Utility regulation, the RAB and the cost of capital[1]

Dieter Helm
University of Oxford
May 6th 2009

1.   **Introduction**

There are a number of reasons why the great British privatisation experiment may have run its course. At the heart of privatisation are two fundamental ideas: that equity is a necessary condition for maximising incentives; and that private balance sheets play an important role in financing investment. Both of these have now been seriously undermined: there has been a gradual but remorseless flight of equity since the mid-1990s, accelerating after 2000; and financial engineering has exhausted balance sheets and broken the link between physical investment and borrowing.

So far the consequences have been limited. For the most part, it has been a very slow-motion crash—though, in the case of Railtrack, there was literally a train crash and a collapse of equity. Welsh Water exited equity too, though attempts to follow suit by Kelda and Anglian into mutuals were not successful. But the process has been remorseless, and now all the companies have geared up and, in consequence, equity has been squeezed. Attempts to inject equity have been notable only in response to distress. These include United Utilities and its defence of its dividend (and its ill-considered acquisitions in the 1990s), BT, and now Anglian and Southern Water to meet covenants.

The process is far from complete. On current trends, it is quite possible that equity will be eroded further, and more mutuals and de facto nationalisations may follow. In the twentieth century, across much of Europe, public ownership or debt-based rate of return regulation were the norm, as private equity existed the utilities—often after investment failures and typically by compulsion. Though privatisation has come to be taken for granted, there is nothing inevitable about this form of utility ownership.

---

[1] Comments to dieter@dhelm.co.uk. Further papers available on www.dieterhelm.co.uk.

With such an outcome, the utilities would gradually fall back on the incentive-poor public sector and rate of return contexts, and in due course the capital budgets would increasingly come out of national and local public budgets, or, in the case of mutuals, the political influences which would come through the appointed 'members' and trustees.

But just as private ownership is not inevitable, nor is the extinction of equity and incentives. The process is in large measure the outcome of a badly designed regulatory framework. British utility regulation does not reward equity properly. RPI – X is designed to promote high-powered incentives, and that requires equity. But RPI – X does not set the financial framework in harmony with these incentives. On the contrary, it offers a weighted average cost of capital (WACC), which definitionally averages between debt and equity. In a normal commercial business this would not matter much, but in the utilities its effects are altogether more pernicious. The reason is the special status and role of the regulated asset base (RAB). The WACC does create incentives, but these are not the ones RPI – X is meant to deliver. Rather, it is an open invitation to financial engineering, to replace equity with debt, and once complete, for equity to complete its exit.

This paper is an attempt to explain what has been going on in the utilities over the last two decades, and in particular to explain the importance of the RAB, and the appropriate incentive regime to combine the protection of past investments through the RAB, while at the same time maximising the incentives to operate and invest in the businesses in an efficient way, and in the process open up the maximum scope for competition.

The paper is structured as follows. Section two deals with the RAB and its central role in utility regulation. This is sharply distinguished in section three from the day-to-day job of operating the networks and investment, on which incentives are concentrated. These two sections provide the conceptual dissection of utilities into their economic parts. Section four looks at the role of regulation, and in particular in allocating the equity risks (which are never extinguished) between shareholders, customers and taxpayers, and how this allocation is reflected in rate of return and RPI – X regulation. Sections five and six then respectively turn to the problems associated with using a WACC approach and to indexation of the costs of debt. Section seven considers the role of government in setting the overall investment framework—the direction of travel. Finally, section eight sets out a series of reforms which would rescue the equity role and revitalise both the balance sheets and the incentives.

2. The centrality of the RAB

In much of the literature on utility regulation, the core problem identified is monopoly, and in particular natural monopoly. This, in turn, leads to a static analysis of the appropriate price and output: focusing on how to limit excessive pricing and insufficient output. But market failures are multiple: for infrastructure complementarity and externalities matter too. Fixing one market failure—like monopoly—on its own may actually make the efficiency of the outcome worse. For example, if price is above marginal costs for monopolies, but below marginal costs for pollution, lowering the price makes the externality worse. Arguably this is exactly what happened in electricity in the 1990s. Multiple market failures call for a coordination of policy instruments.

2.1 The time-inconsistency problem

The control of monopoly is an important problem, and in a period of low investment and general excess supply, it may even be the dominant one. But it is not the only problem, and when investment is required, consideration moves from asset-sweating to creating new assets and infrastructures. Then the dominant problem is how to ensure that there are appropriate incentives to invest and that investors are not exposed to *ex post* exploitation by politicians and regulators. This is the *time-inconsistency problem* familiar in a number of contexts, most notably in monetary policy.[2] It arises in the utilities because they are long-lived, capital-intensive industries, with considerable sunk costs, and because it is in the nature of networks that there is a wide gap between marginal and average costs. The temptation for politicians and regulators is to promise *ex ante* that investors who sink capital will be able to recover their investment and the cost of capital. In other words, they will be able to charge average not marginal costs. However, the time inconsistency arises when the investment is complete and the assets are in place. Now the politicians and regulators can force prices to marginal not average costs, and the services continue to be provided. They renege on their *ex ante* bargain with investors.

---

[2] See Helm, Hepburn and Mash (2005) for an application to environmental policy.

2.2 Previous attempts to solve the time-inconsistency problem—monopoly, nationalisation and rate of return regulation

There is nothing new about the time-inconsistency problem: it has been a feature of utilities for centuries. Not surprisingly, therefore, a number of solutions have been tried out. In the case of the railways in the 1840s, concessions were offered to developers who were then free to charge as they chose. In other words, monopoly and market power 'solved' the problem. In the ensuing speculative bubble, there was a rush to build, with the result that overcapacity meant that competition drove prices below average costs, and many investors lost their shirts. A similar pattern followed in the 1980s and 1990s with telecoms, mobiles and the new Internet services: a free-for-all plus a speculative bubble (the dotcom boom) cost many investors dearly.[3] This process of 'creative destruction', as Schumpeter famously termed the process, relied on permitting monopoly not only as an incentive to invest, but also to stimulate innovation and entry.

A second solution—perhaps the most common in the twentieth century—was nationalisation. Since governments could not credibly commit *ex ante* to investors, investment in the private sector would be deficient. The water and sewerage systems were not built, electricity networks lagged behind, and roads and airports were inadequately provided. So governments internalised the time-consistency problem: in effect, taxpayers provided the fall-back to cover the capital costs where prices were kept below average costs. Indeed, it was argued that this was the proper role of taxpayers—efficient resource allocation mandated marginal cost pricing.[4]

A third solution was rate of return regulation, which was widely adopted in the US, Japan and in many European countries where the private sector remained. Costs—*all* costs—are passed through to final customers. Provided that the guarantee is credible—and in the US the legal protections were very considerable—investors can sink capital, secure in the knowledge that there will be no expropriation through *ex post* marginal cost pricing. Although in the US a 'used-and-useful' test was sometimes applied together with efficiency audits, the credibility was such that investors had incentives not only to invest, but even to over-invest: hence the criticisms that rate of return regulation leads to gold-plating.[5] The drawback was that rate of

---

[3] The M6 Toll Road has a similar freedom to price, but here there is a substitution for the zero-priced M6.

[4] See HMT (1967)

[5] See Averch and Johnson (1962).

return regulation covered *all* costs, and hence OPEX and CAPEX costs were passed through too, with the result that US utilities tended to be inefficient in aggregate.

2.3 The RAB—a targeted solution

As an afterthought, rather than an act of conscious design, RPI – X regulation developed a fourth solution, the concept of a regulated asset base (RAB). It was an afterthought because it only became apparent that assets needed to be given a value after privatisation when the time came to re-set the price caps, and hence evolved as part of the evolution of RPI – X. At the periodic review, regulators needed to provide not only sufficient revenues to cover the OPEX and CAPEX—the day-to-day business of running the utility—but also a return on the amount invested in the business. The RAB is the assets within the regulated entity necessary to carry out the functions of the business. The regulated capital value (RCV) is the number given by the regulators to represent these assets.

Because the RAB was an afterthought rather than considered at the time of privatisation, there was ambiguity as to what it represented exactly and how the RCV should be calculated. That ambiguity remains today and, as we shall see, comes with a cost of capital premium.

The problems which arose at the first periodic reviews were threefold: first, the price levels at privatisation did not remunerate the current-cost value of the assets; second, the value of the initial investment was unclear as between the issue price, the initial premium and the evolution of market value in the first period; and, third, there were no *ex ante* rules as to how the CAPEX in the period was to be incorporated in updating the RAB. Each has an important bearing on the current regulatory regime—we are prisoners of past mistakes.

The most formidable problem is the result of the way in which the time-inconsistency problem was addressed in the nationalised industries. Prices for utility services were typically below the average costs, and hence did not fully remunerate the value of past investments. Part of this was disguised by using historical cost accounting, thereby facilitating through inflation a gradual reduction in the real asset prices. In effect, taxpayers were taking a hit, to the benefit of customers. This probably did not matter much in one respect—taxpayers and consumers were largely the same people. What it did mean, however, was that after privatisation utility services were provided at below the full costs, and as the assets were replaced, the current cost of the assets would be increasingly manifest in prices, and prices were therefore set to rise. It also meant that the prices were distorted, being neither at marginal costs nor at full resource costs. Moreover, because the RABs were all valued

differently, there were distortions between the networks. This mattered most in energy between gas and electricity, and as a result power stations were built in the wrong places. Network prices were both relatively and absolutely wrong.[6] In transport, one of the reasons why it is typically cheaper to fly between British cities than go by train is partly explained by the inconsistency in the approaches to valuing airport and railway assets—and of course the externalities too.

Thus, the current-cost value of the regulated assets was higher than the privatisation value, determined by prices held below full costs. The latter reflected the politics: raising prices at privatisation would have increased the proceeds from the sale, but would have been politically very unpopular. Or, put another way, the assets were worth the capitalised value of the revenues minus the costs in the price cap, after allowing for OPEX and CAPEX. Because the prices were kept low, the balance sheets were smaller than would be justified by the current-cost asset values. This would turn out to be painful later, as the balance sheets were exhausted.

The second problem exercised regulators most: what initial value to set. There were several options: the offer price, the first-day premium, the 100-day average, or a generalised uplift. It will be apparent that there is no right answer: it is essentially arbitrary. Attempts by regulators to justify their chosen base (they were all different) were unconvincing and in some cases (such as the 1994 electricity distribution review) sufficiently implausible to merit a re-opening of the review.[7]

The third problem was how to incorporate the CAPEX in the period which was ending. The utilities argued that this should be the forecast CAPEX, so that the value was increased by what it was agreed would be an efficient spend, as opposed to what was actually spent. The gap was material: in the electricity distribution case, the companies spent about half the predicted amount. The regulators argued that the efficiency incentives in respect of CAPEX applied *within* periods but not between them, and hence the lower of the actual CAPEX costs and the efficient CAPEX costs were transferred to the RAB. Again this had consequences: if the incentives applied within the period only, the most profit-maximising CAPEX projects

---

[6] See Helm (2004), pp. 279–80.

[7] In water, it was the first 100 days; in electricity distribution, at the first attempt, it was 50%.

were short-term ones. The capital programme was therefore systematically biased to respond to the incentive mechanism.

These problems are all serious, and none has been entirely satisfactorily addressed. Prices remain distorted, and importantly still are relatively distorted. At periodic reviews, regulators do not typically concern themselves with asking whether the price level is correct—they take it as given, and focus on the changes from it. The original mark-up on the privatisation sale price was messily resolved, but at least it is in the past. The incorporation of new CAPEX remains a contentious area.

Notwithstanding these difficulties, the RAB remains a vehicle for committing to sunk investment, and therefore it provides a long-term contract between customers and investors. It is what keeps the privatised industries from collapse into nationalised entities. Without the RAB, investment would be value-destroying to shareholders.

2.4 Financing functions

The guarantee which lies behind the RAB—and upon which investors rely—is the duty to ensure that functions can be financed. This is a duty on independent regulators and is backed up by statute. If the regulator fails in this respect, companies can have recourse to the courts, typically first via the Competition Commission—and hence, provided that managers pursue shareholders' interests, investors have protection.

It matters greatly what this financing functions duty means, and how it relates to the RAB. Unfortunately this is not entirely clear. It has never been tested in the courts, and regulators have given different interpretations—indeed, the precise wording differs between the industries. Quite early on, regulators interpreted the duty to finance functions as the duty to *efficiently* finance the *efficiently* conducted functions. This was a substantive change, and open to considerable interpretation. What does it mean to finance functions 'efficiently'? Is this an *ex ante* concept? What are the counterfactuals? What happens to embedded debt? Regulators gradually filled in this detail, but rarely with clarity. For example, notional gearing was assumed, despite the differences between individual company circumstances, not least the variance in their real CAPEX spending. But even here the notional numbers were up for review at each periodic review.

And then there are the *efficient* functions. What exactly is 'efficient CAPEX'? More immediately relevant, how can the RAB be efficiently carried out since it is an *ex post*

number? And, more worrying, if OPEX and CAPEX functions are not carried out efficiently, does this mean that the financing of the RAB is at stake?

These are not mere academic questions. The risk to the RAB is entirely regulatory and political: managers have to do nothing to gain a reward on this number, and they cannot by their actions change the number. They cannot alter the regulatory and political risk. Finance theory tells us that risk is best borne by those able to manage it, and conversely transferring risk to those who cannot raises the cost of capital. Given that the RAB is such an important part of the total financing of utilities, it follows that failure to provide clarity on how the duty to finance functions is defined—indeed, failing to make it crystal clear that it includes a commitment to honour the RAB—is a deadweight welfare cost on consumers.

It is therefore imperative that regulators define this duty with clarity, even to the extent of engineering a judicial review, so that regulatory and political risk can be minimised. Then the RAB can be financed by debt, and with a (very) low premium to (index-linked) government bonds. The result is a material change to the cost of capital, with large implications for customers' bills. How this can be effected is the subject of sections five and six. But before that, there is the other component—the operating businesses.

3.   **The operating businesses: OPEX, CAPEX and competition**

In contrast to the passive nature of the RAB, the day-to-day business of delivering the utilities services requires active management, and is therefore naturally equity-driven. The OPEX is comparable to many asset management tasks in the private sector, and these types of business tend to be asset-light. It is therefore hard to see what role, if any, debt might play.

It is also on the OPEX that the high-powered incentives of RPI – X bear most directly, as witnessed through comparative-efficiency exercises. Since managers can do very little about the RAB, and since CAPEX efficiency is much harder to measure, the cost-minimisation axe falls here.

CAPEX requires project finance. The nature of such finance depends on the exit once the assets are created. In the private sector, project construction is frequently put out to tender, the contractor raises funds for the project duration, and then pays down from receipts from the purchaser. Project finance is typically a mix of equity (including working capital) and debt. The time period is typically short-term, tied to the project and discharged upon completion. In the utility case, this is institutionalised through the RAB; once completed, CAPEX projects

are sold into the RAB, providing the exit. It follows that the financial profile of the OPEX and CAPEX is quite different from that of the RAB.

There remains one further function of utilities which lies outside the RAB: the coordination of the networks to ensure that the functions are delivered. This coordination function is tied to the licence: companies must ensure that they deliver these functions, and the licence is conditional on delivering them.

It has been argued that this location of responsibility for coordinating delivery ties the RAB to the licence since only if companies perform can they keep their licence, and only if they keep the licence do they have a protection that the functions are financed, and hence the RAB is remunerated. Put simply, fail to deliver the functions and the RAB is at risk.

It may well be that investors perceive this risk and in consequence price in an equity risk in respect of the RAB. However, the perception may lack substance, for it depends on what happens if the company fails, and, if this happens, on the role of the special administrator. In principle, a company which fails to comply with its licence functions has its licence withdrawn and then sold on to another that can discharge these functions. The question is whether the value of the RAB remains constant, and hence whether the security of the debt holders is preserved. This is both an empirical question and a matter of regulatory design.

Empirically, in the case of Welsh Water, the company was sold on without a special administrator for a value below that of the RAB. In the Railtrack case, the company went into administration (a form peculiar to rail) and then transformed into Network Rail. The bondholders were protected; the equity owners were not. If Welsh Water had gone into administration, the regulator would have had to ensure that the (efficient) functions would be discharged by a new owner. Therefore the RAB would have had to have been honoured. Indeed, it is possible that a premium to the RAB would have been paid by an acquirer. In the case of Welsh Water, it could be argued that the prior owners took a loss in respect of the operational business and this offset the apparent RAB discount.[8]

As a matter of regulatory design, it makes little sense to ransom the debt holders in respect of past sunk costs in order to encourage the operational side of the business to be efficient. On

---

[8] Alternatively, it could have been the windfall tax.

the contrary, the right answer is to provide sufficient equity incentives and returns to properly address the delivery of functions, and to minimise the cost of debt for the RAB. Unfortunately that is not what happens.

4.   **Regulation, the assignment of equity risk and different types of regulation**

It is the job of regulation to match risk and rewards in the protection of customers—to ensure that investment takes place, and that rewards are sufficient to incentivise the efficient delivery of CAPEX and OPEX, whilst preventing excess returns.

4.1   **The domain of RPI-X**

When RPI – X was first being considered, the main arguments were about the (in)efficiency of the alternative: rate of return. Rate of return regulation has a long pedigree, particularly in the US. (In the UK, we had nationalised industries.) In a classic article, Averch and Johnson (1962) argued that rate of return regulation encouraged cost inefficiency and gold-plating. Costs would be excessive and there would be excessive investment. This line of argument convinced Littlechild (DTI 1983) and the then Conservative government that rate of return regulation was to be avoided. Instead, regulation should mimic competitive markets, and therefore utilities should be price-takers. A fixed-price, fixed-period contract was to be offered so that profit maximisation would be equivalent to cost minimisation.

In setting up RPI – X, little attention was paid to the practical nature of rate of return regulation in the US, as opposed to Averch and Johnson's stylised representation. In particular, the very low beta coefficient of regulated utilities and the corresponding low cost of capital went largely unnoticed. So, too, did the fact that, in utility networks, over-supply was much to be preferred to under-supply: the costs were asymmetrical between the two cases.

It is therefore not surprising that there have been few attempts to demonstrate whether RPI – X has actually been more efficient once cost of capital effects have been taken into account. Such an analysis is particularly difficult because of the counterfactual: if RPI – X's handicap of a higher cost of capital was less relevant when there was excess supply, and hence less investment, it would hide the longer-term consequences.

The real efficiency problem with rate of return regulation lies not with the fact that *some* costs are passed through, but that *all* of them are. As identified above, the RAB costs should be

passed through since these are sunk and beyond the scope of management. However, it is the passing-through of the OPEX and CAPEX that creates the cost inefficiency and gold-plating. These are the areas where the RPI – X incentives might be deemed to be more effective (although not necessarily in five-year contracts, especially for the CAPEX).

But if rate of return covers too much, RPI – X covers too little if it is not explicit about the RAB. RPI – X should apply to the OPEX and CAPEX, but not to the RAB. Once this separation is made, the scope for competition becomes apparent. Whereas it is not possible to compete for the RAB (although there can be financial market competition through tradeable RABs), it is possible to compete for the OPEX and CAPEX. Such competition can be of two forms: competition for the whole OPEX and CAPEX programme, or competition for disaggregated parts. The latter is already widespread, with mixed results (eg, Network Rail). The former raises more interesting issues.

At the extreme, companies could be invited at periodic reviews to submit RPI – X bids. Provided there were sufficient bids, the task of the periodic review from the regulators' perspective would be to define the outputs and then the market could dictate the efficiency costs. In many respects this is similar to the franchise model, with the assets separately owned from the operational side.

This is a neat solution to regulation, and greatly reduces the regulatory burden, but it does raise some obvious objections. Bidders might have only one period, and therefore would have little interest in the long-term quality of the assets, and there would be problems over CAPEX projects of more than five years' duration. This is a problem well researched in the literature on franchises and is one reason why the relatively short-term franchises initially granted to train operating companies have been subsequently extended.

Such objections are serious, and raise the central issue of licence responsibility for failure as well as dynamic efficiency. As a result, the disaggregated competitive tendering model has more appeal since the incumbent remains in overall coordinating control, and holds the responsibility for delivering the functions through the licence. But if the domain of the RPI– X incentives should apply to the coordination, OPEX and CAPEX, it is here that the equity risk should be concentrated.

4.2   Assigning the equity risk

Equity risk never goes away. It can be assigned to shareholders, to customers or to taxpayers, or some combination of these three. In the private sector, companies allocate equity risk between themselves through contracts. Long-term contracts transfer risks, to suppliers and to customers. In the regulated sector, it is the regulators who assign this risk between the three classes.

Pure rate of return regulation assigns all equity risk to customers. In pure price caps, it is assigned to shareholders—this is what 'high-powered incentives' mean. In the case of Network Rail, equity risk is shared between taxpayers and customers as long as the bondholders have a government guarantee. In the case of Welsh Water, the equity risk lies with customers and possibly bondholders (and, in the latter, inefficiently). Welsh Water argues that it has an equity buffer, but this is retained profits—in other words, excess customer contributions in the context of a not-for-dividend model. Since it has no shareholders and since this is customers' money, it is customers who hold the equity risk, not Welsh Water.[9]

The art of regulation is to assign the equity risk in a manner which minimises overall costs to customers once incentives have been taken into account. The principles which help to assist this efficient allocation are:

- assign equity risk to those best able to manage it;
- maximise clarity and certainty about who holds the risk;
- do not assign equity risk to bondholders;
- ensure that the regulatory rules are consistent with the financial structure.

The discussion above indicates how this equity allocation should be organised for regulated utilities. The businesses have three parts, defined by functions: the RAB, representing past sunk investment; the coordination function in respect of the licence; and the day-to-day business of doing the OPEX and CAPEX.

---

[9] See Ofgem's comment on the Welsh Water structure (Ofgem 2000).

In respect of the RAB, the equity risk should be assigned to customers, for whom the investments in the sunk costs have been made. In respect of coordination, the equity risk should be imposed upon management, as it should for the OPEX and CAPEX. Coordination and the delivery of CAPEX and OPEX may be carried out by different companies—indeed, much of the OPEX and CAPEX already is in practice. Coordination goes with the licence, and the reward/risk trade-off comes from outperforming the regulatory contract for OPEX and CAPEX whilst meeting the required outputs. But, even here, it may be efficient for customers to share some of this risk, particularly where there is uncertainty about outputs and where there are exogenous shocks. In the former case, customers may want the option of changing the requirements within periods. In the latter case, if companies bear exogenous shock risk, customers will face a higher cost of capital. As in the private sector, a risk-sharing approach can often be more efficient than a corner solution with either customers or shareholders.[10]

5.   The WACC and the cost of capital

Financial structures should follow the nature of the business and the underlying risk characteristics. They should not—as they began to after the dash for debt—drive the business and its management decisions. Although utilities vary considerably, their core common characteristics identified above are the sunk assets, coordination of networks, and the management of the operations of those networks and investment in them. To determine the cost of capital that utilities face—and hence to estimate the required rate of return—necessitates consideration of each of these components in turn and then aggregating them. Amazingly, this is not what British regulators do; instead they consider all the activities of the business as one aggregate, and then calculate a WACC, typically via the capital asset pricing model (CAPM).

The CAPM approach has considerable merits as a classification of the components of the cost of capital. It separates out the cost of debt from the cost of equity, and treats the latter from a portfolio perspective. Behind it lies the efficient markets hypothesis (EMH), which assumes that capital markets allocate risk efficiently (all trades are exhausted), with the price of a share summarising all relevant information. Market prices are 'correct' because no informational stone has been left unturned. The CAPM (and the EMH) are not 'true', however—capital

---

[10] Ofgem has conflated the regulation with the returns to equity by introducing the concept of 'regulated equity'.

markets are neither perfectly competitive nor perfectly efficient, not least, as recent events have demonstrated, because there can be significant systemic risks which undermine the classic portfolio diversification results.

But the efficiency or otherwise of capital markets aside, the classification is helpful. Taking each part in turn, the cost of debt is calculated as the risk-free rate (approximated by the rate on government bonds) and the risk premium. Recent events have cast a shadow over the risk-free rate—as governments engage in quantitative easing (QE) in an attempt to push down the long-run returns on government bonds, and when government debt itself has become a little more risky.

The risk premium is more difficult still. What is the risk to which the premium adheres? Here it matters greatly which bit of the business is involved. For the RAB, as discussed above, as long as the duty to finance functions applies to it, it is hard to see that there should be *any* premium. Indeed, given QE, it might even be possible to argue that government bonds are more risky since the utility assets are real, whereas the government's asset is its ability to gain a democratic mandate to put up taxes to whatever level is necessary. The security of utility assets is further enhanced by the ability to pass through inflation (the RPI indexing), whereas for the government, a sharp rise in inflation drives up the nominal costs of accumulated debt costs. In the current circumstances, investors may want to price into the cost of government debt the possibility that the rate at which current borrowing is adding to the national debt and the inflationary possibilities from QE may combine to induce a flight from government bonds, sharply raising the 'riskless' rate used in the CAPM.

The risk premium on debt if the RAB is not protected by the duty to finance functions is altogether different, for the utility can carry out its functions even if the RAB is reduced to almost zero—it will still have enough funds to carry out the coordination, OPEX and CAPEX functions. Given, too, the time-inconsistency problem outlined above, the RAB is seriously at risk of opportunist regulators and politicians. It now becomes a very risky asset, given in addition that it is not within the competence of the utility to manage it.

The debt premium for the running of the business is likely to be much higher than that of the RAB protected by the duty to finance functions. Indeed, this higher cost is the reason why companies that are engaged in OPEX and CAPEX tend to be debt-light. This is because the risks are equity ones, and as a result transferring equity risks to bondholders tends to be very expensive. Debt in service companies tends to be largely for cash-flow and working capital,

while, for CAPEX, debt tends to be project-based, tied to specific assets and linked to the exit on completion (in this case, 'sale' into the RAB). The rationale of RPI – X is precisely to heighten these equity risks, to create the incentives (and risks) for management to maximise profits through cost minimisation.

Having tackled the cost of debt, the CAPM separately calculates the cost of equity. This also has two parts: the return on an efficient portfolio of diversified equity (the equity risk premium), and the beta coefficient, measuring the extent to which the particular share price varies with respect to the market. If beta equals one, the share is perfectly correlated with the market portfolio; if closer to zero it is not (and hence has lower equity risk). Historically rate of return US utilities had beta coefficients around 0.2, whereas RPI – X-regulated utilities have betas at much higher levels—even at around one in some limited cases.

In calculating both the market return and the beta, the forward estimate used at periodic reviews is based upon past data. Unless a good case can be made for quick (within the five-year period) mean reversion, it is therefore always prone to error, especially if there is reason to believe that structural breaks have occurred. In the case of the equity risk premium, it has varied very considerably over the last century. After the long boom of the late twentieth century, there are good reasons to believe that history will reveal these decades to be 'exceptional' rather than normal. The credit crunch has further increased the uncertainty, and whatever is the mean to which the EMH assumes the equity risk premium will revert to, it is unlikely to happen any time soon—and not in the next periods for which the cost of capital is currently being set.

When it comes to the data on utility betas, this is narrow, and becoming narrower, as more and more utilities have been taken into private hands. There are now few large quoted utilities. They include: Severn Trent, United Utilities, Pennon, Northumbrian, BT, Scottish & Southern, and National Grid. Of these, National Grid has around half of its business in the US under a different regulatory regime; BT is an international and heterogeneous business blighted by its pension fund deficit; Pennon has a significant waste business; and Scottish & Southern is vertically integrated. That leaves just three 'pure' quoted utilities, covering only one sector—water. The share price series goes back around 20 years, but all of this is in the long boom years, and at least half relates to periods when the rates of return were boosted by the post-privatisation cost savings and by the merger wave. It is very hard to argue that this time series is representative of the 'mean'.

It remains to tie the cost of debt and the cost of equity together to gain a WACC. In a conventional unregulated company, the gearing assumption follows from market data— whatever comparative companies' financial structures happen to be. This changed over time: with the coming of private equity players, leveraged buy-outs led to a substantial rise in gearing—both for targets of private equity and those who wished to keep out of their grasp. The unprecedented (until 2007) monetary and fiscal expansion after the stock market crash of 2000 exacerbated this preference for debt over equity as real interest rates were sharply reduced.

In the case of utilities, it is the regulators who impose an exogenous gearing assumption on the WACC. There are a number of options, including the gearing which represents real CAPEX not paid out of customers' current bills (the intention at privatisation); actual gearing; and some notional assumed 'optimal' gearing. Regulators have overwhelmingly opted for the last of these—on the grounds that it is for companies to choose their own financial structure.

This argument is quite extraordinary, for its implication is that the financial structure of the companies has little or no bearing on its behaviour and performance, and its specific circumstances in respect of CAPEX. High- and low-geared companies must be assumed to behave in similar fashions. But this assumes that the interests of debt and equity holders are the same—that each has a similar view about how much risk managers should take, and how they manage the risks that they do take. Debt and equity are fundamentally different: debt holders are focused on the interest and the return of their capital; equity holders share in the capital gains (and losses) in the business. The influence of bondholders will be greater in higher-geared companies and, as a result, they can be expected to be more risk-averse.

The other two options are better than the one taken. But of these, the second is fraught with difficulties. The reason is that there is no obvious answer to the optimal gearing of a utility considered in aggregate. Indeed, there is little agreement as to the optimal gearing generally. The famous Modigliani–Miller theorem states that, under demanding assumptions, the capital structure of a business does not matter. The most important of these assumptions is arguably about taxation, but it would be hard to argue that the optimal capital structure is primarily a tax issue. Indeed, since the tax treatment between debt and equity tends to favour debt (because interest is tax-deductible) and capital gains over dividends, it remains a puzzle as to why firms pay dividends at all.

This is not to argue that there are not good reasons for having debt and that the disciplines of debt may actually benefit equity holders. Rather, it is to suggest that the complexities and 'puzzles' are such that it is very hard for regulators to decide what the optimal capital structure ought to be.

This leaves the first option: gearing should be based on actual CAPEX not financed through current customers' bills—in other words, to be the bridge between current CAPEX and future (benefiting) customers. That was clearly the intention at privatisation—indeed, it was the explicit reason why the water companies were given an initial green dowry of a cash injection. But even this measure has its complexities: the gearing comprises the value of the new investments that have gone into the RAB and project finance for assets in the course of construction.

The WACC that results is therefore essentially crude, but there is at least an *ex post* check—whether, as a result of re-setting the WACC at periodic reviews, the market value of the utility is in excess of the RAB—representing the RAB itself plus the expected outperformance against the price cap in the period. This is Tobin's Q, and it is apparent that share prices have varied widely from this measure since privatisation. For most of the post-privatisation period (especially in the initial period, and then when takeovers came along), market values have been well above the RAB. But there have been times when they have been below, notably for water at the 1999 periodic review through into 2000, and after the recent credit crunch. One interpretation of the positive premia is that the efficiency gains turned out to be much larger than anticipated. A second is that regulators took time to learn the game. And a third is that exogenous variables—notably the interest rate—kept falling from 1990 onwards, and that at each periodic review, regulators underestimated the subsequent fall. All three in fact played a part, and as the gearing levels increased, so the interest rate effect increasingly predominated.

To these three we need to add a further, crucial, factor: financial engineering. By definition, the WACC is an *average*—higher than the marginal cost of debt and lower than the marginal cost of equity. As a result, there was a massive arbitrage opportunity. The surprise has been that it took time for financial markets to appreciate the opportunity—perhaps in part because it was far from clear that regulators would allow it to be exploited, and in part because it took time for the threat of takeover to focus the minds of both incumbent managers and raiders on how to deliver the maximum short-term financial benefits to their shareholders.

The process started almost by accident. Once the first re-setting of the distribution price control had been determined in August 1994, and the golden shares had been relaxed, Trafalgar House, a company now long gone, made a bid in the December of that year for Northern Electric. The bid itself was not particularly remarkable, except that it valued the target at well above the RAB, despite the re-setting of the price cap. It merely indicated the scale of the then regulator's error.

What transformed the situation was the decision of the Northern Electric board to fight, and in particular its second defence document. In this, it was proposed to mortgage Northern Electric's assets, and return around £5 in cash for every share that had been purchased only five years earlier for £2.40. In the end, shareholders received not only the £5, but also in a subsequent bid around £7 a share as well—multiplying the initial investment by at least a factor of five (and there were several years' worth of dividends as well) for what on any reasonable assessment was a low-risk business.[11]

Northern Electric's defence started the ball rolling, and from this point onwards gearing was increased through a variety of measures, including special dividends, share buy-backs and acquisitions. Although, over time, the regulators gradually raised the bar through higher notional gearing assumptions and eventually by clawing back the tax benefits from higher debt, the returns were very considerable. Assuming that the RAB—the overwhelmingly dominant part of the asset value—could be backed by debt, because, as argued above, the equity risk in the RAB had been transferred to the customers through the duty to finance functions (as a solution to the time-inconsistency problem), the companies could carry on mortgaging their assets and banking the difference between the WACC and the marginal cost of debt.

Regulators have duties to protect the interest of customers. This arbitrage leaves customers worse off in at least three ways. First, they are paying a premium on the RAB (the difference between the cost of debt and the RAB) and at the same time taking the equity risk on the RAB. Second, the balance sheets have been depleted, so pay-when-delivered will gradually have to be replaced by pay-as-you-go for CAPEX (since the WACC is below the cost of equity necessary to refresh the balance sheets now that they have been significantly impaired

---

[11] For a detailed description of the Northern Electric episode and subsequent takeovers in the electricity sector, see Helm (2004), pp. 214–7, and chapter 12.

by the financial engineering). Third, the 'day job' of coordination, OPEX and CAPEX is under-rewarded because the WACC is below the cost of equity and, as discussed above, these aspects of the utilities are where the equity-based incentives are located. Hence, the high-powered incentives towards efficiency, which is at the heart of RPI – X, are not properly remunerated.

There is a fourth, less obvious but potentially more serious, detriment to customers. Very highly geared companies may fail completely, and in the run-up to collapse may focus on cash management, cutting back on OPEX and CAPEX. Although there is now a special administration function for most utilities (but not yet for airports), the practical imposition of such a special administration regime is fraught with difficulties. Before collapse and special administration there is the insidious gradual failure to deliver the functions as cash management predominates. It might take years to realise what neglect of maintenance and some aspects of CAPEX implies for customer services in essential network utilities.

The obvious solution to the arbitrage opportunities offered by using the WACC is to split the cost of capital into two main parts: a cost of capital for the RAB, and a cost of capital for the coordination, OPEX and CAPEX functions. Since there are established RCV numbers for all the main utilities, and since this number is formally rolled forward with added efficient CAPEX at periodic reviews, the distinction is a sharp one.

The equity risk on the RAB lies with customers. That is what the 'deal' between them and investors incorporates. Investors sink capital in the full knowledge of the *ex post* gap between marginal and average costs, but do so because the RAB represents their long-term contract. Therefore it is efficient to finance the RAB through debt, backed up by the duty to finance functions. (We return below to the need to clarify this duty further.) The gains to customers of matching the financial structure to the allocation of regulatory risk via the regulatory framework are potentially very large, and probably larger than any conceivable efficiency gains.

On the other hand, the 'day job' should be financed largely on the basis of the cost of equity, adjusted for some project finance (at higher debt costs than for the RAB). The day job would therefore be higher-risk/higher-reward, as befits the incentive-based RPI – X regime. So some of the customer gains from a low return on the RAB would be offset by higher returns on the equity.

It is important to recognise just how important this equity return is: without the possibility of abnormal returns, the incentives are blunt. Why outperform an RPI – X price cap if there is no extra reward to shareholders? Indeed, the consequences can be seen in the three cases where RPI – X incentives are applied in the absence of private shareholders: Welsh Water, Network Rail and the Royal Mail. The only traction for incentives is that managers are interested in the public benefits to customers and, if they are publicly minded, they do not need any incentives. The managers themselves can be incentivised, and this could simply be by tying their salaries and bonuses to performance. But then RPI – X price caps are not needed at all if the management pursues the public interest. And if performance contracts for managers were difficult, private equity ownership could be dispersed across much of the economy.

Since the split cost of capital was first proposed,[12] a number of objections have been raised. The scale of the opposition is not surprising: the split cost of capital lowers the aggregate returns to shareholders, puts a stop to much of the new industry of financial engineering and the gains to the City that have accrued, and challenges the regulatory practice developed by the regulators and the Competition Commission. It is hard to construct a more formidable set of opposing interests.

The criticisms can be divided into objections in principle, and objections in practice. The objections in principle have to validate the claim that there is some equity risk in the RAB ascribed to shareholders. It is claimed that the RAB is not separable because the assets are required to fulfil the licence and the RAB is therefore at risk if the company fails to deliver the functions and hence loses its licence, presumably through special administration. This criticism has merit in so far as the financing functions duty in respect of the RAB has never been tested in the courts, and regulators have not given sufficient clarity on its definition. However, whilst it is correct that shareholders can lose the licence, the position of debt holders is much more protected. This is because, in the special administration process, the licence will be offered for resale, and the value of the resale should be equal to, or slightly above, the RAB. Thus the value of the business remains comfortably in excess of the gearing levels currently witnessed in most, if not all, regulated utilities. A situation might arise where the liabilities of the utility in special administration are so large as to reduce the sale value below the RAB. The most obvious example is pensions. However, this is not an argument

---

[12] See Helm (2006).

which requires that the RAB attracts a WACC to reflect equity risk; rather, it is an argument about the functions which the regulators are obliged to finance. Since pension contributions come out of the returns on the business operations, if there is not full cost pass-through then the risk is an equity one, linked to profits. In most cases regulators have allowed for pension cost pass-through.

The crucial, and simple, point here is that there is only equity risk in the RAB to shareholders if the government and the regulators expose them to the time-inconsistency problem. It would be extremely inefficient to do so since shareholders would then be carrying risks over which they have no control, and the equity cost of capital to the RAB would be very large. The principle that risk should be assigned to those best able to manage it would have been violated. Since there is nothing whatsoever the managers can do to change the accounting number, which is the RAB, it would be to the detriment of customers.

Threatening to undermine the RAB is one aspect of the proposals to introduce 'competition' to the water industry. So far, as Nourse (2009) has shown, shareholders appear to place little weight on this possibility. A further example is where assets are taken out of the RAB—as is the case for metering in electricity distribution. Perhaps the most significant cases arise where the income of utilities depends upon volume—here the returns to the RAB are dependent on the behaviour of the economy more generally and technical change. In BT's fixed network this is very apparent. Where these risks in respect of sunk assets are exposed in this way, investors will expect higher *ex ante* costs of capital or regulatory assurances. The case of broadband is illustrative of this risk of stranded assets—indeed the relative backwardness of broadband roll-out in Britain may in part reflect the lack of a RAB-type regulatory protection.

The objections in practice are more credible, and fall into two categories: the practicalities of implementing the split cost of capital; and the implications for investors' legitimate expectations given the current WACC approach. In the case of the former, the approach is straightforward. The RABs are already established by the RCV numbers, as is the process of updating at periodic reviews. The regulators will need to be explicit about their interrelation of the duty to finance functions, and the underpinning of the RABs. Given the number, the cost of debt can be applied. For the rest of the business a WACC can be calculated, based upon the costs of equity and debt and the gearing that comparative businesses engaged in these sorts of business activities receive in the market. This calculation would need to take into account project finance costs in respect of CAPEX, as well as network management and operations. Wherever possible, competitive tendering can be used to test market-related costs.

There are a number of practical advantages of this approach. It would not need to revisit the relationship between CAPEX not financed from current customer revenues in relation to the gearing, since this would now be taken care of through the upgrading of the RAB. Financial ratios would apply to the operating side of the business, net of the RAB. It would also bring to bear the costs of project finance, and as a result put these aspects of the business on a par with PFI – type approaches.

The main practical consequence would, however, flow to the incentives—there would be a higher and more appropriate return to equity. In turn, rights issues would be a business option, though there is no obvious reason to believe that, with equity properly rewarded and with assets sold into the RAB at periodic reviews, these parts of the business would need to be highly geared.

The impact on investors' expectation is more serious. Investors have been led to believe that they are entitled to a WACC, and the methodologies for calculating these numbers are well established. It is therefore argued that to change the methodology now undermines the predictability of regulation, and hence raises the cost of capital.

This claim is indeed correct. But it does not follow that the methodology should be preserved, for a number of reasons. First, being wrong in the past does not carry with it the implication that it is appropriate to be wrong in the future. The scales of the gains to consumers—both through the cost of debt in respect of the RAB and the benefits from a greater equity incentive and return on the rest of the utility activities—are likely to be significantly greater than the cost of capital effect as a result of the change to expectations. This indeed was what happened in respect of British Gas and the MMC, where a change in methodology invited the debate about correcting past errors as against regulatory continuity. Second, regulators have in any event been continuously changing the methodologies as they go along. The RAB and even the cost of capital did not figure in the original privatisations—they were developed afterwards, notably by Oftel and Ofwat. Then there have been changes to the tax treatment of debt, clarifications to the pensions, and, crucially, the introduction of common notional gearing across companies in the same sectors independent of their actual CAPEX. The exercise of regulatory discretion in respect of the financial and other aspects of regulation have been the norm, not the exception.

It is therefore debatable how far account should be taken of legitimate expectations. But since there is a powerful set of lobbies in favour of the existing approach—including the CC and

the regulators, the utilities and the City—and these will create a major obstacle to change, it may be appropriate to phase the change in, perhaps over more than one period. The extent of any share price movements will depend upon the accompanying force of the clarification of the duty to finance functions, and in addition the increase in the totality of debt in respect of the RABs. The higher return on the running of the businesses will also need to be taken into account. Timing will be important too: if regulators set the parameters as recommended by the split cost of capital, the lower risk on the RAB and the higher return on the rest of the business should result in an exit value for the shareholders in respect of the RAB—the remainder of which is in effect sold to debt holders. Furthermore, an uprating of the value of the rest of the business, such as the total impact in terms of valuations, should be (very) small. If, at a periodic review, the market values had in any event fallen below the RAB (as they have in 2008 for water), there may be a gain to shareholders from this change in methodology.

6. *Ex ante* **and** *ex post* **costs of capital—the role of indexation**

At its core RPI – X has the idea that prices should be set *ex ante*, to make profit maximisation equal cost minimisation, in order to mimic competitive markets. We have already noted how crude this is in respect of long-term contracts, and linked this to the treatment of the RAB. But it is also true that, in competitive markets, prices vary—in particular, in respect of exogenous shocks. We would not expect oil companies to fix the price of retail petrol once every five years, whatever happened to crude oil prices; similarly for bread if the price of grain varies. Thus RPI – X deliberately acts in a non-competitive market way in order to sharpen incentives.

It was noted above that this is not efficient in respect of the RAB. In addition, it is far from obvious that financial costs should be fixed in five-year blocks since costs of debt, in particular, vary continually, not least in respect of exogenously set interest rates. *Ex ante* estimates by regulators have so far been consistently generous, and sometimes by a large margin. Why, then, set a fixed *ex ante* cost of debt in a context of an exogenously determined number? The defence is that company treasurers should have incentives to beat the market. But what exactly is this innovation? Do company treasurers know better than the market what the Monetary Policy Committee of the Bank of England will decide on interest rates or QE? Indeed, it could be argued that it is precisely because the risk of an exogenous variable has been placed upon those who cannot influence it that there is a debt premium, and at the level

currently observed in the market. From a customer perspective, this is hardly an efficient bargain.

There is therefore a competitive markets reason for indexing the cost of debt to the market rates, and on a more frequent basis than the current five-year indexing. In the current context, where the uncertainty is greater than at any time since privatisation, the case for indexation is overwhelming. The practical question is what form the indexing should take, to which there is no obvious correct answer. Any index less than five years will improve on the current arrangements. It could be an annual adjustment, based on forward rates; it could be monthly, or even weekly or daily. The case for indexation does not depend on the exact index and time interval chosen. Perfect exogeneity is unlikely to be attained in any event. Almost all indexes are, however, better than the five-year crude approach currently used by regulators.

This indexing does not of course mean that prices vary on the same frequency as the index. Prices are currently changed annually. There therefore needs to be an *ex ante* yearly prediction, with an *ex post* error correction for the divergence between the assumed rate and the outcome. The error correction amount is what is determined by the chosen index.

A subsidiary question is what rate the index should be linked to, once the time interval has been determined. There is a good theoretical case for using an index-linked rate—perhaps a small premium to index-linked gilts. Although the market for index-linked debt is less developed than for fixed rates, the utilities' prices are themselves index-linked (as is the value of the RCV). For pension funds and similar long-term institutions, the creation of a utility-wide index-linked debt market may be a significant benefit to financial markets more generally, and to the provision and funding of index-linked pensions. If the RAB debt becomes widely tradable, this would add to the liquidity and depth of the index-linked markets.

Given the current uncertainty, it may seem surprising that regulators faced with fixing *ex ante* costs of debt do not reach for indexation. A joint Ofgem and Ofwat paper (2006) toyed with this reform. The reasons for this reluctance are several—not least that to embrace indexation implies that regulators have been wrong not to use this methodology earlier. But there is a second, more worrying, reason for the observed reluctance; namely that regulators have wanted to keep the financial variables in play with the whole of the periodic review. Thus, if it turns out that the cost of debt rises well above the *ex ante* assumed rate, it is argued that utilities can always come to the regulators and ask them to re-open the price cap. Critically,

this re-opening would involve not just the financial variables, but also the OPEX and CAPEX too.

The mistake here is to confuse the endogenous with the exogenous. It is very important to maintain the efficiency incentives for those aspects of the business which managers can control. It is important not to incentivise managers for those exogenous variables they cannot control. Thus the case comes back to the exogeneity of the interest rate: regulators are in effect saying that the cost of debt is endogenous.

7. **Infrastructure, systems and public goods: the role of sectoral plans**

The arguments set out above in respect of the RAB and the split cost of capital have a wider resonance than just the utility sector. They have implications for the way in which the infrastructure of the economy as a whole is developed and maintained, and for the economic borders between the state, the infrastructure and capital markets.

The conventional focus of economic policy is split between macroeconomics (fiscal and monetary policy) and microeconomics (tax, competition policy and regulation). Much less attention has been paid to the role of infrastructure in the overall performance of the economy in British policy—though not so in France and Germany. Infrastructure matters for several reasons: it affects the costs of all firms and individuals, and its costs are strongly asymmetric between too much and too little infrastructure capacity. It will be inadequately provided by the private sector not only because of the time-inconsistency problem, but also because the asymmetry points to the public need for excess supply in order to ensure security of supply. It is, with law and order and defence, a key public good.

In the twentieth century, infrastructure was largely provided in the public sector, and still is in a number of European countries. The advantage has been to the cost of capital; the disadvantage has been to efficiency. The RAB-based model defended here provides a very general model of how the roles of the public and private sectors can be defined. The state guarantees that efficient sunk costs are paid for; the private sector ensures that they are efficiently delivered.

For the economy as a whole, both the cost of capital and the efficiency of delivery matter greatly. But it also matters how much infrastructure is provided. For network utilities this is not a question which can be addressed purely from a private sector perspective. The optimal public goods are necessarily publicly defined. For the utilities, a crucial mistake in the early

post-privatisation days was to leave this decision to be negotiated between the companies and the regulators during periodic reviews. Gradually it has been recognised that an overall national view needs to be taken of the required outputs and supporting networks.

For each of the sectors this has gradually been put in place. For energy, the government and the EU specifies the amount of renewables to be provided by 2020, and this entails network investments to support the decentralised generation and the smart meters necessary to manage demand. In water, the government and the EU specifies the quality of rivers, the quality of water and the measures necessary to address climate change. Metering, too, is a government decision. In railways, the government provides subsides and guarantees Network Rail's debt. It sets out the transport strategy. Roads are government-driven (both local and national). In airports, new runways are government decisions. Even in communications, broadband roll-out is government-driven.

None of these is precisely specified, and much remains to be sorted out between companies and regulators. But that does not distract from the prime responsibility for decision-making. Having recognised this, it is also widely recognised that much of British infrastructure is at best imperfect and in many cases not fit for the purposes of the coming decades. A very major investment programme is required, to deliver new transport systems, new energy networks, communications networks, and airports. In the nationalised industry framework, the overall quantity of investment was constrained by fiscal policy and the government's cash-based budget. Borrowing and debt targets further constrained spending, and politically it was usually easier to cut capital rather than current spending. This was the impasse reached in the early 1980s which kicked off utility privatisation.

The RAB-based model enables the investment spending to be independent of short-run government budget constraints in an efficient manner. The Private Finance Initiative schemes, as well as utility privatisation, have made some of the necessary steps, but they still lack clarity. Suspicion that PFI in particular has been an off-balance-sheet borrowing exercise has led to a gradual readmission of the debts onto the public balance sheet. But in sorting out what is and what is not public spending, the crucial issue is the allocation of equity risk. Where equity risk lies with the taxpayer, it should be in the public sector accounts. Where it lies with consumers, it should be treated as private. This simple principle allows an allocation to follow which keeps much of the infrastructure out of public accounts. Water, energy and aviation lie outside; rail and roads lie partly inside. The role of government is to guarantee that customers will pay in the former case; in the latter case, it is taxpayers' risk.

This principle enables us to view infrastructure in a broader perspective. An example is renewables. The government sets the targets and through the Renewables Obligation Certificates compels customers to pay for the extra costs of renewables, notably wind. But in making this compulsory, it neglects to recognise that the major costs are sunk. Providing a RAB-based approach to the sunk costs would reduce the costs of the Renewables Obligation substantially. (Currently it is amongst the most expensive in the developed world.) The case of the utility-like services of banks is also interesting: if there are taxpayer guarantees then there is a corresponding reduction in the cost of capital which should be passed to customers; if the guarantee is from a company-funded insurance scheme, it is customers who are ultimately absorbing this equity risk. What has been created in the credit crunch is, in effect, an enormous banking RAB. It remains to be seen whether this should be entirely handed back in the current form to the private sector, or whether part of banks' activities should be ring-fenced into utilities, and, if so, in what regulated form.

8.   **Reforming regulation—an evolutionary approach**

In classic British pragmatic style, what began as a crude and simple fixed-price, fixed-period regulatory rule has migrated into a model for a more efficient allocation of costs and risks between customers, investors and taxpayers. RPI – X is now in practice almost unrecognisable against the initial model. The process of this evolution has been driven by regulators exercising discretion, by events, and by public pressure over both the quality of service and the prices.

This process has had many benefits, but it has costs too, not least the degree of micro-management and the sheer complexity of the periodic review process. At the limit, water companies now produce enormous submissions of their business plans. The end product is pregnant with an efficient approach to regulation, but so far it remains expensive and the end result for customers and the economy more generally is far from a satisfactory level of provision, service or costs.

At the heart of a more efficient system of regulation is the RAB. It is the time-inconsistency problem that matters for dynamic efficiency as much as, or more than, the current focus almost exclusively on monopoly market failure. Time inconsistency requires much of the political and regulatory processes, since it applies over time periods often generational. Neither the private unregulated monopolies of the nineteenth century, nor the nationalised

industries of the twentieth century provided satisfactory solutions. The evolving RPI – X framework provides the germ of a better solution, but it remains poorly developed.

The reforms necessary to make the transition to a better balance between the roles of the public sector and the private sector, and in the allocation of equity risk are not themselves revolutionary. They do not require what has been put in place over the last two decades to be torn up. They are evolutionary. There are two key steps: to further define the duty to finance functions so that it guarantees the sunk costs in the RABs; and to follow this through by implementing the split cost of capital.

From these two steps, much else follows. The RABs become potentially tradable, and are likely to be debt-financed by institutions such as pension funds. Since the investments tend to be long-term, and pension funds are necessarily generational, the match is clear. By indexing the cost of debt appropriately, a major extension of index-linked debt becomes available. For the business of delivering the OPEX and CAPEX, the split cost of capital enhances the role of equity and hence the power of incentives. It harnesses the private sector to the challenge of efficient delivery. At the limit, major parts of the OPEX and CAPEX could be delivered through competitive tender—although there are drawbacks here too, and coordination has to remain with the licence itself. Nevertheless the split cost of capital approach maximises the scope for competition and cost minimisation in a way that would probably be appreciated by the original architects of RPI – X.

These reforms are not just ideas—not pursuing them has real costs which are currently shared by customers and in some cases taxpayers. Customers are both paying the WACC premium on the RAB and taking the equity risk in respect of the RAB. Financial engineering has been largely to the gain of shareholders and financial institutions. And the efficiency incentives for the OPEX and CAPEX have been blunted by the failure to provide a proper equity return on these parts of the businesses. The result is that companies whose balance sheets have been exhausted have not been able to launch rights issues (except in distress). In consequence, they have either collapsed (Welsh Water and Railtrack), or regulators have had to move towards pay-as-you-go for CAPEX. The resulting state of British infrastructure is hardly the envy of Europe.

So far these costs of the existing system have been masked by the initial efficiency gains and the gradual process of gearing up the balance sheets. Going forward, the position is likely to deteriorate more rapidly. With the WACC approach, and with normal rights issues unlikely,

either prices will have to rise more steeply to finance investment from current bills (pay-as-you-go), or the companies will gradually collapse by default into mutuals as equity exits. Already the impacts can be seen in the airports and railways. The challenges for the electricity networks are enormous, requiring major enhancements and upgrades to meet the low-carbon obligations. How these will be paid for, by whom and at what costs are questions which are sufficiently important to weigh on the economy as a whole, and hence to draw the attention of economic policy as well as sector regulators.

References

Averch, H. and Johnson, L.L. (1962), 'Behavior of the Firm Under Regulatory Constraint', *American Economic Review*, **52**:5, 1052–1069.

Beesley, M. E. and Littlechild, S. C. (1989), 'The Regulation of Privatized Monopolies in the United Kingdom', *RAND Journal of Economics*, **20**:3, 454–472.

DTI (1983), 'Regulation of British Telecommunications' Profitability', Department of Trade and Industry, London: HMSO.

Helm, D.R. (2004), *Energy, the State and the Market: British Energy Policy since 1979*, Revised Edition, Oxford University Press.

Helm, D.R. (2006), 'Ownership, Utility Regulation and Financial Structures: An Emerging Model', January 14th, available at www.dieterhelm.co.uk.

Helm, D.R. (2009), 'Credibility, commitment and regulation: ex ante price caps and ex post interventions', in *Populism in Natural Resources*, MIT Press, forthcoming.

Helm, D. R., Hepburn, C., and Mash, R. (2005), Credible Carbon Policy, pp. 305–321, in D. R. Helm (ed.), *Climate-change Policy*, Oxford, Oxford University Press.

HMT (1967), 'Nationalised Industries: A Review of Economics and Financial Objectives', Cmnd 3437, Her Majesty's Treasury, London: HMSO.

Littlechild, S. (2009), 'Regulation, Over-regulation and deregulation', Centre for the Study of Regulated Industries Occasional Lecture 22 held November 24th 2008.

Nourse, R. (2009), 'Competition proposals and financing issues: A report for Ofwat', January.

Ofgem (2000), 'New Ownership Structures in the Water Industry: A Response to the Director General of Water Services' Consultation Paper', July, London: Office of Gas and Electricity Markets.

Ofwat and Ofgem (2006), 'Financing the Networks: A discussion paper', February, London: Office of Water Services and Office of Gas and Electricity Markets.