Cost boundaries and the speed of cost recovery

A NOTE PREPARED FOR CE ELECTRIC

Introduction

It has long been recognised that the existence of the present cost boundaries can encourage inefficient input choices by DNOs and result in a lack of comparability of cost levels within each category across the DNOs.

In this paper we argue that rewriting the boundaries on a different basis could remove these distortions and result in a more efficient outcome. However, the rewriting of the boundaries will lead to the creation of a new cost base (or bases) that contains some activities and costs that have hitherto been recovered on a pay-as-you-go basis, and some that have been recovered over a longer depreciation period of twenty years. It is therefore appropriate to consider the definition of the proposed new cost boundaries together with the speed with which those costs are recovered by DNOs in order to satisfy two desirable objectives:

1. DNOs need to operate in a framework that avoids perverse substitution incentives; and
2. the speed of recovery of DNOs' costs promotes a credible regulatory environment and sends the right signals to customers.

Perverse substitution incentives

It is well-understood that the present regulatory arrangements embody cost boundaries that could encourage game-playing and create perverse substitution incentives. For example, at present Ofgem treats a high percentage of costs incurred in repairing a fault as opex, while costs incurred in network enhancement projects are treated as capex. In practice, many projects include both repairs and enhancements and it is not necessarily clear how the costs associated with such projects should be most appropriately allocated between the repair and enhancements "pots", and to a certain extent, such allocations are made at the discretion of network managers. However, since opex and capex are treated differently within the present price control framework, there is a financial incentive to increase the level of recorded capex in preference to increasing the level of recorded opex. As a consequence, where there is an element of managerial discretion, there is scope for the allocation of costs between categories to be distorted. Such distortions would be likely to limit the extent to which Ofgem can depend on cost data to inform its decisions, reducing the

ability of Ofgem to regulate effectively. These distortions could also alter investment choices and act to increase the costs that are recovered from customers over the longer term. Given the potential negative consequences of these distortions, it is right that Ofgem regards their presence as a weakness in its prevailing arrangements.

In order to address this concern, Ofgem has proposed three potential options for an alternative treatment of costs at DPCR5. Each of these alternative treatments is focused on removing arbitrary boundaries between costs and seeking instead more equal treatment of potentially competing costs, in order to remove or reduce financial incentives to distort allocations[1]. The three treatments of costs proposed by Ofgem are:

- Option 1 - treating all costs in the same way (i.e. capitalising the same percentage of all costs into the RAV);
- Option 2 - treating all network costs in the same way, but fully expensing business support costs; or
- Option 3 - attempting to identify key trade offs to try and eliminate most of the boundary issues that can give rise to distortions.[2]

Ofgem has stated that a fixed proportion of costs covered by the first or second options would be allocated to the RAV (with the remainder being expensed), but has not yet indicated what proportion this would be[3]. Neither has Ofgem yet defined the boundary points that would be required to implement options 2 and 3. However, Ofgem has provided some guidance on the definition of the network costs building block that it might use under the second of these options in its Initial Consultation paper and its December paper, although it has not yet gone as far as clearly identifying all of the cost items that might be included. However, our understanding is that the scope of costs that could fall within the network costs building block, at least under the second option, could be relatively broad, including many costs that have traditionally been treated as opex (such as head office costs related to network planning).

---

1 We note that incentives to distort the undertaking of activities (or the allocation of the costs associated with those activities) within a single building block might continue to exist if costs associated with particular activities within a building block are more likely to be disallowed relative to costs associated with competing activities in the same building block. Consequently, any benchmarking of subsets of costs and activities within a building block will need to be considered in the round as part of an assessment of the block as a whole, in order to avoid this risk of distortion.

2 Ofgem DPCR5 Policy Paper, December 2008, pages 86-87.

3 Ofgem has further suggested that all costs under the first option and all network costs under the second option could be subject to the IQI incentive framework. This approach would reduce the incentive to allocate costs to enhancement when they are in fact repairs, for example, since it would ensure that all costs of this kind were treated identically under the price control.

Notwithstanding these uncertainties, it is still possible to characterise the arrangements as they currently apply, and in stylised terms how they could apply under options 1 and 2 for each category of costs, in the table below. The terms 'fast' and 'slow' are used in the table below and throughout the paper to denote whether the costs are expensed (i.e. fast) or added to the RAV (i.e. slow).

Cost category	DPCR4 basis		Option 1		Option 2	
	Fast	Slow	Fast	Slow	Fast	Slow
Direct operational capex	0%	100%	x%	(1-x)%	y%	(1-y)%
Faults, inspections & maintenance	76.5%	23.5%	x%	(1-x)%	y%	(1-y)%
Tree cutting	76.5%	23.5%	x%	(1-x)%	y%	(1-y)%
Direct non-operational capex	76.5%	23.5%	x%	(1-x)%	y%	(1-y)%
Engineering indirects	47.43%	52.57%	x%	(1-x)%	y%	(1-y)%
Network/investment support	47.43%	52.57%	x%	(1-x)%	y%	(1-y)%
Business support	47.43%	52.57%	x%	(1-x)%	100%	0%
Pensions (ongoing)	42.3%	57.7%	x%	(1-x)%	y%	(1-y)%

Table 1: Potential treatment of different cost categories[4]

Source: CE Electric and Frontier Economics

In order to remedy the most significant of the perverse substitution opportunities and incentives arising from the cost allocation rules, there is merit in our characterisation of Option 2, namely distinguishing between:

- network costs (i.e. costs that relate to the management and operation of the network assets or directly facilitate customer service activities such

4 We do not consider the allocation of the pensions deficit contributions in this paper, since a different (albeit related) set of considerations apply to where in the cost base they should be allocated, and over what time period.

as fault repair, network reinforcement and network design or call centre activities); and

- business support costs (such as finance or regulation).

If this categorisation is made it would be possible to subject each of these categories to different (or, indeed, to the same) incentive arrangements. We do not believe that network costs and business support costs as defined compete with one another to any material extent and hence it almost entirely resolves the distortion that concerns Ofgem. As a consequence, under Option 2, business decisions could be based on a comparison of the relevant alternatives, without reference to differences in the regulatory treatment of those alternative options[5]. Like Option 1 therefore, Option 2 solves all material distortions arising from existing boundary issues, a crucial development in the light of the sizeable investment programmes that are being undertaken by all DNOs.

Furthermore, not only would this arrangement encourage optimisation of the network cost base by the individual DNO, it would also enable Ofgem to make its efficiency comparisons and forward projections on a more consistent set of information across the DNOs. With the clarity provided by the distinction created for cost allocation under Option 2, business support costs (and to a similar but probably lesser extent indirect costs) could be readily mapped to those costs incurred by other DNOs permitting more rigorous efficiency judgements as part of the allowance setting process. The benchmarking of subsets of network costs could also be undertaken as part of a holistic efficiency assessment of the network building block as a whole – in order to avoid perverse substitution incentives arising out of the benchmarking process. In the light of these benefits arising from Option 2's categorisation of costs, Option 2 is preferable to Option 1.

Consequently, Ofgem's option 2 has much merit in promoting efficiency across the DNOs. The one outstanding issue that remains in the treatment of the cost boundaries is the recovery of those network costs over time.

The speed of cost recovery

From a purely accounting perspective, at first sight it would seem sensible to recover the costs of long-lived assets over the lifetimes that those assets are expected to be operational. However, such an approach confuses accounting depreciation with economic depreciation of assets, a point made by Ralph Turvey nearly forty years ago:

---

5 It is worth noting in this context that the incentive rates in the IQI may need to be increased in order to preserve the present overall incentive power of the regime once all network costs are treated as a single item.

While accounting depreciation is determined by some simple rule – straight-line, diminishing balance, etc – and by an arbitrary choice of life, economic depreciation is the result of an explicit optimisation calculation[6].

Whilst it would not be desirable to replace the present arrangements for calculating price controls with an optimisation calculation of the type Turvey describes, this does not mean that a regulator should not try to seek to mimic the economic effect of an economic depreciation schedule in a regulatory accounting schedule. In our view, it is still possible to infer in broad terms what an efficient depreciation schedule should look like, that takes account of the following factors:

- the price signals that customers should be exposed to in order to encourage efficient location and consumption decisions;
- the impact of the speed of cost recovery on the commitment regulators are able to provide to enable businesses to recover efficiently incurred costs;
- the impact of the speed of cost recovery on regulatory discipline and accountability; and
- the impact of the speed of cost recovery on the financeability of the businesses.

All of these factors would point towards a depreciation period far shorter than the technical life of the assets in order to facilitate efficient outcomes[7].

Price signals to customers

The concept of economic depreciation – as opposed to accounting depreciation rules – is relatively straightforward: when capacity is plentiful relative to demand the depreciation charge embodied in prices to customers should be low to encourage utilisation of the assets; and when the capacity/demand balance is tight, the depreciation charge should be high to both signal the need for new investment and to ensure that the assets are consumed by those who value them most highly.

---

6 R Turvey (1971), Economic Analysis and Public Enterprise, Allen and Unwin, London.

7 However, there are limits to reducing the cost recovery period within the present incentive arrangements applied to the DNOs. For example, a pay-as-you-go model for DNOs would fundamentally alter the balance of incentives between stewardship of the existing assets and incentives for new investment – including incentives for deferral and cancellation of plans, and if such a move were contemplated it would require a major revision of the incentive arrangements applied to investment. We do not believe it is necessary or desirable for Ofgem to make a radical move in this direction at this stage.

Economic depreciation cannot be applied mechanistically under price control regulation because the calculation of the level of the price control has evolved – for various reasons – as an accounting identity. That is, the allowable revenue for a DNO is equal to the (discounted) sum of expected operating expenditures plus a return on an index-linked asset value plus an arbitrary straight-line depreciation charge on an arbitrary asset life. It is unlikely to be feasible, or desirable, to replace that well-understood rule with a pricing rule based on an optimisation calculation of the level of economic depreciation since to do so would fundamentally alter the dynamic of the relationship between the regulator and the DNOs. However, simply because one should not abandon arbitrary depreciation rules in favour of optimised economic depreciation rules does not mean that a regulator should not try to seek to mimic the economic effect of an economic depreciation schedule in a regulatory accounting schedule.

A sensible starting point in evaluating what an economic depreciation schedule might look like is to recognise that economic depreciation is closely linked to the concept of marginal cost pricing, which Ofgem is presently seeking to introduce through the common charging methodology applied to DNOs. When capacity is plentiful, the marginal cost of serving additional demand is low; and the marginal cost is high when capacity is scarce. The economic depreciation charge should therefore be seen in the same light as marginal cost pricing in that it seeks to encourage efficient locational and consumption decisions within and across networks.

In its consultations on DUoS tariff charging Ofgem has indicated a preference for tariff setting models that are based on long run incremental costs[8] in order to encourage economic efficiency, an implication of which is that tariffs should signal to users the cost consequences of their consumption decisions.

However, the difficulty under the present regulatory arrangements is that the target revenue base – a key determinant of tariffs - is not only not based on a forward looking assessment of how incremental demand might drive future costs, but does not even accord closely with *recent* historic capital expenditure. Instead, revenue allowances are based on recovering historic costs and a proportion of costs in the immediate five-year future.

Consequently, even if the tariff "benchmarks" produced by typical DNO tariff charging models are well designed and appropriately embody long term incremental costs, these must be scaled in order that they match the allowed revenue target. The prices that result from this process will not (unless by accident) signal the degree of scarcity or the extent to which future consumption decisions might drive incremental cost. The present environment of increased

8 See for example Section 3.42 of "Structure of electricity distribution charges: Consultation on the longer term charging framework", May 2005

Cost boundaries and the speed of cost recovery

investment needs is suggestive of an increasing scarcity of capacity that would merit further regulatory consideration of whether the target level of revenue recovery to which tariff yardsticks will be scaled should more closely reflect expected future network costs, rather than historic network costs, while still allowing recovery of historic expenditure.

At the very least, it would clearly be inappropriate to slow down the overall speed of cost recovery under new boundary definitions of cost.

Regulatory credibility and discipline

Ofgem has considerable discretion in its regulation of networks, in terms of the level of costs it allows the businesses to recover through the price control, the speed with which those costs can be recovered, and the incentives embodied in the cost-recovery mechanisms.

This discretion has both advantages and drawbacks, in the sense that it enables flexibility of regulatory response to unforeseen events, but also makes it very difficult for the regulator to commit to a particular course of action. Credible commitment to long-term objectives is important because it prevents the regulator from taking benefits in the short run that could compromise those long-run objectives. Well known examples of these problems from the regulatory economics literature include the regulator reneging on high-powered incentive contracts to capture profits once the profits have been revealed (to the detriment of longer-term efficiency gains), and reneging on the recovery of long-lived investments (to the detriment of future investment). Good regulation strikes the right balance between flexibility to deal with short-run shocks and credible commitment to long-term goals, which is crucial to promote dynamic efficiency.

Dynamic efficiency

Dynamic efficiency gains represent the most significant efficiency benefits that can be obtained from a well constructed incentive regime, especially when applied to an industry with long-lived assets. Whilst productive efficiency gains can be made by making existing assets work harder, dynamic efficiency gains are made through optimising investment in new assets and new technology.

The constant pressure of incentives – derived either from market disciplines or well constructed regulatory rules – creates an ongoing and sustained pressure on management to ensure that investment decisions are optimised on an ongoing basis. This optimisation process encompasses:

- the robustness of the asset management programme;
- the decision about the type of investment
- the decision about the scale of the investment
- the decision about the timing of the investment
- the decision about the procurement of the kit
- the implementation of these decisions

The optimisation process described here covers all types of investment decision, not just in physical plant, and sub-optimal behaviour at a single point in time has a cumulatively detrimental effect, relative to the outcomes associated with the right decision.

The credibility of the regulatory regime in respect of the incentives for efficient behaviour and the recovery of efficiently incurred long-lived investment costs is crucial to the promotion of dynamic efficiency.

The speed of cost recovery raises two important credibility issues that could impact on incentives to maintain and improve dynamic efficiency. The first is the familiar one that regulators are unable to credibly commit to very slow recovery of costs (for example through long depreciation periods), and regulatory risk through possible stranding of assets will increase with the effective length of time over which companies are able to recover their costs. Adoption of longer depreciation periods may reduce the cost borne by current customers through depreciation charges in any particular period, but it may also have the offsetting effect of raising the cost of capital required to finance the assets and/or reducing the investment that is actually made.

The second credibility issue is rooted in the strand of economic theory that emphasises government failure as a source of inefficiency, analogous to the better known problem of market failure. Whereas markets may fail to generate efficient

outcomes for a variety of reasons (the presence of externalities, information asymmetries and conditions that give rise to monopolies, etc), government failure may also fail to deliver efficient outcomes due to, for example, rent-seeking by government agencies leading to capture by interest groups; and populist or electoral pressure that forces agencies to reject efficient solutions in favour of electorally acceptable ones.

If the speed of cost recovery is too slow, then this can have the effect of increasing the attractiveness of rent-seeking by regulatory agencies because the cost to customers of that behaviour is obscured by the long depreciation period. For example, there may be an increased concern that regulators may have less incentive to evaluate thoroughly the cost submissions made by businesses, or may acquiesce in passing through costs that are not necessary for the provision of the service but are politically convenient to load onto customers. Consequently, a faster rate of cost recovery can have the beneficial effect of disciplining the regulatory agency to act effectively and in accordance with its obligations.

Financeability

Financeability considerations have been a relevant component of price control determinations since the electricity, gas, telecoms and water businesses were privatised and subjected to incentive-based regulation, and have probably been made most explicit, over the longest period of time, in the water sector.

The role of the financeability tests is, in essence, to maintain investment grade status, usually by ensuring sufficient cash-flow is generated by the business. In principle it is worth reflecting on why such tests are necessary if every component of the price control calculation is objectively correct. The answer is likely to be that the price control calculation itself does not capture all of the factors that provide the right signals for management to invest and for investors to provide the funds. The most obvious example of this problem is that discussed above, which is that in some situations it is likely that investors and businesses may not believe that the regulatory authority will be able to continue to make the objectively correct decisions at all times in the long lives of the assets that are going to be sunk. In other words, the regulatory authority is not able to credibly commit to the recovery of sunk costs over long periods. Indeed, the explicit financeability metrics designed by Ofwat in the early 1990s resulted in part because Ofwat, as a new regulator, had not been able to establish credibility in the financial markets, and so the tests had the effect of restricting Ofwat's own room to manoeuvre to push cost recovery too far into the future.

It is often argued that if there is an impending cycle of investment that regulatory depreciation on historic investment would not be able to cover, then this could create cash-flow difficulties for the business, and hence a threat to its investment grade status, which financeability criteria exist to deal with. In this case however,

it is clear that the problem is that the depreciation lives are likely to be too long from an economic perspective. As discussed above, in situations where the demand-capacity balance is tight, or if new demands are being made on the network for which capacity is currently inadequate, then these resource costs should be being signalled to the customers who are creating them.

These two examples have served to show that financeability considerations are important to regulatory regimes since they serve as a backstop to the formal arrangements and capture characteristics in the formal regime that are either inadequate (e.g. regulatory reputation) or wrongly specified (e.g. the regulatory valuation of the assets, the WACC or length of depreciation period). Indeed, the low cost of capital that regulators have been able to derive from market data, and to make use of for the benefit of customers, is in part a function of the financial market's expectation that financeability criteria will be applied and also that the timescales within which investments will be recouped will be consistent with the depreciation lives assumed in the price control reviews of the recent past.

The relevance to the speed with which costs are recovered is clear – if the underlying depreciation schedules either do not reflect the capacity-demand balance or provide an inadequate prospect of cost recovery over the life of the assets, then financeability considerations will be brought into play.

Ofgem practice to date and options for DPCR5

As discussed above, the regulatory depreciation span for costs incurred in maintaining and developing a network should be well below the technical lifetime of distribution assets. Indeed, in practice we observe that in competitive markets effective depreciation periods that companies use in appraising similar major investment projects are well below the technical lifetime of the assets. For example, evidence from the entry of independent power generators (IPPs) during the 1990s suggested that these plants operated on business plans that embodied asset lifetimes significantly shorter than the technical life, and long-term contracts for the sale of the power.

At DPCR4 those costs treated as capex were depreciated at a 20 year rate, while a significant proportion of network costs which were classed as opex were returned to companies on a pay as you go basis.

Whilst all of the factors discussed above have led to the use of 20 years as the regulatory asset life, it is likely that it was primarily due to the fact that the regulatory asset value in the price control calculation was too low from an economic perspective – which first necessitated bringing financeability considerations into play. The DNOs were privatised in 1990 at values significantly less than the economic value of the assets (which for these purposes can be approximated by the modern equivalent asset value), and it was these values that were used as the basis for regulatory asset valuation. This under-

valuation in turn meant that both the expected returns and cash-flows generated by the DNOs were lower than their respective economic values. Within this environment, lengthy depreciation periods would have further diminished cash-flow and compromised the investment grade status of the DNOs. So, whilst the regulatory valuation may have promoted an equitable settlement between customers and shareholders, this valuation would have led to unsustainably low cash-flows which would have needed to be remedied either by shorter depreciation lives or by a higher cost of capital. As it is, the depreciation life has been the parameter used to ensure an adequate flow of cash to ensure that the businesses remain viable.

At DPCR5 these issues will continue to influence the extent to which cost recovery can (or should) be pushed too far into the future, but the key difference with the recent past is that the scale of investment that is expected over the next twenty years is clearly indicative of a tightening demand/supply balance on the networks, and Ofgem should therefore consider whether a higher effective depreciation rate would be appropriate, in order to strengthen the price signal being sent to current customers.

At the very least, it seems clear that Ofgem should seek to maintain the speed at which costs were recovered at DPCR4 under new boundary definitions, which implies that it would need to expense about 20-25% of the costs contained within the network cost building block. If Ofgem wished to send stronger signals to customers then this proportion would need to increase further.

Conclusions

In this paper we have made the following suggestions for the treatment of costs at DPCR5:

- Costs should be clearly separated between network costs and business support costs in order to:
  - promote efficient input choices by DNOs;
  - enable effective cost comparability and assessment of business support costs and (to a similar but probably lesser extent) indirect costs; and
  - permit cost assessments within the network building block to be undertaken in the round, which necessarily would take into account the scope for substitution of competing activities.
- Within the network building block, the speed at which costs have been recovered in the past have depended on a variety of factors, all of which have pointed to a regulatory asset life significantly lower than the technical life of the physical assets.

- At DPCR5, there is no compelling reason to suggest that the speed with network costs are recovered should be slower than that which has prevailed so far. Indeed, in the light of the fact that the investment requirements of the networks are forecast to increase considerably, it may be desirable for Ofgem to signal the tightening demand/supply balance to customers by increasing the speed of cost recovery.

Frontier Economics Limited in Europe is a member of the Frontier Economics network, which consists of separate companies based in Europe (Brussels, Cologne, London & Madrid) and Australia (Brisbane, Melbourne & Sydney). The companies are independently owned, and legal commitments entered into by any one company do not impose any obligations on other companies in the network. All views expressed in this document are the views of Frontier Economics Limited.