrangements should be replaced by a dynamic time of use charging structure (with advance notification of high charge periods) and ex ante capacity charges with a high penalty for exceeding the contracted capacity.
- As the above highlights, the issues raised by network charging are highly technical but could have huge ramifications for consumers, including vulnerable consumers. Further thought needs to be given as to how to bring the consumer voice into this debate. The Charging Futures Forum is helpful in widening the debate but remains very heavily industry dominated. A programme of consumer research and engagement is required.

References

Bell et al 2011: Bell, K., Green, R., Kockar, I., Ault, G. & McDonald, J. ‘Academic review of transmission charging arrangements’ for Ofgem

Boiteux 1971: “On the Management of Public Monopolies Subject to Budgetary Constraints” Journal of Economic Theory 3:19-40

Bonbright 1957: Bonbright, James C., American Economic Review. May57, Vol. 47 Issue 2, p386. 30p.

Bonbright 1961: *Principles of Public Utilities*

CEER 2017: Electricity Distribution Network Tariffs: CEER (Council of European Energy Regulaors) Guidelines of Good Practice (Jan 17)

CEPA 2017: International review of cost recovery issues: A report by CEPA and TNEI. Final Report (February 2017)

Citizens Advice 2016: The Tariff Transition – Considerations for Domestic Distribution Tariff Redesign in GB - Brattle group for Citizens Advice (April 2016)

Decker 2014: Literature review: principles of tariff setting and revenue recovery for network monopolies in the absence of significant network capacity constraints Final Report for Ofgem Dr C Decker & S Jones (April 2014)

Decker 2016: ‘Regulatory networks in decline’, Journal for Regulated Economics, 49, pp. 344-370

Eakins 2011: A Peer Review of Commissioned Academic Analysis for Ofgem

Energy UK 2016: Electricity charging arrangements report – Energy UK (Autumn 2016)

Evans 1993: Evans, R. (1993). "Environmental and economic implications of small-scale CHP." Energy Policy 21(1): 79-91.

GTM 2013: Greentechmedia URL https://www.greentechmedia.com/articles/read/are-utilities-ready-for-the-coming-death-spiral#gs.MftNM9k (accessed 22.1.18)

Helm 2017: *Cost of Energy Review* Dieter Helm published by BEIS (October 2017)

MIT 2016: *Utility of the Future*

NARUC 2016: NARUC (National Association of Regulatory Utility Commissioners) “Distributed Energy Resources Rate Design and Compensation” (Nov 2016).

National Grid 2018 Power Responsive – Demand Side Flexibility Annual Report 2017.

NEA 2017: Heat Decarbonisation: Potential impacts on social equity and fuel poverty. A report for NEA by Maxine Frerk and Dr Keith MacLean. (September 2017)

Newbery et al, 2005: Long Term Framework for Electricity Distribution Access Charges report for Ofgem. Tooraj Jamasb, Karsten Neuhoff, David Newbery, Michael Pollitt. (March 2005)

Newbery D.M. 2011, High level principles for guiding GB transmission charging and some of the practical problems of transition to an enduring regime, report for Ofgem

Ofgem 2015: Regional Differences in Network Charges Ofgem (October 2015)

Ofgem 2017a: Targeted Charging Review – a consultation. (March 2017)

Ofgem 2017b: Embedded Benefits: Decision on industry proposals CMP269 and CMP270 (June 2017)

Ofgem 2017c: Reform of electricity network access and forward-looking charges: a working paper. (November 2017)

Ofgem 2017d: Ofgem – Targeted Charging Review launch of a Significant Code Review (August 2017)

Ofgem 2018: Getting more out of our networks by reforming access and forward-looking charging arrangements (July 2018)

Pollitt 2016: Electricity Network Charging for Flexibility by Michael G Pollitt EPRG Working Paper 1623 (September 2016)

Ramsey 1927: "A Contribution to the Theory of Taxation" Economic Journal 37 (145):47-61

RAP 2018: Regulatory Assistance Project - Cleaner, Smarter, Cheaper: Network tariff design for a smart future by Christos Kolokathis, Michael Hogan, Andreas Jahn (January 2018)

Raymond 1918: "*What is Fair: A study of some problems of utility regulation*" by William G. Raymond reprinted by Forgotten Books (page 60)

Regen SW 2017: Grid Charging for a Flexible Future (September 2016)

RMI 2014: Rocky Mountain Institute "Rate design for the distribution edge" (August 14)

RMI 2015: The Economics of Load Defection. Rocky Mountain Institute. (April 2015)

Strbac et al 2005: Framework and Methodology for Pricing of Distribution Networks with Distributed Generation A report to OFGEM by Professor Goran Strbac and Dr Joseph Mutale. (March 2005)

Annex 1: Network Charging Principles – Summary of Views

Table 1: Network Charging Principles – Regulators' Views

	Ofgem	US (NARUC)	EU (CEER)
Cost recovery and economic signals	Cost Reflectivity Minimise distortion for "residual" elements (ie taking cost recovery by networks as read)	Efficiency in discouraging wasful use of service while promoting all justified uses (looking at both total amount and eg peak v off peak) Effectiveness in yielding total revenue and revenue stability year-on-year Avoidance of undue discrimination	Cost reflectivity (reflect costs they impose and send appropriate incentives to avoid future costs) Non distortionary Cost recovery Non-discriminatory (between users)
Practicality	Proportionality and practical considerations (ie prioritising where greatest justification, simplicity, reduced volatility)	Related practical attributes of simplicity, understandability, public acceptability and feasibility of application Freedom from controversy over interpretation	Transparency Predictability (ie can estimate costs but will vary) Simplicity (easier to respond to)
Distributional impacts	Fairness - Inc in particular impact on vulnerable customers	Stability of rates to minimise seriously adverse impacts on existing customers Fairness in how apportioned	
Public policy goals	-		
Other considerations	Ideally robust to changes in other elements of charging Do note that some people may view some distortions as "preferable" to others		Future proof Reflect different cost elements and timeframes Other tools to signal to customers (eg flexibility procurement) esp where firm response needed

Sources: Ofgem 2017d, NARUC 2016 (referencing Bonbright 1961), CEER 2017

Table 2: Network Charging Principles – Academics' Views

	Prof David Newbery (Cambridge)	Prof Goran Strbac (Imperial)	MIT
Cost recovery and economic signals	Cost reflectivity - Encourage efficient connection and operating decisions Cost recovery Minimise distortion for "residual" element subject to other policy goals the regulator wishes to pursue Aid competition	Economic efficiency (cost reflectivity) Future investment signalling Deliver on revenue requirements	Allocative efficiency Recovery of regulated costs
Practicality	Simplicity Transparency Predictability	Stable and predictable prices Determination of prices must be transparent, auditable and consistent Practical to implement	Transparency Simplicity Consistency with the rest of the regulatory framework Stability Implementation costs
Distributional impacts	Equity / fairness (option)		
Public policy goals	Meeting environmental objectives (option)		
Other considerations	Consistency with other elements of regulation (eg DPCR incentives) Aligning T&D		

Sources: Newbery et al 2005, Strbac et al 2005, MIT 2016

Table 3: Network Charging Principles – GB Stakeholders' Views

	Regen SW (Distributed energy)	Energy UK (Industry)	Citizens Advice (Consumers)
Cost recovery and economic signals	Cost Reflective Incentivise long term reductions in network costs Encourage network balancing by strengthening the appropriate locational and temporal signals Support competition	Cost reflectivity Locational signals Market signals (ie problems where signal is ex-post)	Cost reflectivity (revenue recovery implicit)
Practicality	Ensure charging regime is transparent and charges are visible to all customers	Stability and predictability Transparency	Price and bill stability Simplicity and transparency Technical feasibility
Distributional impacts	Appropriate balance of charging for generation and demand		Fairness and affordability for vulnerable customers
Public policy goals	Ensure that grid charging aligned with other energy policies around long term decarbonisation and energy security		
Other considerations	Changes made in open consultation Support integration with EU markets	Harmonisation across Europe Long term outlook	

Sources: Regen SW 2016, Energy UK 2016, Citizens Advice 2016