
Annex D

Public submission on vertical integration and separation

Professor George Yarrow

June 2008

Regulatory Policy Institute

Reflections on policy issues raised by next-generation access networks in communications

Professor George Yarrow
Dr Christopher Decker

June 2008

Views expressed in this and other Regulatory Policy Institute reviews and reports are those of the authors, and they should not be attributed to any organisation(s) with which the authors are associated. The Institute is a charitable organisation dedicated to the promotion of the study of regulation for the public benefit, and it does not itself take positions on the relevant policy issues.

The Institute welcomes general and specific support grants for its various activities. In this case funding for the work was provided by Telstra, and we would like to express our thanks to that company for granting this opportunity to contribute to thinking in an important area of public policy.

© Regulatory Policy Institute, 2008

31-33 Westgate, Oxford OX1 1NZ, UK.

Reflections on policy issues raised by next-generation access networks in communications

The introduction of next-generation access networks (NGANs) in communications changes the market and technological context in which public policy operates. It does so in a way that, in terminology used in the economic analysis of innovation, can be said to be ‘drastic’ in nature. This poses a number of new challenges for the development of regulatory policy, some of which are identified and explored in this paper.

When new problems arise, one of the first analytical tasks for regulators is to discover perspectives that might be of assistance in thinking through the issues that emerge. These alternative ‘frames’, which will also be explored below, are not solutions to the policy problems, but rather are ways of approaching the issues to assist the development of policy positions.

On this basis, this paper proceeds as follows:

- We begin by identifying the core policy issue: that changing economic and technological circumstances and public policy imperatives are creating requirements for major capital investments in NGAN assets in telecommunications, and that these investments are characterised by significant uncertainties. The issue of facilitating major investments in the context of substantial uncertainty is not unique to telecoms, and broadly similar policy issues have recently arisen in the energy and transport sectors. There are, however, a number of distinguishing features that characterise the policy issues in relation to NGANs, including: the nature of the demand uncertainty; the likelihood and pace of future technological change; and, the impacts this can have on the incidence of ‘network bottlenecks’ and on the appropriate scope of regulation.
- While some, similar policy issues have arisen on a smaller scale in relation to telecoms networks in the past, the nature and scale of NGAN investments presents a much wider and more complex set of challenges for the policy framework. More specifically, it is noted that the contextual challenges associated with NGANs require a fairly fundamental reassessment of how policy is developed and applied, and, in particular, of the appropriate trade-offs between the different categories of efficiency, and of the different potential ‘balances of competition’. We suggest that, if policy does not adapt to reflect a different balance of potential sources of consumer welfare gains in the NGAN context, this may introduce new regulatory risks and may adversely affect the form, scale, timing and location of investments undertaken.
- The paper then briefly reviews how the ‘standard regulatory model’ (as developed in several jurisdictions over the past twenty years or so) has been applied, both in general terms and more specifically in relation to telecoms networks. We suggest that while the standard regulatory approach may have

been ‘fit for purpose’ for traditional telecommunications networks – albeit with some shortcomings – simply ‘rolling-forward’ or ‘force-fitting’ this approach to the changing NGAN environment is likely to be detrimental both to the incentives to invest in NGAN assets, and to the development of competition over time. A particular issue identified is that the standard approach to pricing in telecoms has effectively conflated the regulator’s role in determining the *average* level of allowed revenue with a role in determining the *structure* of relative prices. It is argued that such an approach will no longer be appropriate in the NGAN environment.

- We then address a number of general issues relevant to the overall design of regulatory frameworks for NGANs. First, we examine the issue of regulatory discretion and note that, in the NGAN context, there should be a preference for longer term regulatory reviews, with the possibility of regulatory intervention only in limited, specific and predefined circumstances. We note that this does not equate to the setting of long-term fixed prices, and that complete pricing rigidity could be detrimental to efficiency in both the short and long-term. Second, we note that the design of the regulatory framework should probably not be unduly complex, and that a ‘principles-based’ approach, which provides market processes with considerable scope and incentives for discovery (of all types of relevant information – about technologies, products, consumer demands, effective business arrangements, etc.), rather than a more prescriptive approach based on market ‘models’ and ‘scenarios’ is likely to be a much more effective approach, particularly given the significant uncertainties surrounding NGAN investments and the scale of those investments.
- In relation to the often misunderstood issues of vertical integration and structural separation, we begin with some ABCs of the economics of integration and separation. The underlying economic landscape is that of *complementary* activities. The general presumption about integration of complementary activities is that it tends to lead to lower prices and higher output (and hence to be pro-competitive), and, particularly in the developmental stage of new, complementary activities, to higher investment. Strong incentives to abuse dominant positions characterised by vertical integration are caused chiefly by price regulation, which heavily constrains profits at a particular point in the vertical chain. Structural separation is, in effect, usually a remedy for incentive distortions that would not exist but for tight price controls. Our view on leverage/foreclosure incentives in the absence of cost-of-service regulation of NGANs is that they can be expected to be weak because, absent such regulation, the network operator has strong incentives to promote demand for use-of-network.
- We conclude by considering the features of alternative ways forward for regulatory policy. First, we note that the significance of the distinction between persistent/permanent and temporary monopoly positions since, far from being a problem, temporary monopoly tends to be a key driver of dynamic competition and of the realisation of dynamic efficiencies. Given the significant uncertainties surrounding NGANs, we suggest that a regulatory commitment not to bear down on average use-of-network prices/revenues for a

defined period could be used to rebalance incentives towards promotion of investment and discovery. On the other hand, such price deregulation could give rise to unwarranted price increases for what, in effect, are existing products provided via existing network assets, at least in locations where inter-network competition is weak. Price deregulation could also have a potentially chilling effect on investment in services, arising from possible fears of hold-up. We suggest that adaptations of familiar market responses to such problems, in the form of longer-term agreements, can likely play significant roles in addressing these specific issues. Agreement between the network operator and the regulatory authority on a limited set of ‘anchor products’ is one possibility here, as would be opportunities for network users to contract, with the operator, for use-of- network or network services on a longer term basis.

1. The core issue

In the standard analysis of network utility regulation it tends to be taken for granted that *but for* price regulation of the relevant activities, a privately owned network operator would raise prices to levels substantially above those that could be justified on any cost-of-service basis. The supplier might not be able to do this for each and every regulated activity – there may be stranded assets in some areas – but, for the most part, and taking the regulated activities as a whole, the ability to set an unconstrained price level high enough to recover efficiently incurred costs is not considered a problem. Rather the perceived problem is that prices may be set ‘too high’; that, after all, is the official *raison d’être* of price control.¹

The presumption here is so commonplace that it is sometimes not even noticed. However, with increasing frequency, changing technologies, changing policy priorities (e.g. climate change), and changing factor prices are presenting problems and challenges to regulators where the underlying assumption breaks down, and its continuing presence in policy thinking, often implicit and unarticulated, poses risks to sound policy development.

Among the most important of these challenges are the issues surrounding major new infrastructure projects in regulated sectors. A quick glance at relevant episodes in economic history will quickly illuminate one, generally forgotten lesson: major infrastructure projects are frequently characterised by heavy financial losses. Thus, railroad bankruptcies were common in the nineteenth century and the first half of the twentieth century.²

¹ In practice, of course, the rationales for many types of regulatory intervention are very frequently somewhat less elevated than is implied by official narratives. One of the central, consensual findings of a large economic literature on the effects of regulation is that “As a positive theory of regulation, the normative theory of welfare economics [i.e. the public interest theory of regulation] is obviously incorrect” (P. Joskow and R. Noll, “Regulation in theory and practice: an overview”, in Fromm (ed.), *Studies in Public Regulation*, MIT Press, 1981). See also P. Joskow and N. Rose, “The effects of economic regulation”, in R. Schmalensee and R. Willig (eds.), *Handbook of Industrial Organization, Volume II*, North-Holland, 1989.

² See, for example, Daniel Schiffman, “Shattered Rails, Ruined Credit: Financial Fragility and Railroad Operations in the Great Depression”, *Journal of Economic History*, Cambridge University Press, 2003.

Looking forward rather than back, three current UK examples illustrate today's policy issues:

- Very major investment requirements have emerged in the gas and electricity sectors. Whilst it is reasonably certain that large capex programmes are necessary for efficient network development, there is much greater uncertainty when confronted with the question: *where should the capacity be added, and when?*

In gas, where the scale of the investment requirements is probably lower, the issue is related to changing patterns of gas flows (older North Sea fields to the north and east are being depleted and replaced by new LNG and interconnector flows, which tend to be located further south and west). Since the locations of new points of origin of flows are uncertain – for example, LNG terminals can be constructed at a number of different coastal locations – there is a consequential risk of asset stranding: pipelines may be built, but, in the event, the level of demand that turns up may not be sufficient to allow cost recovery.

In electricity the driver is environmental policy, which is leading to major shifts in the demands placed on distribution and transmission networks. New windmills tend to be sited in rather different, and more scattered, locations than thermal power stations.

Whilst stranding risks are not a new phenomenon, what *is* different is the scale of the investment requirements. If a project is minor in relation to the activities of the network operator as a whole, revenue deficiencies connected with the project can easily be made good from other activities. When the projects are large in relation to, say, the regulatory asset value of the operator, such pragmatic adjustments may be either infeasible or, if feasible, so damaging in their economic effects as to be impracticable for public policy purposes.

- Major expansions in the capacity of airports tend to be lumpy decisions, and there is currently particular controversy over the projected expansion of London Stansted (London's 'third' airport). Stansted, which has always faced more uncertain demand prospects than Heathrow and Gatwick airports, has, until very recently, been unable to price up to levels allowed by the Civil Aviation Authority's price cap.

The perceived demand uncertainties have recently been enhanced by a combination of factors: rising and more volatile fuel prices; an increased but inevitably variable propensity of government to increase airport taxes, justified in terms of environmental effects; and, prospects for further substantial increases in input costs as a consequence of climate change policy measures yet to come, for example by including airlines within the EU emissions trading scheme. Whilst the questions of *where and when* additional airport capacity might be added have always been around, the uncertainties now are such as to

raise the more radical question of whether more capacity should be added at all.

In the face of these challenges, the Civil Aviation Authority has expressed the view that the traditional ‘building blocks’ approach³ can no longer be considered satisfactory, chiefly because of its potentially distorting effects on investment decisions.⁴ It has further recommended that Stansted airport should be deregulated, although the Department of Transport has rejected that proposal for the time being.

- In the case of the communications regulator, Ofcom, the equivalent challenge concerns the development of a policy approach appropriate to the issues raised by NGANs. In this case, the situation may appear somewhat different in that NGANs are replacements for an existing network, which at first sight appears unlike the opening up of new routes for pipelines and transmission wires, or the addition of new airport capacity. The major interest in NGANs, however, lies in the possibility of developing new services for consumers, at least when the policy focus is on a broad roll-out of new connectivity, rather than, say, on serving a more focused customer group, such as large, urban, commercial customers. The trade-off is between the costs of a major investment programme and the value of the *incremental* benefits (over and above what could be provided in the relevant counterfactual). As in relation to the *wheres and whens* of the energy and air transport investments just discussed, such incremental values are subject to considerable uncertainties. However, in the telecoms case, the benefits may also depend to a large extent upon the characteristics of activities and services which do not yet exist and which remain to be invented or ‘discovered’.

Our focus will be on NGAN issues in telecoms in what follows, but the above discussion serves to illustrate the generic nature of the underlying or core issue. It can be characterised broadly as follows:

- large capital requirements;
- indivisibility and irreversibility in investment;
- uncertain demand, such that there is significant risk that efficiently incurred costs cannot be recovered, even with fully deregulated pricing; and
- inability to resolve major uncertainties ahead of investment requirements.⁵

³ The ‘building blocks’ approach is a form of cost of service regulation in which the overall price cap or revenue allowance is built up from detailed consideration of contributing components (such as opex, capex, the regulatory asset base, depreciation, and the cost of capital).

⁴ See *Inquiry into airport charges at Stansted Airport: Issues statement*, Competition Commission, May 2008, available on the Competition Commission’s website.

⁵ This is not to imply that it is impossible to reduce uncertainties and risks at all, since, for example, the latter are clearly affected to some extent by the conduct of public policy. Rather, the important point is to recognise the practical limits on what can be done in this direction.

2. The NGAN context

Although there are commonalities among the major investment issues currently presenting themselves across regulated sectors, each case also has its own, more specific features, which differentiate it from the others. One of the fundamental principles of economic case studies, which is fully shared with legal analysis, is that “context is everything, circumstances alter cases.”⁶ It is not our purpose here to attempt to develop a full analysis of policy options for NGANs, but rather to develop and explore ways of approaching such an analysis. Nevertheless, it is relevant to take account of a number of specific features of the issues.

Perhaps the major difference from the cases discussed above is that the incremental revenues that can be expected from investment in NGANs are driven by the anticipated introduction of *new* products and services. It is not just a matter of anticipating more demand for the same product (airports) or more demand in different places (energy), but of creating new demand via product/service innovation.

Two consequential points might be noted at this stage:

- demand (for use of NGANs) assessments are likely to be particularly uncertain. In the colloquial way of putting things: we are not only dealing with known unknowns, but also with unknown unknowns; and
- the incremental revenues attributable to the prospective investment should be measured against a counterfactual based on the continued operation of existing networks, including possible future developments of those networks. Since increasing bandwidth and new products/services can be expected as a result of technical advances applied to existing access networks, there is also non-trivial uncertainty surrounding the counterfactual.

The incremental dimension of the relevant policy evaluation is of particular significance. At least over some geographic areas, access is considered to be a bottleneck, but it is only a bottleneck in relation to the provision of those products/services requiring bandwidths in excess of those available from any existing, alternative access technologies, such as wireless and/or HFC cable.

Suppose we divide consumer services into a set (S_o) that can be provided with existing technologies over competing access networks (of which voice telephony is the obvious example), and a currently fuzzy and undefined set (S_n) that requires higher bandwidth. The prospective bottleneck relates to S_n only.

It can also be expected that these sets will be in a continuous state of flux, with newly discovered products/services being added to S_n , and migration of products from S_n to S_o as technical progress in other access networks take place.

⁶ Sir Christopher Bellamy, “The Competition Appeal Tribunal – five years on”, in Colin Robinson (ed), *Regulating Utilities and Promoting Competition: Lessons for the Future*, Edward Elgar Publishing, 2006.

The fact that a product belongs to set S_n (i.e. is not capable of being supplied at an acceptable quality by an alternative access network) does not, of course, mean that there are no substitutes for NGANs. At a general level, there is always some substitutability among products and services because of what, in economics, are called income effects, and in this case lower income households in particular may well choose to spend their money on things like food, energy and housing rather than on obtaining the latest communications innovations.

More specifically, there will be more direct substitutes for some of the products/services in S_n from other communications services, the most obvious of which is digital television which can be transmitted terrestrially, by satellite, and by cable. Increased broadcasting capacity, coupled with innovations in household data storage and retrieval equipment, provide increasing flexibility in the 'consumption' of audiovisual content acquired from broadcasters, which inevitably affects willingness to pay for at least some of the kinds of services that might be available using NGANs.

Such substitutes would likely not be classed as direct competitors to a NGAN, in the sense that they are unlikely to be close enough substitutes to be judged to be in the same relevant antitrust market (for competition law purposes). They can, however, be expected to affect both the level of demand for use of NGANs, and the price sensitivity of that demand. It is a little like football. There may be only one club in town, and only one match available to go and watch on a Saturday afternoon, but there may be plenty of red ink in the accounts of the club if performance is not so good and punters switch to any of a large number of alternative uses of leisure time.

The policy problem is therefore one of how to deal with a possible bottleneck input for a fuzzily defined, and probably constantly changing, set of (currently) non-existent products (denoted S_n in the above notation), whose levels of demand and price sensitivities are uncertain and are likely to change over time as alternative technologies move forward. That is quite a challenge.

A NGAN is only a *possible* bottleneck because, in high demand density locations, there will frequently be more than one, and in some cases several, NGANs available. Thus, in the Australian case there are likely to be important parts of the national footprint of the proposed National Broadband Network in which there will be inter-NGAN competition, whether from existing HFC networks, higher speed broadband wireless, or city-based fixed broadband networks.

This (now factual) observation about the feasibility of inter-NGAN competition in many local areas leads immediately to questions concerning geographical market differentiation, which is generally a matter of central importance for the understanding of network economics but which, at least in the application of competition law, is often a relatively neglected issue.

In the EU, for example, the lumbering behemoth that is telecoms regulation provides for an exercise that proceeds from market definition, to assessment of market power, to, where considered appropriate, adoption of (proportionate) remedies. In principle, this allows for the definition of geographically differentiated markets according to customer density and for the application of what are, in effect, different policies in those different markets. In practice, the flexibility potentially available is generally

underutilised, and the arcane complexities of the procedures appear to be one of the obstacles here.

For example, focusing first on ‘antitrust markets’, defined by procedures that are commonly adopted in the enforcement of competition law but at least some of which are conceptually suspect, is not necessarily how an unencumbered economist would approach policy assessments in the context of a network characterised by a complex and shifting matrix of substitutable and complementary activities. The natural analytic tendency would be to evaluate the matrix as a whole, substitutes and complements together, and not to slice-and-dice activities into sub-sets of close substitutes, which are then analysed as if detached from the rest of the activities.

To the extent that this tendency to hold on to the bigger picture influences regulatory thinking, the consequence can be a tendency to define relatively broad geographic markets. That is fine so far as it goes, but if policy measures are then dependent on market definition – for example, because policy constraints artificially link market definition, assessment of market power and remedies (as in the EU approach) – the result can be a ‘one-size fits all’ approach to locational issues in which policy is poorly targeted and, since decisions are often heavily influenced by the position in lower density locations where market power issues tend to be greater, to generalised over-regulation.

This last point leads on to another issue of particular significance in communications, namely the precise meaning of the concept of universal service. As technology has moved forward, the tendency has been to broaden the scope of the services that, for one reason or another, policymakers consider should be available to all, at reasonably similar prices. This can get very expensive, raising the question of how it is to be paid for.

Historically, USOs were funded by cross-subsidisation, but that meant suppression of market competition. Recognition of the scale of the *indirect* costs of such suppression of competition has been one of the motivating factors for the liberalisations of recent decades. However, the fundamentals do not change, and wherever the prospect of substantial funding requirements for subsidisation of a part of a wider supply exists, there is always a risk of measures that seek, one way or another, to restrict markets in ways that help sustain/finance the cross-subsidies.

In principle, the approach of unbundling supply subsidies, adopted for example in the NGAN policy of the Australian Government, can allow for the achievement of distributional objectives without distorting the functioning of the relevant markets. In practice, there always remains a risk that, in order to reduce the quantum of public subsidy required, unnecessarily restrictive constraints will be introduced on how a supplier can operate in serving unsubsidised accounts.

Finally, in this short discussion of some of the more specific sectoral characteristics of significance for regulatory policy development in communications, two further aspects of NGANs are worth highlighting:

- The NGAN technologies of current interest change the structure of the relevant networks at the access level. In particular they substitute in whole

(in the case of fibre to the home) or in part for existing local loops. This indicates that in communications it is not only the set of products/services that is changing relatively quickly over time; network structures are also changing, at rates and on scales generally unmatched in energy and transport networks. This emphasises the importance of framing economic assessments at the ‘whole network’ level, and the dangers of over-focusing on particular sub-sets of activities. To illustrate, compared with some other EU Member States, the UK came rather slowly and slightly reluctantly to policies of local loop unbundling (LLU), having previously preferred policies of promoting network competition at the access level. Having made the jump to this more narrowly focused approach, Ofcom now finds that NGAN roll-out would strand investments (at local exchanges) that policymakers have only relatively recently started to promote and encourage. There are, therefore, clear risks in basing regulatory approaches too closely on existing network structures in circumstances where those structures are themselves changing as a result of technical progress.⁷

- The introduction of local NGAN tends to *displace* an existing access network, at least in fibre-to-the-node (FTTN) implementations.⁸ This is a rather fundamental point for regulatory policy since, if the NGAN were simply an additional alternative to the existing local access network, it might be expected that it would simply be welcomed as a competing alternative that lifted the monopoly bottlenecks that exist in some locations. In our view, therefore, the policy issues that arise are centred on the consequences of the fact that, in rolling out a NGAN, an alternative means of delivering a number of basic products/services, of which voice telephony is the central example, would be removed. The structure of competition among access networks is therefore largely unchanged: in those locations where there is already access competition, the number of competitors will not change, and in those locations where there is an access bottleneck there will continue to be a bottleneck.

3. What is really different in all this?

At the risk of repetition, issues/questions that become much more important when considering such a major change in network structure as NGAN roll-out can be summarised as follows:

- *Whether or not new network will be built at all.* Traditional regulatory approaches tend to take the basic network structure as given and address investment in terms of incremental adjustments (although it can be noted that, even in these less challenging contexts, capex issues are often the most difficult for regulators to address in price reviews).

⁷ It is tempting to use the metaphor of an over-rigid regulatory policy or ‘model’ itself becoming “stranded” as a result of technical progress, but regulators can sometimes have the (monopoly) power to change the market realities when the ‘model’ that they are using is not appropriate, in which case there is, unfortunately, nothing “stranded” about the policy.

⁸ When fibre-to-the-home (FTTH) replaces a copper network, it is possible that the original copper network might continue to be used (as it is, for example, in some of Verizon’s territories in the U.S.).

- *The time at which the investment should be made.* Uncertainty (and particularly demand uncertainty) coupled with requirements for irreversible commitments implies the existence of option values, which tend to become a much more important aspect of investment appraisal. These values depend upon assessment of future information flows (what might we know better tomorrow that we do not know so well today?). When the projects are large in scale, option values become larger on that count also (i.e. not just because of greater uncertainty)
- *Choices in relation to technologies and alternative network architectures.* For example, different network layouts, with potentially different implications for public policy, are associated with FTTN (fibre to the node) and FTTH (fibre to the home).

Speaking generally, there is a wider menu of choice than is normally relevant to policy decisions in regulated sectors, mirroring the wider range of demand outcomes that are associated with greater uncertainty (what potential products might be available, what will consumers want, and how much will they be willing to pay?).

At a relatively deep level, none of this affects the broad economic and policy principles to be applied, but the new context does affect, in potentially radical ways, how those principles are best developed and applied. This can be illustrated by reference to two important, inter-related sets of trade-offs concerned respectively with efficiency and with competition.

Efficiency trade-offs

It is now conventional in many policy areas to distinguish between (a) allocative efficiency, (b) productive or cost efficiency, and (c) dynamic efficiency. The first two of these categories can be given precise definitions, whereas the third – which typically refers to the capacity for growth in consumer welfare over time – is a little more nebulous, largely because economic models of market dynamics are much less developed than static theories. There are good reasons for this imbalance in precision: the driver for economic progress over time is discovery, and it is hardly surprising that it is difficult to analyse that which is yet to be discovered

The key point for current purposes is that there are trade-offs between what can be achieved in each of the various dimensions of performance. Consider, for example, the evaluations surrounding the introduction of RPI-X or price-cap regulation. A conscious policy decision was taken at the time to move away from strict cost-of-service regulation, deliberately weakening the links between prices and costs so as to create greater incentives for improvements in productivity and reductions in costs.

The contextual judgment underlying this decision was that the utilities of the time were characterised by cost padding, by virtue of having operated as public or private monopolies for extended periods. Hence, the potential for gains from increases in productive or cost efficiency was perceived to be much greater than any associated allocative efficiency losses that would flow from the fact that prices might diverge from costs by predetermining prices for a period.

Over time, the RPI-X price-capping approach has reintroduced a greater emphasis on costs, via the building block approach that has been widely used in price reviews. In part this has reflected re-emergence of allocative efficiency considerations associated with investment, which were, of course, influential in the development of rate of return regulation in the USA. In particular, for a number of different reasons, price-cap regulation, in which rates are periodically re-set without any very clear and explicit linkage to costs can have a negative effect on investment (most obviously when regulated prices fall below costs). Moreover, as productive/cost inefficiencies are driven out of utilities, the relative balance in the potential gains from further productive/cost efficiency improvements, as compared to prospective gains from other types of efficiency improvement, necessarily change. One period's successes change the context of later assessments, so that applying the same principles inevitably has led to changes in the balance of policies implemented.

Whereas the shift in the balance between sources of potential efficiencies just described is largely endogenous to the regulatory process itself, the much bigger changes, and those that are now giving rise to the greatest challenges to policymakers, are arguably exogenous to network regulation. The two most important are:

- radically shifting priorities in environmental policy, which are driving the problems in energy and air transport networks referenced in section 1, above; and
- technological change, which is the more important influence in communications.

Both of these developments come with large uncertainties attached to them, albeit for different reasons. In both cases it is reasonable to expect that, for network companies and for policy makers, the future will be significantly different from the past, in as yet unknown ways. An immediate corollary is that issues of dynamic efficiency should be very much to the fore in policymaking.

In a sector such as energy, dynamic efficiency has been a factor that, to say the least, has not been particularly prominent in regulatory policy development. In the UK, for example, R&D expenditure in energy fell significantly following privatisation, and whilst there have more recently been some attempts to introduce schemes to incentivise certain types of network innovation, these are typically *ad hoc* add-ons to price control reviews, with marginal effects at best.

Dynamic efficiency has tended to be referenced more in regulatory documents in the communications sector, although even here, in the absence of any very clear sense of what precisely is at stake, it is hard to discern that analysis of its trade-offs with other forms of efficiency have been particularly prominent in policy thinking.

Competition trade offs

Dynamic considerations have featured more in assessment of different 'balances of competition'. The idea here is that, just as there are trade-offs in the pursuit of different sources of efficiencies, so there are trade-offs among the different forms that

competition might take. The best known example of this is the efficiency case for patents (or IPRs more generally) that, by establishing monopoly rights, allow the charging of prices for the use of information which may far exceed the costs (of use of the information). The familiar argument here is that it is the prospect of monopoly returns from control over the use of information that provides the incentives for the production of new and valuable information in the first place. Moreover, the textbook analysis of the determination of optimum patent lives involves an explicit trade-off between losses of allocative efficiency (which arise because prices exceed costs) and gains from dynamic efficiency. The contrast with cost-based regulation is obvious.

On the other hand, what is often not recognised in such ‘balance of competition’ assessments is the weight of the evidence pointing to the centrality of dynamic competition in promoting economic welfare at more or less all times and in more or less all circumstances, not just when in formal IPR contexts. It is in the process of ‘discovery’ – by which we mean the ubiquitous creation, processing and commercial use of information of all types, from abstract and general scientific knowledge to highly idiosyncratic business know how – that competitive markets have their most decisive advantages over bureaucratic forms of resource allocation. As John Kay has written recently⁹:

If the partial genius of market economies lies in their capacity to achieve co-ordination without a co-ordinator, the greater genius lies in their ability to innovate and adapt in an environment of uncertainty and change.

Like for all other economic activities, the prospects for gains from discovery will vary from context to context, and from time to time. There is an obvious, general relationship with uncertainty – since without uncertainty there is nothing to discover and learn. The increased uncertainty whose implications are of interest here – and which might arise from environmental drivers in energy and airlines, and from technological change in communications – can, on the basis of experience, reasonably be expected to raise the prospects for discovery, although we obviously cannot be absolutely sure of this in any particular case. That is another way of saying that the scope for dynamic efficiencies is likely to be enhanced, relative to other sources of efficiency gains.

If, therefore, the same general principles are applied to the new contexts of major and highly uncertain investment projects, regulatory policy will/should shift toward implementations that are more conducive to the achievement of dynamic efficiencies (i.e. more conducive to discovery) than current measures.

4. New regulatory risks

It might be argued that the large, uncertain projects of interest here present nothing more than an enhanced challenge to get the investment allocation incentives right. The big difference is that whereas previous concerns about over-investment (from rate-of-return regulation) or under-investment (from price-cap regulation) related to distortions at the margin of capex programmes, the big-project distortions are

⁹ “The failure of market failure”, *Prospect Magazine*, August 2007.

potentially on a much larger scale, encompassing the possibilities that a major reconfiguration of a network is either (a) not undertaken when it is economically justified, or (b) undertaken when it is not economically justified. The harm caused by ‘getting it wrong’ is much larger.

Whilst this argument is convincing on its own terms, we think it would be a mistake to suppose that this is the only thing that is different in a NGAN context. The second important characteristic of the problem under discussion is uncertainty about demand, and in the NGAN case about supply also (what is the best network architecture to adopt? What sorts of new products and services might be possible in consequence of greater bandwidth for large numbers of households?).

Regulators are required to develop policies governing how this uncertainty is best addressed; which is another way of saying that what is required are policy principles for approaching issues of ‘discovery’, ie, the dynamic process of creating, discovering and using new information. Importantly, (avoidable) risks can be expected to be introduced if policy fails to reflect this different balance of potential sources of consumer welfare gains; if, in effect, regulators do not get their priorities right.

For example, undue priority to the achievement of allocative efficiency may lead to cost-based pricing which causes one or more of the following:

- abandonment of the relevant projects, because cost-based regulation truncates returns in the event that demand turns out to be high, but cannot guarantee reasonable returns in the event that demands turns out to be low;
- undue delay to investment, because projects are held back until such time that it becomes clear, beyond peradventure, that capital costs will be recovered, even given the regulatory constraints; or
- inappropriate choice of technology, for example because opting for less capital intensive projects mitigates some of the risks that capital will not be adequately remunerated.

More generally, major regulatory risks may be associated with:

- Use of a ‘wrong’ or inappropriate economic or regulatory model to inform decisions. This is always a danger, as is manifest in much of the ‘market failure’ literature in economics, which tends to conclude that a market has ‘failed’ if it exhibits characteristics that are not captured by a particular economic model.¹⁰ The alternative hypothesis, that the model may be wrong, is rarely considered. Thus, ‘externalities’ and ‘asymmetric information’ are frequently characterised as sources of market failure whereas they are, in reality, typical features of effectively functioning markets. However, the risks of ‘model failure’ are likely to be greater in circumstances of uncertainty and change.

¹⁰ See Kay, *ibid.*

- Undue retention of regulatory discretions, which tends to exacerbate policy credibility problems, leading to a tendency toward under-investment (a classic argument which acquires increased significance because of the increased significance of investment issues).
- Increased regulatory complexity, which may itself contribute to regulatory uncertainty even in the absence of other factors.
- Failure to understand changing incentive structures and trade-offs, particularly in relation to ‘discovery’ (see above) and ‘vertical supply arrangements’.

5. The ‘standard’ regulatory model in changing contexts

As indicated earlier, by the ‘standard regulatory model’ we mean the hybrid approach that has been developed in a number of jurisdictions, including Australia and the UK, and which is broadly characterised by the following features:

- protection against monopoly pricing for consumers by the capping of prices or revenues;
- encouragement of cost efficiency by predetermining¹¹ prices or revenue allowances for a fixed period, typically a few years, but with cost-related resetting of prices/revenues at the end of each period to promote allocative efficiency; and
- promotion of competition where feasible, with removal of price regulation as competition develops.

5.1 *Distinguishing between monopoly activities and competitive services*

In many countries where liberalisation policies have been introduced it is standard practice for public policy and regulatory agencies to distinguish between the different activities undertaken by a firm, and, in particular, to focus on the regulation of those activities in which a firm is seen to enjoy a monopoly position. A common example observed across different utility sectors has been the distinction made between network activities (at least some of which tend to display monopoly characteristics) and the various activities that make use of those networks (which are more likely to be competitive). Somewhat loosely, but for ease of discussion, we will refer to those network activities that are monopolistic in nature as ‘network bottlenecks’.

The recognition that regulation should be focussed only on those activities in which a firm occupies a monopoly position – and should be gradually withdrawn from activities that become more competitive over time – is a cornerstone of good regulatory practice. However, the distinction between what constitutes a monopolistic

¹¹ For example, prices may be fixed in nominal terms, or linked to the RPI or CPI indices, or to other indices (e.g. wholesale energy prices in the case of regulated retail energy prices). The key requirement is that, so far as possible, any indexation is to variables beyond the control of the regulated company, so that financial benefits from reductions in controllable costs are kept by the company, at least until prices are reset at the end of the period.

activity and what constitutes a potentially competitive activity is not necessarily always straightforward. Nor is any classification necessarily static in nature, since the nature and boundaries of various types of network activities can change with technical progress. Nowhere is this more true than in telecoms networks.

5.2 Approaches to the regulation of monopoly activities.

Despite regulatory efforts toward unbundling based upon better identification of potentially competitive activities, it remains the case in most utility sectors that some core or network activities are unlikely to be capable of being subjected to significant competitive pressures, and therefore require some form of on-going regulatory oversight. Where this is the case, a variety of approaches to the regulation of such monopoly network activities can be observed across jurisdictions and utility sectors, reflecting factors such as: different inherited industry structures; different public policy philosophies and motivations; and a range of other sector specific considerations.

In broad terms it is useful to distinguish between two standard approaches that have been adopted in respect of the price regulation of monopoly/bottleneck activities in both the UK and Australia:

- the first – most commonly applied in the energy and transport sectors – has been to focus on setting the average price control based on the estimation of the average level of (efficiently incurred) costs associated with the provision of the monopoly activities *considered as a whole*; and
- the second – most common in telecoms – has been to focus on setting the individual prices for different services on the basis of estimates of the incremental costs associated with providing such services.

In both cases, the broad regulatory approach has been motivated by a desire to provide appropriate incentives for the regulated firm to reduce operating costs associated with the provision of the services on the network, and therefore to improve cost efficiency.

A number of complex issues tend to arise when implementing either of these approaches and, as a consequence, significant resources of both the regulated company and the regulator are generally devoted to determining what costs should be included in such price determinations, and the precise detail of how prices are to be calculated from the cost data. A notable feature of this process is the principle that prices should only reflect those costs that are efficiently incurred, and that expenditures by the regulated entity that are considered to be manifestly excessive are not reflected in prices. The intention of this formulation is to try to ensure that appropriate pressures are applied on the regulated firm for the efficient provision of network services, while maintaining the principle that prices should allow for the recovery of average costs.

5.3 Applying average price controls to network services

It is widely understood that in order for a firm subject to regulation to have the appropriate incentives to supply services, it must be the case that the overall revenue

received for supplying network services should cover the overall costs associated with the supply of those services. As noted above, in the energy and transport sectors the general approach adopted to satisfy this requirement has tended to be to ensure that the average level of prices charged for use of the network reflects the average costs associated with the provision of those services.

It follows that the primary task of regulatory authorities in these circumstances is to determine the overall (total) or average revenue requirement for the set of services subject to regulation, for a specific period of time. Under this broad-brush approach, the regulator may not play a major role in determining the prices for individual services supplied, and the firm may enjoy significant discretion in structuring its prices for different regulated services, subject to the constraint of the overall revenue cap. Such discretion might be bounded by competition law (in the event that a charging structure could be shown to have anti-competitive effects in related markets) or by requirements that the structure of charges be determined according to a published methodology that requires regulatory approval (the approach adopted for electricity transmission tariffs in Great Britain).

5.4 *Determining average prices for each network bottleneck service*

Where a monopoly network provider supplies a range of services that differ in their economic characteristics, it is sometimes the case that a regulator focuses principally on setting the average price charged for each specific service, rather than the average price level for the provision of these services as a whole. In practice, this approach – which is most commonly applied to telecommunications networks – typically involves the regulatory agency modelling the direct and indirect (*including joint and common*) costs associated with each service supplied for a specific period, which is then divided by the expected demand for that service to determine the average price.

An important consequence of this approach is that the regulated firm has less discretion to determine the structure of prices for the different services it provides, and the structure of prices for the different services provided over the network is effectively determined by the regulator. This approach to the regulation of network activities potentially conflates the regulator's role in determining the average level of revenue for the regulated firm, to address general concerns regarding excessive pricing and cost efficiency, with a role in determining the structure of prices, which can be expected to have effects on how intensively different parts of the network are used, and which is therefore more a matter of promoting allocative efficiency.

5.5 *Price structures and approaches to the allocation of common costs*

It is widely recognised in regulatory discourse that, *other things being equal* (a critically important condition), it is desirable on grounds of efficient use of economic resources that the structure of prices for different services supplied reflects the structure of costs involved in supplying each of those services. The underlying rationale is that when the prices for different services reflect the costs associated with their provision this implicitly ensures that users of the network take account of the cost implications of their use-of-network decisions.

At one level of economic theory, the most efficient price structure would entail setting the price of each service to equal the marginal cost associated with its provision. However, in network industries it is generally the case that there are substantial fixed and common costs that cannot be attributed to any one service. This requires that some method for the recovery of these fixed and common costs be employed. It is commonly argued that the least detrimental method (in terms of allocative efficiency) of deviating from marginal cost pricing to account for fixed and common costs might be to allow prices to reflect the willingness to pay of different users, an approach known as Ramsey pricing.¹² However, in practice, various cost based allocation methods have been employed by regulatory agencies for the allocation of fixed or common costs.¹³

As discussed, the average price control approach applied outside the telecoms industry has principally focussed on controlling the *overall levels* of prices and costs associated with a *set* of regulated products or services, and has left considerable discretion to the regulated firm in determining relative product or service prices. In contrast, service-by-service, long-run incremental cost (LRIC) based approaches have, in effect, not only focussed on controlling the level of costs of providing different services, but have also implicitly allowed the regulator to determine the structure of relative prices for the wider set of network services. Thus, in the UK, whereas Ofcom has tended to devote significant resources to supervising the allocation of fixed, common and joint costs among regulated services, the energy regulator Ofgem has tended to take a much more ‘hands off’ approach to this allocation exercise.

5.6 Relevance of views on trade-offs between allocative, cost and dynamic efficiency

A centrally important feature of the standard model is the predetermination of allowable prices/revenues for a designated period, so as to provide strong incentives for reductions in unit cost. Since, in the short- to medium-term, the regulated firm typically has the greatest control over operating costs/expenditures, the incentives have been particularly directed toward opex reductions. This reflects a starting position in many jurisdictions in which it was believed that the scope for opex reductions was significant or substantial.

As time has gone by, however, the position has changed: the very success achieved in eliminating opex inefficiencies has led to an endogenous shift in the trade-offs. All forms of price control tend, in one way or another and to greater and lesser extents, distort investment incentives. As opex inefficiencies have been reduced, these investment distortions have tended to acquire greater significance in policy trade-offs. To illustrate with the most usually quoted example, overly tight price controls may have relatively little effect on capex if there is cost padding in operations, but it might be much more damaging if operating expenditures are at or close to efficient levels and the only way of reducing cash outflows is to cut capital projects.

¹² There are, however, some potential qualifications to this principle when the goods or services at issue are intermediate in nature (i.e. they are inputs into downstream supply activities), and there may be more efficient alternatives if some degree of non-linearity in pricing is feasible in the relevant circumstances.

¹³ In telecommunications the most common approach is the equi-proportionate mark up, which allocates fixed and common costs to each service in proportion to that service’s incremental cost.

Even in circumstances where new investment in networks is of an incremental, non-drastic nature, regulators using the standard model have struggled with capex issues; which is unsurprising because the regulatory approach was simply not designed with the difficult task of promoting efficient capex in mind. However, the limitations of the standard framework become greater in the face of large investment projects that can be expected to have major effects on the structure of networks, and greater still when, as in the case of NGANs, the issues encompass major demand discovery problems. For just as the standard model was not developed to prioritise capex efficiency, it was likewise not designed to promote longer term dynamic efficiency.

6. The ‘standard’ model in telecommunications

6.1 *Business/activity/market boundaries in telecommunications*

Perhaps more than in other sectors, the standard model applied to the regulation of telecommunications displays a number of detailed common characteristics across jurisdictions, for example in relation to the widespread use of a common pricing methodology. However, this commonality potentially conceals some important differences in the context in which telecommunications regulation has been applied and has evolved over time. A good example of this diversity is the distinction made between those activities that constitute ‘network bottlenecks’ and those activities that are considered to be potentially competitive. In some jurisdictions – such as in many European countries and in Australia – an initial focus of market opening efforts was to develop competition in the form of resale of services, in effect treating all activities except those associated with retailing or service provision as part of a monopolistic network. In other countries, such as the UK, the possibility of resale of services was accompanied by the development and expansion of network investments by competing operators, some of whom utilised specific elements (but not the entirety) of the existing incumbent’s monopoly network, and there was a greater emphasis on network competition from the outset.

In all jurisdictions, however, the boundaries of what constitute ‘network bottlenecks’ in telecommunications have not remained static. As in other network industries these changes appear to have been, in part, prompted or facilitated by technological developments, which include not only new approaches to the use of the copper access network but also in management systems and information technology. As a consequence, a range of new resale and other wholesale products have emerged, each of which has required different interconnections with the existing copper access network. An important development here has been the emergence of access products associated with the so-called ‘unbundling’ of the local loop. These products allow entrants to effectively ‘own’ a portion of the access network, giving them an ability to install their own equipment in exchanges, and therefore to have greater control over the type and quality of final products they supply.

The key point is that the boundaries of what constitute ‘network bottlenecks’ in telecoms are not always straightforward to define/identify and, perhaps more importantly, the boundaries have been subject to continual change as a consequence

of economic and technological developments. The boundaries appear less stable, less persistent, and more temporary than in, say, energy.

The net effect of these developments is that the scope of potentially competitive activities in telecoms has generally increased. Under the theory of the standard model, it should imply a shrinking footprint for prescriptive regulation. In practice it has not quite worked out that way, at least in relation to speed of regulatory responses – which may be another indicator of the limitations of existing regulatory approaches in telecoms. For example, retail household market opening in UK telecoms took over 20 years from the commencement of liberalisation to the final deregulation of retail prices, whereas the same process in electricity, which was arguably a more difficult technical exercise in which the prospective gains from liberalisation might be expected to be lower, took 5 years.

6.2 *Network vs service competition*

Notwithstanding the difficulty in precisely defining the scope of network bottlenecks, it is more straightforward to make the broad distinction between those activities associated with network operation and those activities that are related to service provision. In the traditional access network, it has often been the case that a single network operator has provided the transmission links and exchanges which allow for end-to-end connectivity between users, while the service providers (including the network operators themselves) have used this network connectivity to provide specific communications services to end-users.

In general terms, the access transmission network in telecommunications comprises a range of different types of connections and links. However, as indicated above, the precise boundaries and characteristics of what has constituted the transmission network has changed over time, principally as a result of developments in information technology. In addition, as the services/applications offered over the networks have increased – to include services such as broadband and video on demand/TV – this has led to the different use of existing transmission networks and to the development of potentially competing transmission networks (cable, wireless, satellite etc.) on which at least some of the relevant services can be provided.¹⁴

It is still the case, however, that in a number of jurisdictions (including Australia) the extent of ‘network competition’ – in the form of entrants seeking to duplicate the traditional customer access transmission network – has given rise to a rather uneven degree of competitive influence on existing incumbent network operators, when viewed across their entire national networks.¹⁵ In these jurisdictions the main arena of competition has tended to be service provision.

Initially, the principal focus of service competition was limited to the provision of basic voice services, an activity that remains an important aspect of competition in liberalised markets. However, following the expansion of the internet, resale

¹⁴ In the UK, Germany and parts of the US, network competition for some services emerged relatively early as a result of the development of cable networks.

¹⁵ Again, the competitive influence has tended to differ substantially across different geographical areas, and across different customer groups (e.g. business customers and customers in highly dense areas have benefited most from some form of network competition).

competition now occurs across a range of other services and applications that utilise the access network. For example, retail competitors now typically offer both traditional voice services and broadband services (of varying speeds and download limits) that allow customers to access various applications such as VOiP, social networking, or video download sites.

The development and rapid expansion of applications and services that use the digital portion of the lines comprising the access network has in many ways transformed the nature of the access network. Put simply, while the physical assets comprising the access network have in many cases remained the same, the *purposes* to which these assets are being put has changed substantially.

Such technological and market developments have major implications for the scope of regulation of the services that are transmitted on the network, and correspondingly on the extent to which the network assets themselves should continue to be considered to be a natural monopoly. For example, as noted above, when these assets were used principally for the transmission of voice services, the scope for network competition was limited in most jurisdictions, although even this is changing with the rapidly increasing extent of substitution of mobile services for fixed telephony. However, given that the assets are increasingly being used for other services – such as accessing the internet or TV – the potential substitution possibilities afforded by the existence of other ‘network’ assets such as cable, wireless and satellites has expanded, suggesting that increasing possibilities for network competition may well develop over time.

6.3 *The ladder of investment and the balance of competition*

The above discussion has highlighted the difficulty in precisely defining the boundaries of the network in telecommunications, and has noted that the boundary between network competition and service competition has varied across jurisdictions (and indeed within jurisdictions) over time. This suggests that there is not a single implementation of the standard regulatory model in telecommunications, and that in many cases the form of regulation applied has evolved to reflect, and has subsequently influenced, local contexts and conditions.

Notwithstanding the apparent diversity in policy approaches, a number of jurisdictions around the world have been attracted to similar conceptual paradigms when approaching the issue of how to regulate the network elements of the telecommunications network. One such conceptual paradigm is the so-called ‘ladder of investment’ approach to the regulation of network assets. In a nutshell, this approach sees regulation as playing a role in encouraging competitors to progressively invest in network assets that are considered to be sequentially less replicable.

The most common application of this approach is to be found in regulatory efforts directed toward encouraging competitors to move away from resale services towards the uptake of the unbundled local loop, which requires them to install and operate their own DSLAM equipment in exchanges.

The ‘ladder of investment’ approach has been particularly attractive to those jurisdictions where the initial focus of competition was principally service-based, such as many Member States of the European Union and Australia. As the European Regulators’ Group (ERG) has noted, service-based competition is, in this conceptual

framework, seen as a first step (or a vehicle) for competitors toward infrastructure-based competition.

There is a large element of presumption in all of this. It is assumed that what competition can do (in the form of assumptions about replication possibilities) and how it will develop is relatively clear, whereas it might be hoped that technical progress in telecoms would encourage an approach that was more open to the notion of competition as a discovery process, from which regulators themselves will learn what might and might not be possible. As Hayek famously expressed it: competition is only really valuable because and insofar as its outcomes are unpredictable and on the whole different from those that anyone would have been able to strive for at the outset.

This raises an important point about the standard model. Since one of its component elements is the promotion of competition, the regulatory approach needs to embody a ‘view’ of competition, of what it means to say that a market is competitive. If that view is essentially static, it is liable to lead to very different regulatory implementations of the regulatory approach than if the view of competition is more dynamic. And if the implicit view or model of competition is uncertain and/or unstable over time, the expected consequence is regulatory uncertainty.

Related to this last point is the risk arising from application of rigid conceptual paradigms in regulation, which by definition require that market developments be adjusted to the dominant regulatory framework rather than the other way around (where regulation is adapted to specific market contexts). The ‘ladder of investment’ seems to us to be just such a paradigm. Competition may simply not pan out this way, unless forced to do so by regulation, in which case such forcing of competition can itself be described as anti-competitive in that it tends to restrict the potential for discovery of other ways of doing things.

6.4 Peculiarities of TSLRIC+

Despite the diversity observed in the regulatory models applied to telecoms services across countries, it remains the case that the general approach to pricing of network assets is broadly consistent, and is based on estimates of the forward looking (average) long-run incremental costs of providing a defined service, including an appropriate mark up for fixed and common costs.

Determining the forward looking, long-run costs associated with network assets is a complex and often controversial task in most utility sectors, but this is particularly the case in telecoms. The complexity can, in part, be attributed to the general approach to pricing that has been adopted, which typically involves the modelling of a hybrid network and consequently allows considerable discretion in the estimation of relevant parameters. This discretion derives from the wide range of assumptions that can be incorporated into the modelling of costs associated with specific network services, including the assumptions adopted about: network configuration and the rate of technological change; the degree of optimisation of the network, and its assumed capacity utilisation; the approaches to depreciation and the valuation of assets over time in the light of potential obsolescence and changes in asset values; and, how fixed

and common costs should be allocated across the different services/elements over the network.

In consequence, LRIC methods can, depending upon the particular combination of assumptions used (which are at the discretion of the regulator), yield quite a wide range of possible estimates of the unit cost of supplying a defined service.

In Australia, the telecommunications network services pricing has been based on a total service long run incremental cost approach (abbreviated to TSLRIC+). Without entering into a discussion of the relative merits of this specific pricing approach, there are certain general peculiarities of this approach – common to similar implementations elsewhere – which suggest that it is unlikely to be an appropriate pricing approach for the regulation of NGAN assets, and more generally may act as a strong disincentive to investment, particularly investments on a large scale.

A number of general points highlight the potential inappropriateness of this pricing approach for major new investments in network assets. One argument for TSLRIC is that it is an approach that, because it is forward looking, can be used in contexts where the actual costs associated with the development of the network are largely unknown. In the NGAN case, however, the context is rather different: the actual capex programme will be observed in real time, from start to finish.

A second argument for TSLRIC+ is that it approximately reflects the costs that would be faced by a hypothetical and reasonably efficient new entrant, and that it therefore provides appropriate signals to real world potential entrants who might be considering the alternatives of making use of an existing network to provide services (service competition) or of replicating the relevant parts of the network (network competition).

The approach, which requires repeated revaluations of network assets to reflect current replacement costs, is unfortunately wrong in theory because it ignores real option values. In circumstances in which there is some uncertainty about demand and cost conditions, and in which some of the investment is sunk, the efficient ‘entry price’ for new-build of capacity is likely to be above (efficient) TSLRIC+. By under-rewarding investment, access at TSLRIC+ prices therefore tends to create adverse incentives for new investment.¹⁶

The error is an interesting one, however, because it is another pointer to an overly static world-view, which is much closer to classroom economics textbooks than to commercial realities or higher economics. Specifically, neglect of option values is linked to neglect of changing information conditions in the market, and hence to neglect of the central dynamics of market processes.

Finally, it is widely recognised that the potential for revaluation of assets in regulatory contracts can, by and of itself, have potentially adverse impacts on the incentives for network investment. This is one of the major criticisms of TSLRIC+ as a pricing approach for network assets: the potential for investments to be continually re-valued exposes the network operator to *all* of the risk associated with changes in demand,

¹⁶ See J. Hausman, “Regulation by TSLRIC: Economic Effects on Investment and Innovation”, in J.G. Sidak, C. Engel and G. Knieps (eds.), *Competition and Regulation in Telecommunications: Examining Germany and America*, Springer, 2001.

technological change, and reductions in costs, but provides no mechanism for compensating for those risks. It follows that the more uncertain the expected changes in demand, technology or costs – and hence the higher are the neglected option values associated with access to network assets at regulated prices – the more problematic are likely to be the investments in network infrastructure under this pricing approach.

6.5 *Limitations of these approaches in the new contexts*

We have described some of the problems associated with the standard regulatory model as applied to legacy networks and, for reasons indicated, the difficulties can only be expected to become significantly greater in the context of major investment programmes that lead to a major reshaping of the relevant network, and in the face of the very major demand uncertainties that are likely to be associated with such projects. This leads to the question of whether the standard regulatory approach is likely to remain ‘fit for purpose’ for the regulation of investments in next generation network assets?

In our view, the answer to this question is likely to be in the negative for a number of reasons, suggesting that new regulatory thinking will be required, and that simply rolling-forward the existing regulatory approach or force-fitting it to changing market circumstances is likely to be detrimental to the incentives to invest.

To summarise, the reasons why we consider the standard model in telecoms to be inappropriate for dealing with next generation network asset investments include:

- The question of what activities constitute the ‘network bottlenecks’ in telecommunications is not straightforward to answer, and the boundaries of both this set of activities and of what can be classified as network activities more generally are likely to continually shift over time as a consequence of economic and technological developments;
- The diversity in services/applications now offered over the networks has increased significantly in recent years, and this has potentially introduced ‘network competition’ from other forms of transmission of these applications (cable, wireless, satellite etc). The scope of applications that can be provided using such transmission networks can be expected to increase significantly in the future, suggesting increasing competitive pressures on telecoms network providers from the aggregate of substitution possibilities that will be available.
- The classic distortions of investment incentives arising from price control can be expected to weigh more heavily when the level of capex is expected to be very high. In particular, under-investment problems associated with the retention of regulatory discretions over price determination at recurring price reviews will tend to be greater in circumstances where there is a much greater degree of uncertainty to be resolved, over time, than has typically been the case in the past. Put simply, uncertainty about how regulators will conduct themselves in the face of major market uncertainties and of changing information conditions is itself a problem for investors.

- Simple roll-forward of current conceptual paradigms such as the ‘ladder of investment’ approach to these new economic circumstances appears to us to be based upon an unjustifiably high degree of confidence as to how precisely competition will develop, which itself creates uncertainty (for example, about when and how it might be abandoned, and what will follow next) and is likely to result in unintended and adverse consequences for so long as it is retained.
- The TSLRIC+ approach is unlikely to remain robust as a pricing methodology for NGAN services. At a general level, this is because the cost conditions associated with investments in next generation assets are likely to be known, implying that there is no longer the need to estimate costs on the basis of a hypothetical network. In addition, the context in which these network assets are being deployed is very different to that which characterised the initial origins for the use of TSLRIC+, where markets were only gradually being opened to competition and where there was a strong focus on trying to establish appropriate incentives for investment in network activities, albeit that the implementation itself (in the form of TSLRIC+) was flawed.

7. Regulatory discretions and the length of pricing periods

The choice of the length of the regulatory review period is one of the most important decisions to be made across utility sectors. As is well understood, it can potentially have major impacts on the incentives of the regulated firm particularly in relation to investments in network assets. For example, the use of pure price-cap regulation with relatively short regulatory reviews can lead to an under-investment bias for regulated firms.

This is because shorter regulatory periods are, in effect, similar to a short-term supply agreement between the regulator and the regulated firm to supply services at a set of pre-determined prices. A problem with such contracts, however, is that they do not provide any assurance as to what prices might be expected at the end of the relevant period. Consequently, for those investments in assets that are productive beyond the end of one regulatory period this necessarily creates some regulatory uncertainty. In some contexts, this uncertainty can be amplified by the fact that, once the investment in the assets is made, the regulator will be in a strong bargaining position, since the relevant capital costs are likely to be non-recoverable/sunk. This potential for the regulator to act opportunistically can therefore diminish the incentives for investments in long-term projects, and is an example of the policy credibility problem.

To address these concerns, regulators need to provide longer term assurances that they will not behave opportunistically. In principle, the most straightforward way to address such concerns is to opt for longer pricing periods between regulatory reviews (potentially extending up to the entire useful life of the assets). This can provide regulatory credibility by limiting the exposure of the regulated investment to potential opportunistic intervention/behaviour by the regulator, and can provide the necessary assurance to the regulated firm that that it will be able to recover the costs associated with its investment.

Such long term regulatory reviews may not be suitable in all circumstances. In particular, where the market circumstances are such that there are serious concerns that the prices charged for the regulated service and the underlying costs may deviate significantly for an extended period – for example leading to an outcome in which prices are substantially greater or less than costs – this may result in a preference for shorter regulatory periods that allow the regulator to periodically realign prices and costs. If the period between reviews is of very limited duration, this type of approach comes to resemble cost-of-service regulation insofar as it maintains a close association between price and the underlying costs of service provision. More generally, it is often argued that the choice as to the length of the regulatory period should be related to the perceived stability of market conditions, especially cost conditions. In principle, longer review periods introduce the potential for more sustained deviations between prices and costs, and are arguably more suitable in sectors where costs are relatively stable and known.

Perceived stability is not the only factor of relevance however. A lower relative weight attached to allocative efficiency compared with dynamic efficiency in assessments of the prospects for future benefits could likewise point to longer review periods.

The underlying issue is the trade-off between the potential adverse impacts on investment and on discovery incentives of shorter pricing periods on the one hand and, on the other hand, the likelihood that prices will deviate substantially from costs for a sustained period where longer regulatory reviews are employed. As always, this trade-off needs to be considered within the context of the specific market circumstances. In the case of new and substantial investments in network assets – where the investment might not be undertaken in the absence of sufficient assurance by the regulator concerning opportunistic behaviour – there may be a strong argument for employing longer regulatory reviews, which allow for the possibility of regulatory intervention *within* review periods only in specific pre-defined circumstances and only in ways in which discretions are heavily curtailed.

8. Regulatory complexity and its implications

While there is undoubtedly great complexity associated with the technology driven developments surrounding NGNs, the shifts in the underlying uses and structure of telecoms infrastructure does not necessarily imply that the regulatory framework that governs such developments need be equally complex. High levels of complexity associated with regulation provide greater scope for potential regulatory uncertainty, which can itself act as a significant disincentive to making new investments.

The principle that regulation needs to be stable and certain and that undue complexity can create unnecessary risk for network operators, users and capital markets is well enshrined in many regulatory regimes. Indeed, one of the principal benefits associated with the introduction of RPI-X forms of regulation over rate of return approaches identified by Professors Beesley and Littlechild was (at the time) its perceived lack of complexity, which made it simpler to operate by the regulator and the company.¹⁷

¹⁷ M. Beesley and S. Littlechild “The Regulation of Privatized Monopolies in the United Kingdom”, *The Rand Journal of Economics*, Vol. 20 (3), page 456.

The acknowledgement that, to be effective, regulation must be ‘user-friendly’ and that undue complexity in regulation can have adverse effects on investment and innovation has recently attracted renewed attention in some utility sectors, most notably in the UK energy sector where Ofgem is conducting a comprehensive two-year, high-level review of the RPI-X regulatory framework. In part, this review is motivated by a recognition that the regulatory framework and processes have evolved in ways that have become exceedingly complex in some areas, and that have given rise to increasing resource costs for both the regulator and for regulated companies. As the Chief Executive of Ofgem noted when announcing the review:¹⁸

What would the framers of RPI-X say now to a process that after 20 years takes 2 years of consultation, taking more data, and arguably starting to lose the ability to identify big differences between companies?

What do the consumers say to such excellent but horribly complex concepts such as sliding scale IQI regulation (see Appendix for example!)?

More specifically, when considering the complexity associated with the regulation of investments in NGANs it is useful to contrast two approaches that appear to be emerging in different jurisdictions. The first approach might be termed a ‘principles-based’ approach, focussed on setting out a limited number of high-level principles to govern any future regulatory framework, but which does not speculate or specify in advance any particular set of regulatory arrangements. An example of this type of approach can be seen in Ofcom’s ‘Future broadband’ consultation, where the regulator notes:¹⁹

In an environment of uncertainty, it is important that regulatory policy is clear and transparent in order for industry to make informed choices on the technology, timing and reach of next generation access investments. In defining our approach to regulation of next generation access we think we should adhere to two underlying principles:

- *ensure that disproportionate regulatory policy does not inhibit efficient and timely investment; and*
- *ensure that the timing of regulatory decisions, or inaction, do not result in foreclosure of options for competition in the future.*

Elsewhere in the consultation, Ofcom notes that the issues associated with the timing and scope of NGAN deployments raise a number of complex issues for regulation to deal with, and suggests that because of this complexity it may be better to leave as many decisions as possible to the market:²⁰

It is worth highlighting that we feel these principles should be applied to ensure efficient and timely investment. This is a more complex issue for

¹⁸ A. Buchanan *Ofgem’s RPI at 20* project, presentation to SBGI, 6th March 2008, slide 19, available from www.ofgem.gov.uk.

¹⁹ Ofcom, *Future Broadband: policy approach to next generation access*, September 2007, page 27.

²⁰ Ofcom, *ibid*, page 30.

regulation to deal with than sometimes indicated. Efficient and timely investment involves investment in the right technology at the right time and in the right location, rather than a simple black and white decision between investing and not investing. At the same time, investment decisions may also need to consider overlay versus new build networks. There is a large set of potential options for investments of differing cost and value to end customers in different locations; efficient and timely investment involves choices across the changing set of options over time that maximises expected total welfare. The complexity involved is one of the reasons why we believe these investment decisions are best left to the market rather than regulatory policy or public intervention.

This approach to dealing with the complexity associated with the regulation of next generation assets in telecommunications can be contrasted with that adopted in other jurisdictions, which might be termed ‘scenario-based’. The latter approach is focussed on constructing a range of different scenarios and then developing an elaborate set of regulatory arrangements to deal with the identified potential problems.

An example of this scenario-type approach can be seen in the European Regulators Group (ERG) opinion on the regulatory principles for NGANs. In this document, rather than set out broad principles, the ERG examines the possible implications for the European regulatory framework of a number of possible NGAN scenarios. It describes the general approach adopted in deterministic terms.²¹

Implications for existing regulation and challenges to the regulatory framework are analysed, with regard to Markets 1 and 2 and more particularly with regard to Markets 11 and 12 of the Recommendation. The scenarios previously described are taken as a reference to describe possible barriers to NGA deployment. The issue of occurrence of old and potential new bottlenecks is thus raised together with considerations on appropriate wholesale products applicable to mitigate envisaged competition problems for each scenario. Some additions, with regard to changes of the Recommendation and the Review of the Framework, are suggested as a result of the analysis. Further sections of this chapter will be dedicated to focusing on procedural issues of the migration phase and highlighting possible changes in the ladder of investment in an NGA environment. Although the aforementioned developments might not lead to a fundamental change in the regulatory approach, it is necessary to analyse adjustments needed in order to preserve a level playing field for competition and to provide the right incentives for efficient investment.

Despite acknowledging the significant complexity and uncertainty regarding the potential form/scope/timing that NGAN investment may take across – and indeed *within* – Member States, the ERG document draws the definite and prescriptive conclusion that the existing regulatory framework should remain robust to such developments, implying that it will be able to accommodate these investments irrespective of the form/scope/timing they ultimately take, noting that:²²

²¹ ERG “ERG Opinion on Regulatory principles of NGA”, ERG (07) 16rev2, page 4.

²² ERG, *ibid*, page XIII.

Given the impact of scale effects on competitive conditions in different areas of a country, the national market structure may become more heterogeneous as the NGA roll-out may not happen everywhere. Summing up it can be said that in order to maintain the level of competition reached, NRAs may have to adjust the access products to fit to the NGA hierarchy, potentially followed by a lot of movements of operators, the general concept of the ladder will stay in place....Service competition based on regulated access at cost oriented prices (or retail-minus prices) can be seen as a vehicle for long term infrastructure competition. Therefore, regulators should impose remedies that enable the new entrants to reach a point of the investment ladder which makes economic sense and which tends to maximize the extent of economically efficient competing infrastructure.

Each of the two broad approaches described has pros and cons, and they contain elements that are not necessarily mutually exclusive. Principles-based approaches that recognise the limits of current knowledge and the limits of regulatory capacities can be said to start from a sound intellectual position, but principles need be made operational and brought closer to market realities so as to promote regulatory certainty rather than erode it, via discretionary interpretations of what the principles mean in future contexts. Similarly, scenario assessments can be useful in ‘stress-testing’ alternative regulatory approaches in different conditions, so as to ensure they are robust and not likely to be undermined by unfolding events. However, care needs to be taken not to assume that more is known about the future than is actually known, which tends to lead to more prescriptive, more interventionist policies than are appropriate to the circumstances.

9. Business/vertical separation issues

Policymakers are concerned about vertical integration in network sectors for two quite distinct reasons. The two are often conflated in regulatory discourse, but it is important to recognise the distinction because each carries quite different implications for public policy. The two concerns are:

- leverage of market power from activities in which a company enjoys substantial market power to activities that are, or can be, significantly more competitive; and
- the use of vertical integration to counteract the intended effects of regulatory constraints, particularly price controls.

Leverage/foreclosure

The first of these concerns is a general issue that arises across economic sectors, not just in networks, and there is a policy framework targeted on the potential problems, namely competition law. Leverage and foreclosure strategies are just one of several different ways in which a position of (horizontal) dominance or market power might be abused.

The economic analysis of such exclusionary conduct is widely understood, and it can be summarised as follows:

- There is no basis for a general presumption that vertical integration or vertical supply agreements are harmful, since the relevant economic links are between *complementary*, not substitutable, activities. Adam Smith's famous attack on conspiracies – "*People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices. ...*" – is directed at suppliers of substitutes, people in the same trade, not of complements.²³
- For complementary activities, the economic presumption is the other way round: for example, it is standard theory that merger of two companies, each with significant market power, will tend to lead to incentives to reduce prices and expand output when the products/services provided by the companies are complements. Unfortunately for public policy, whilst obvious when reasoned through, this result is not entirely intuitive.
- In 'vertical cases', therefore, and unlike in 'horizontal cases', this positive effect of agreements/integration always stands to be weighed in the balance. Its effect, other things equal, is to counteract any incentives for abuse – because restriction of the market at one level of activity tends to be bad for the profits of the integrated company at other levels of activity.
- Nevertheless, there can be circumstances in which companies with strong (horizontal) market positions can use vertical agreements or vertical integration to protect or enhance market power, and the gains from doing so may be sufficient to more than counteract the expansionary incentives arising from complementarity. It is to address these cases that competition law encompasses vertical, as well as horizontal, issues.

We know of no evidence to suggest that competition law is generally inadequate when dealing with leverage/foreclosure problems, as and when they arise. When such problems are addressed it is rarely the case that a remedy such as structural or ownership separation is adopted. There are exceptions to this, of course, but they *are* exceptions and they exhibit no discernable track record of success.

The most famous case in the UK was brewing, where the Government forced divestiture of public houses following a report on the sector by the Monopolies and Mergers Commission in 1989. Contrary to the presumption suggested by economic theory – that separation of complementary activities would lead to upward pressure on retail prices – the MMC concluded, with one dissenter, that market circumstances were such that vertical separation could be expected to reduce retail prices. The weight of subsequent evidence indicates that retail beer prices rose following the mandatory divestments, and no analyst has found evidence of the price reductions expected by the MMC.

²³ The next sentence of the *Wealth of Nations* is usually omitted, for fear of turning over a few apple-carts: "*It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty and justice.*"

Easing regulatory constraints

It is the presence of price regulation in network sectors that is the primary driver of incentives to use vertical integration to restrict competition. Put simply, by restricting the returns that can be made from the set of price-controlled activities, incentives are created to use the complementarity between price-controlled and non price-controlled activities to leverage the (otherwise unused) market power available in the former to secure higher returns from the latter.

Even when firms are gaming regulatory rules, however, seeking returns from anti-competitive effects is generally not the principal method by which regulated companies use vertical structures to their own advantage. The much more direct, and generally less risky, route to higher profits is simply to persuade regulators to accept a higher cost base for regulated activities, so as to induce higher regulated prices, for example by seeking to alter the allocation of costs between regulated and unregulated businesses.

It is towards the potential problems created by the imposition of price controls that measures such as accounting, functional, business and structural or ownership separation are chiefly addressed. In general, it is to be expected that the strength of adverse incentives, and hence the likely severity of the policy problems, will be related to the stringency of the price controls themselves. The intended severity of the price controls is, therefore, a major factor in assessing the proportionality of any accompanying business separation measures. Lighter touch price regulation goes along with lighter touch structural regulation or, put another way, the two types of measure tend to be policy complements.

Since, for reasons given above, NGANs cannot sensibly be regulated on a strict cost-of-service basis, and given that, once built, margins between access prices and (forward looking) economic costs can be expected to be high – because a high proportion of costs are fixed in nature and, once incurred, become sunk – whatever the precise approach that is eventually adopted, it can be expected to lead to a significant weakening in incentives for exploitation of vertical links compared with cost-of-service regulation applied in the circumstances.

The underlying economic logic is straightforward. With high margins (over short-run incremental cost) in the network, returns will be highly geared to demand for use of the network; the level of demand will in turn depend upon discovery/innovation in service markets; discovery will be better when there is active competition in such markets (discovery is what competition does best); and the likely incremental returns to a vertically integrated operator available from the achievement of greater market power in service markets can be expected to be modest relative to the potential loss of returns in the network arising from any restrictions in service markets.

Put even more simply, a network operator who is not heavily price constrained should want, above all, to see active, vibrant service markets, and restricting competition in such markets would not be a smart way of achieving this.

The advantages of vertical integration in the NGAN context

The advantages of business integration across complementary assets tend to be particularly pronounced in the developmental phases of economic activities, as is revealed from the early stages of the life histories of many industries/markets, from railroads (railroad ownership of complementary land assets) to video games (integration across hardware and software). Integration leverages the returns from business expansion, and it therefore tends to be expansionary rather than restrictive in its overall effects in the marketplace. In technical jargon, integration internalises what would otherwise be positive externalities which (a) tend to be larger at developmental stages of economic activity, (b) can be expected to be particularly significant for NGANs because high fixed costs create a large wedge between average and marginal costs, and (c) are not at all straightforward to fix by other means (e.g. more complex contracts) in the fact of high uncertainty about demand in the NGAN context.

Precisely because integration has real economic effects, intuition can sometimes mislead analysts into thinking that it is necessarily discriminatory and anti-competitive in its effects. The argument usually goes something like as follows: integration gives economic advantages, the playing field is therefore not level (competitors do not have access to equivalent opportunities), competition is therefore harmed. More specifically, in relation to the level playing field point, it is sometimes argued that, whatever the accounting formalities, the downstream business of an integrated firm will have access to incremental inputs at marginal cost²⁴ whereas, in the absence of sophisticated tariffs, competitors will likely face a higher price for incremental input.

Whilst it is true that integrated and non-integrated firms face different incentives, that is just the other side of the coin from the fact that integration has real economic effects. It would be folly to argue that, as a matter of principle, equivalence or non-discrimination means that the incentives of all competitors in, say, service activities should be fully aligned, irrespective of whether alignment has pro- or anti-competitive effects. Alignment of input costs is, for example, one of the textbook examples of anti-competitive cartel activity, and we have yet to see anyone seriously argue that a dominant firm necessarily behaves in an anti-competitive manner simply because, as a result of higher investment, it may have lower marginal costs than less capital intensive competitors.

Equivalence/non-discrimination is a key principle in access regulation, but it is important to distinguish it from full equalisation of business circumstances and of incentives. Equivalence/non-discrimination is appropriately promoted insofar as it serves to prevent anti-competitive outcomes, but to go further would itself be inefficient and anti-competitive in its effects, leading to more restricted markets and lower investment levels.

²⁴ However, where the extra service output of the integrated firm displaces the output of a competitor (business stealing) there will be an additional opportunity cost to the integrated firm, equal to the margin on network activities. The economic cost to the integrated firm of supplying network services to its own downstream arm is then simply equal to the charge made to others (marginal supply cost plus margin).

Moreover, the tests for illegitimate anti-competitive conduct are clear enough under competition law. For example, in assessing an allegation of an anti-competitive margin squeeze, it might be asked: can an efficient competitor turn a profit at the available access prices. There is no reference in that test, nor should there be, to the (implicit or explicit) transfer prices of an integrated company.

10. Alternative ways forward in a re-framed approach

If traditional cost-of-service regulation, including variants that include regulatory lags of up to around five years (such as RPI/CPI-X controls with cost-based reviews), is likely to be an inappropriate regulatory instrument for addressing NGAN issues, it might be asked: are there obvious, alternative policy approaches? To start to answer this question, it may be helpful first to ask two further questions:

- what are the perceived policy problems? and
- have we seen problems like these before?

Fear of monopoly outcomes

The first possible concern/problem is that, although demand is uncertain, circumstances may develop in which the NGAN becomes a major, monopoly bottleneck – exactly the kind of monopoly that is a prime candidate for price regulation.

At this point it may be well to recall again that the issue with which economists since Adam Smith have been concerned is *persistent* or *permanent* monopoly, not the temporary monopolies that play key roles in providing incentives for dynamic efficiency. Thus, although it is not known *ex ante* whether a serious monopoly problem will or will not eventuate, if the risk of such an outcome is considered to be sufficiently damaging as to call for mitigating measures, the obvious way of addressing the issue is to follow policies that target the *persistence* problem.

The issue is familiar in patent policy. Temporary monopoly is encouraged to provide incentives for discovery/dynamic efficiency, but the intellectual property rights are time limited. In the case of NGANs, a broadly analogous policy would be a promise not to regulate the average price level of access services for a minimum, defined period of time (e.g. for say, 10-20 years). Subsequently, it would be open to a regulator to introduce some or other variant of cost-of-service regulation in the event that, at the relevant future time, it appeared to be warranted by the economic circumstances (i.e. if, on assessment, it were judged that there were bottlenecks that could be expected to persist). Whereas in the patent analogy the end of the protected period signals the potential entry of competitors supplying similar products, in the suggested approach to NGANs the end of the period signals an assessment which, depending upon the facts at the time, might or might not signal a reversion of average prices to costs brought about by re-regulation.

Note that this approach, based upon removal of constraints on average prices for an extended, and possibly indefinite, period is likely to be significantly different in its

economic effects from the ‘access holiday’ approaches that have been advocated or adopted in some jurisdictions.²⁵ In our view, the ‘access holiday’ route may turn out to be substantially more restrictive because it creates a market structure with less competition in the ‘discovery’ of uses of network capacity. Open access arrangements have the great merit of encouraging wider participation in the discovery process.

Addressing potentially harmful discrimination: statics

One immediate potential policy issue is that, if charging constraints were lifted for NGANs, in those locations where the NGAN replaces an existing, copper-based access network and where there is no closely substitutable access network providing existing products/services (e.g. voice), the network operator would potentially be able to derive extra revenues by increasing charges for these pre-existing services, particularly since, given the nature of the demand, some of them might be able to sustain particularly high price-cost margins. That is, in terms of notation used above, retail products/services falling in the set S_0 might be vulnerable to monopolistic pricing in certain geographic areas.

It is then easy to see that, in addition to the consumer protection issues that would arise, the prospect of deregulation could cause distortions in investment decisions: the prospective returns from new-build would be augmented by returns resulting from the linkage of deregulation to investment (because higher investment could lead to higher prices for *existing* products/service), whereas efficient incentives require the investor to balance the *incremental* gains (over existing provision) against incremental costs. Thus, potential extra financial returns from, say, voice traffic should not be included in the assessments, except insofar as they derive from lower costs or from improvements in the quality of the final service to consumers that might justify higher prices.

Any perceived risks of consumer harm can be mitigated by targeted protection of consumers in relation to those existing products/services that are judged to be potentially vulnerable to re-pricing. In competitive markets, the incentive mechanism that tends to provide most protection is *reputation*, but where reputational effects are unlikely to be sufficiently strong, long-term contractual agreements might be adopted.

Recent suggestions from the UK regulator Ofcom, under the label of ‘anchor pricing’, have followed this line of thinking. The idea is to focus regulatory price influence on a sub-set of NGAN ‘anchor’ products that support key retail products/services currently provided over the copper network, whilst leaving prices of other NGAN products, including those yet to be developed, unconstrained. One way of looking at this is to think of anchor pricing as being akin to retaining a ‘virtual’ legacy network in competition with the new NGAN, and this may help clarify when and where this approach might be judged appropriate. For example, in geographic areas where actual, competing access networks exist and are capable of supporting the relevant retail products and services, anchor pricing may be redundant.

²⁵ Access holidays amount to time-limited rights to refuse to supply use-of-network services to others. Our focus here is on an approach based on non-discriminatory access at deregulated levels of price.

For reasons just given, the general approach that is being explored by Ofcom appears to us to be well-targeted, although, as always in regulation, the precise details of implementation will matter a great deal. If there was constant re-setting of anchor prices and/or services, this could undermine the improvements in investment incentives that the approach would otherwise achieve, since the investor might perceive an obvious risk of regulatory opportunism. Thus it is easy to see that, in the future, the toughness of any re-setting of anchor product prices might tend to be positively correlated with the returns then being earned on other, deregulated products, thus effectively reducing those returns. The effect of such an influence/linkage might reasonably be described as a ‘partial regulation’ of new and innovative products, and it would have obvious, adverse consequences for discovery/dynamic efficiency incentives.

One ‘market solution’ to opportunism is long-term contracting, and one approach to the pricing of anchor products could be based around a long-term agreement or undertaking. In Australia, the SAU mechanism provides a legal structure in which this could be done. In the UK, it would require a little more innovation in institutional arrangements.

Addressing potentially harmful discrimination: dynamics

The second aspect of discrimination that is likely to be important is linked to a potential perception on the part of network users that, should they innovate successfully, part of the returns to innovation might be vulnerable to capture by a NGAN operator, who, if unconstrained, could potentially re-price products on a ‘what the market can bear’ basis. This is another form of potential opportunism, but on the part of the network operator rather than the regulator. The concern is that the risk of such opportunism might dampen incentives for investment and discovery in retail service markets.

One manifestation of this general set of issues is the risk of hold-up of sunk investments made by network users and by end customers. This problem has been recently considered by Darryl Biggar in a paper that serves as a reminder of the importance of taking account of investments that are complementary to those of the network operator.²⁶

Biggar’s account of regulatory preferences for stability in pricing and for non-discriminatory pricing (rather than Ramsey pricing) is highly consistent with regulatory practice in UK energy regulation, which has maintained the firm distinction, inherited from US regulation, between the different exercises of determining the average *level* of charges and the *structure* of charges for use of networks. Whilst the latter tariff structure decisions have been left to the discretion of the network operators, subject to regulatory approval of methodologies, regulatory influence has tended to favour methodologies that provide relatively stable charging patterns. And the reason for that is clear: it is an approximation to some of the effects of a longer-term contractual structure in an economic environment that, for other

²⁶ D. Biggar, *Is protecting sunk investments by consumers a key rationale for natural monopoly regulation?*, January 2008, available at: http://works.bepress.com/cgi/viewcontent.cgi?article=1000&context=darryl_biggar.

policy purposes, calls for networks that can easily be accessed by new entrants. A potential network user about to construct a major new power station, an interconnector, or an LNG terminal will obviously be rather keen to have some certainty as to the future course of use-of-network charges, or at least as to the methods by which such charges will be calculated.

The telecoms context is rather more complex than energy, for both substantive and for much less convincing reasons. The substantive reasons are related to the rather greater uncertainties about future products and services in telecoms, and about their interactions with existing products and services, as discussed above. Given the uncertainties and the associated prospects for discovery, there is nothing obviously problematic about the existence of strong linkages between network returns and successful innovation in the products and services that make use of the networks as inputs. These are, after all, complementary economic activities.

The precise form of the linkage will determine how demand risks are shared, and hence will affect the relative strengths of incentives for investment and innovation at each of the network and service levels. What is required is something that mimics an efficient long-term contract, since what is important is that the relevant framework for charging be determined *ex ante*, at least to an extent sufficient to impede harmful opportunism *ex post*.

The non-substantive complexities arise from the peculiar way in which LRIC pricing has evolved in telecoms, which might be described in Conan Doyle language as ‘the strange case of the increments that are averages.’ Perhaps as a response to the rather fluid boundaries between monopolistic and potentially competitive activities in telecoms – a fluidity illustrated, for example, in the impact of NGAN technologies on the economics of local loop unbundling – the sharp, traditional, regulatory distinction between price level and price structure issues has tended to get lost. As argued above, different issues tend to be conflated in the adoption of one or other of the several potential versions of TSLRIC+ pricing. TSLRIC+ is a slave compelled to serve more than one master, and consequently it serves none of them very well.

These regulator-generated complexities, and the problems to which they tend to give rise, can be eliminated by policy makers by the simple expedient of abandoning TSLRIC+ as a pricing methodology for NGANs. The more substantive issues are, however, unavoidable, and in our view they point to the centrality of risk sharing and non-discrimination issues in developing frameworks for investment in, and operation of, NGANs.

Longer term contracting with more explicit risk sharing, and more explicit links between risk and reward, is the favoured way forward in UK energy networks. Thus, faced with the uncertainties surrounding new build of transmission links to serve renewable energy sources, Ofgem is pressing for greater *ex ante* network user commitments in return for traditional regulatory guarantees concerning price caps for use of system charges.²⁷ At the same time, the regulator has indicated willingness to increase traditional price caps to levels above cost-based levels in circumstances

²⁷ Part of the uncertainty arises from the difficulties in co-ordinating investment in vertically separated, complementary activities (new wind farms and new transmission capacity).

where network operators build capacity ahead of securing user commitments. The *quid pro quo* here would be lower returns in the event that the anticipated demand for use of network capacity did not eventuate.

In a sense, the regulator is now trying to re-introduce some of the incentives that are present under vertical integration of transmission and demand, but via contractual forms of arrangements, since electricity transmission (though not local electricity distribution) was fully separated from generation at the time of privatisation. The developments are at an early stage, and the problems are by no means easy to resolve, so it will be some time before the effectiveness of any reforms can be reliably assessed.

In telecoms, some of the uncertainties that arise in energy are mitigated by vertical integration, but we would expect that longer-term contractual approaches could also play important roles, both in relation to interactions between the regulatory authorities and the network operator and between network users and the network operators. In this context it may be worth remembering that alternative approaches to the underlying economic problems are not mutually exclusive, and that in many economic sectors the relationships between players at different stages of vertical supply chains are characterised by a spectrum of co-existing trading arrangements, from vertical integration, to long-term contracts, to medium- and short-term contracts, to informal but ongoing understandings, to spot trading. Further, such diversity of relationships is a source of strength in economic arrangements, particularly in conditions of high uncertainty. When it is unclear what events might occur, and unclear what probabilities might be attached to different outcomes, putting all eggs in one basket is a particularly risky strategy for a nation's policy makers.

As to what might replace the standard model as a regulatory approach, it is clear that national authorities are themselves in a midst of a discovery process of their own. On the basis of earlier points, however, we suggest, in a tentative way, that new approaches might be characterised by structured agreements/undertakings of relatively long duration (compared with the past durations of pricing periods) containing the following type of commitments by network operators:

- to access arrangements that embody the appropriate principles of equivalence/non-discrimination (see bottom of page 32), among other things to provide greater stability in pricing, constrain *ex post* opportunism and facilitate enforcement of general competition law;
- not to price above certain levels for a defined set of access products/services (anchor pricing) in those locations where relevant retail services cannot be supported by competing access networks; and
- to negotiate, on a commercial basis, risk sharing arrangements of various kinds, such as longer-term use-of-network contracts that take account of any *ex ante* commitments a user or group of users might be willing to make.

On the regulatory side, the main requirement is a familiar one: to stick with commitments made, even in the face of information that indicates that the regulator would not have made the agreement if he/she had known earlier what becomes known

later.²⁸ If discovery is perceived to be threatening to a regulatory settlement, then discovery will be chilled.

The precise balance of any agreements/undertakings will depend upon the specifics of the relevant factual context. However, the general principle at work in all cases can be expected to be the same. Just as RPI/CPI-X shifted the ground away from strict cost-of-service regulation because of developing perceptions that additional cost efficiency benefits were considered ‘worth’ some sacrifice of allocative efficiency, the new challenges call for a rebalancing of priorities away from the realisation of static efficiency gains and toward establishing the kind of incentive structures that will be conducive to discovery and to the realisation of dynamic efficiency gains, upon which the prospects for future improvements in consumer welfare for the most part depend.

²⁸ Subject to the obvious exception when a regulator has been misled.

About the Authors

George Yarrow

george.yarrow@rpieurope.org

George Yarrow is Director of the Regulatory Policy Institute, Oxford (www.rpieurope.org); Emeritus Fellow of Hertford College, Oxford University; Visiting Professor at the Newcastle Business School; a Board Member of the Gas and Electricity Markets Authority (GEMA, www.ofgem.gov.uk), the GB energy regulator; economic adviser to the Civil Aviation Authority; a member of the National Audit Office's academic panel on regulatory impact assessment; and a director of the companies DKY Ltd and Davison Yarrow Ltd. He also recently served on the Republic of Ireland's Aviation Appeals Panel, which handles appeals against regulatory determinations of airport charges.

After graduating from Cambridge University, he held appointments at the Universities of Warwick and Newcastle before moving to Oxford, where he spent most of his academic career. During sabbatical periods he was a visitor to Harvard University, the University of California at San Diego, and the University of Urbino.

His principal work has been on the economics of privatization, regulation and competition. Among his books are: *Privatization: An Economic Analysis* (MIT Press) with Sir John Vickers, which has also been published in Spanish and Chinese translations, and *Privatization* (Routledge), a four volume anthology of papers with Piotr Jasinski. He has written numerous academic papers and policy studies on issues of competition and regulation.

Recent public lectures have included: The Enterprise Act (London), Electricity Market Reform (Paris), the Changing Dynamics of Europe's Liberalizing Energy Markets (Amsterdam), Modernization of EU Competition Law (London, to UK High Court and Appeal Court judges), Economic Assessment in Competition Law Cases (Berlin, to the Association of European Competition Law Judges), EU Energy Policy (for the annual ACCC conference, Australia), and Energy Policy: A Time to Stop Pretending? (in the Beesley Lecture series in London). In the second half of 2008 he will be speaking at the ACCC conference on general issues in regulation and on environmental aspects of energy regulation, and giving a Beesley Lecture on 'Discovering the value of water').

Recent policy studies have included work on: Regulatory reform and the promotion of competition in communications (DTI/DCMS, London), Economic regulation of air traffic management (DGTREN, Brussels), State aid and nuclear power (DTI, London), Assessing the burden of regulation on business (Cabinet Office, London), EU Member States Regulatory Impact Assessments (for EU Directors of Better Regulation), Regulatory instruments in farming and the agri-environment (Defra), Reverse eAuctions in NHS procurement (ABHI), , and most recently The effects of maintaining price controls in liberalising markets (Australian Energy Market Commission), and The prospects for competition in the water services sector (Water UK).

Professor Yarrow has considerable experience of policy advisory work, both at international agency (OECD, World Bank, UNDP, etc.) and national government level. Areas of regulation covered by this work include: banking and financial

services, communications, competition policy, energy, environment, health services and pharmaceuticals, and transport (air, rail and road).

In addition, he has a longstanding interest in competition law and policy, including having at various times acted as an advisor to enforcement authorities and to companies, at both national and at EU levels. He has given evidence in competition cases to the Competition Commission, the Office of Fair Trading, the High Court (e.g. in *Crehan* and *Arkin*) and the Competition Appeals Tribunal in the UK, and to the European Commission and the European Court of First Instance. His current cases include DRAM interface technology standard setting, the determination of mobile telephone call termination rates, and conditional discounting in the supply of microprocessors.

On the specifics of regulatory policy, Professor Yarrow has experience of virtually every major aspect of policy development over the past twenty years. He was economic adviser to the National Grid for the initial design of the transmission use-of-system charges for the high voltage electricity grid, and later to British Gas for the development of similar entry/exit arrangements for gas pipeline capacity. Later, first as economic adviser to Ofgem and now as a board member of GEMA, and in addition to working on all aspects of price controls (gas transmission, gas distribution, electricity transmission, electricity distribution, system operator incentives, gas storage, meters) he has been involved in the full range of regulatory reforms introduced in the UK from the mid-1990s on, including: retail market opening, retail market deregulation, gas storage deregulation, the new electricity trading arrangements, the new gas trading arrangements, the establishment of the first energy exchanges, the integration of the Scottish and England & Wales electric systems; and the enforcement of the Competition Act in the energy sector. In telecoms he has been a member of expert panels set up to assist UK ministers in the development of the new Communications Act and to assist EU Commissioners in the development of policy responses to technological convergence in communications. In air transport, he has been an advisor to the CAA for successive price reviews of Air Traffic Control Services and of London and Manchester airport charges.

Chris Decker

chris.decker@rpieurope.org

Chris Decker is Research Director at the Regulatory Policy Institute, Oxford and an Associate Research Fellow at the Centre for Socio-Legal Studies, Oxford University. Prior to returning to the Institute in October 2007, Chris was a Principal Economic Advisor at the Australian Competition and Consumer Commission. Before that he was a research associate at the Regulatory Policy Institute, and also worked as an economic consultant. He currently acts as an independent economic advisor to Ofgem and is a member of the panel of experts for the Commission for Energy Regulation (Ireland).

Chris's work is focused on economic regulation, competition economics and public policy. He has been involved in a range of research and consulting projects for both the public and private sectors, including for: the OECD; the European Commission (DG Transport & Energy); the Australian Competition and Consumer Commission; the Australian Energy Markets Commission; ENARGAS (Argentina); the South African Competition Tribunal; and in the UK, the Competition Commission,

Department of Trade and Industry, Cabinet Office and Office of Gas and Electricity Markets. In addition, he has worked on projects for private organisations such as: British Telecom, British Gas, J-Power (Japan), and EPDC (Japan).

Recent research projects he has been involved with include: a study examining the prospects for competition in the UK water industry; a review of International approaches to transmission access for renewable energy; an assessment of the competitive and economic impacts of the introduction of the Single European Payments Area; a study for the European Commission on the development of implementation rules of economic regulation within the Single European Sky initiative; and two studies examining the impact of regulation on investment and innovation for the Department of Business, Enterprise and Regulatory Reform.

The major focus of his academic work has been on the issues relating to the application of economic techniques in regulatory processes and in competition law enforcement. His book *Economics and the Enforcement of European Competition Law* (Edward Elgar) is forthcoming in February 2009. Chris has a first class honours degree in economics from the University of Melbourne (Australia) and a PhD from the University of Oxford.