



2001 Annual Forum
at Misty Hills, Muldersdrift

Regulatory Reform: Lessons from the UK

Simon Cowan

University of Oxford

10-12 September, 2001

Regulatory Reform: Lessons from the UK

Simon Cowan

University of Oxford
Department of Economics
Manor Road Building
Oxford OX1 3UL, UK
simon.cowan@economics.ox.ac.uk

Paper prepared for the TIPS Forum 2001: New Directions in the South African Economy,
Johannesburg, 10-12 September 2001.

1. Introduction

In the 1980s the UK (and Chile) began the processes of privatizing and restructuring state-owned enterprises, liberalizing the markets in which they operated and regulating their conduct. Since then many countries at all levels of development have implemented their own programmes of regulatory reform. Almost two decades after the process started it is time to take stock and to reflect on the general lessons to be learned from the UK experience.

In the rest of the introduction I outline briefly the major events and themes. For detailed surveys of UK regulatory reforms see Armstrong *et al.* (1994), Helm and Jenkinson (1998) and Newbery (1999). Cowan (2001) covers the theoretical principles of regulation and relates them to UK experience. There were many reasons for privatization and the accompanying regulatory reform. State-owned enterprises had performed poorly in terms of both productive and allocative efficiency because of imperfect monitoring, unclear objectives and lack of competitive pressure. Privatization and, particularly, competition would weaken the power of trade unions in the enterprises. The government was unwilling to finance the major investment programmes required by the enterprises. Shifts in demand and in technology reduced the scope of natural monopoly conditions. The government was keen to obtain an immediate inflow of cash from asset sales. Equity sales offered the opportunity to extend share-ownership.

The main landmarks in the process were the privatization of telecommunications (1984), natural gas (1986), airports (1987), water and waste-water (1989), electricity supply (1990/91) and rail (1994/95). Sector-specific regulators, with a large degree of independence from the government, were established around the time of privatization to regulate dominant firms' conduct, especially their pricing. The main innovation as far as conduct regulation was concerned was the use of price caps, backed up by yardstick competition in the regionally organized industries (water and electricity). Setting price caps, however, has not proved as straightforward as was imagined by early proponents, and it turned out that regulators have to make substantial use of the tools used when applying rate-of-return regulation. A difficult aspect of price control has proved to be the determination of access or interconnection prices, which are the prices that a vertically integrated firm charges rival suppliers for the use of its network services.

Issues of industry structure were initially left to one side, and British Telecom (BT) and British Gas were privatized as vertically integrated dominant incumbents. This contrasts with policy in the USA at the time, which separated AT&T vertically into a long-distance company and the regional Bell operating companies. In the UK a single firm, Mercury, was licensed to compete with BT in the long-distance market with the guarantee that there would be a duopoly for seven years. Since the end of the duopoly policy in 1991 entry into all telecommunications markets has been fully liberalized, though a decade later it is still the case that BT has a dominant market position. British Gas was left untouched as a monopsonistic buyer of gas, as the only gas transporter and with a *de facto* monopoly in the supply of gas. Although entry into gas supply for large customers was liberalized in a legal sense there was no competitive entry for almost ten years after it had become feasible. Thereafter the regulatory authorities chiselled away at BG's structure and conduct, until BG split itself vertically in 1997 into separate transportation and supply companies. Full competition in supply has existed since 1998.

The regional water and wastewater companies were privatized as vertically integrated entities, and only recently have there been moves to introduce some product market competition in water. The electricity industry, though, was subject to large-scale restructuring. Transmission was separated from generation and the generating company was split horizontally into three competing businesses. A power pool was established to allow spot market trading of electricity. Distribution and retail supply remained regionally organized, with a rolling programme of allowing entry into retail supply. The retail supply markets were fully liberalized in 1998. Ownership structures have changed significantly over time, with many of the regional companies being owned by foreign utilities, and some vertical reintegration has been allowed. Distribution and supply businesses have built new generation capacity and generators have been allowed to buy supply businesses once market power was deemed not to be an issue in generation.

The railway industry was subjected to the largest restructuring of all at the time of privatization. The monolithic state-owned British Rail was transformed into more than eighty companies. Provision of track, stations and signalling was separated from train operation, which was itself split into separate routes and regions and franchised. Train ownership was

separated from train operating companies, and track maintenance was likewise divorced from track ownership.

In the next section I consider the regulation of conduct, including price controls, quality and investment monitoring and the use of spot markets for intermediate products. Section 3 covers structural regulation and entry conditions. Section 4 considers the role of regulatory institutions. Section 5 concludes.

2. Conduct Regulation

Kay and Vickers (1988) in an early analysis of UK regulatory reform distinguish conduct regulation from structure regulation. We follow that distinction and focus in this section on the regulation of behaviour taking industry structure and entry conditions as given. The main focus is on the regulation of pricing of dominant firms, but I shall also briefly discuss the regulation of quality and investment, and the role of spot markets in power generation and gas production.

Price-cap regulation, which has been applied across the utilities, has had strong effects on productive efficiency and is generally thought of as a successful innovation. It has been copied worldwide. An index of prices is allowed to grow by at most the general increase in consumer prices less a predetermined X factor. The X factor thus determines the real reduction in prices. Price caps are characterized by: (i) long periods (four or five years in the UK, but less in the USA) between formal price reviews; (ii) a commitment by the regulator to avoid resetting the price cap during the lag between reviews in spite of new information being available during that period; (iii) some flexibility to choose relative prices. The idea is to provide incentives for the firm to cut its costs before the next price review occurs. These incentives exist because cost cutting does not entail immediate price reductions. The most obvious lesson from the study of price caps is that the design of regulatory mechanisms should take careful account of the incentives that the regulated agent faces under each proposed scheme. In particular it might be harmful to use a form of regulation that is optimal when information is symmetric when there are information asymmetries in fact.

The two main assumptions necessary for a price cap to be optimal are that the firm has better knowledge than the regulator about its cost-reduction opportunities and that it is

not too costly to society to allow the firm to make supernormal profits or rents. Without these assumptions prices should be set to just cover the cost of service, i.e. there should be rate-of-return regulation. For example if a firm has exhausted its ability to cut costs, and the regulator knows this, then there is no need to give the firm incentives to cut costs further, since the provision of incentives is costly. One hypothesis is that it is plausible that the ability to undertake cost reduction is greatest at the start of the period of regulatory reform and thereafter, as opportunities to cut costs are exhausted, the regulatory scheme might evolve to one based on rate-of-return. A counter-argument, however, is that a regulator is trying, albeit imperfectly, to mimic the pressure that a firm in a competitive market would face and that these pressures are unlikely to change over time. In other words there might always be opportunities for cost reduction. Price caps in the UK were initially relatively lax, and the resulting high levels of rents caused regulators to tighten the caps over time. More recently, though, caps have become less tough (except in telecoms), perhaps indicating that opportunities for cost-reduction in the other industries are lower than before.

The purest form of price cap would fix prices once and for all to maximize the incentives for cost reduction. In practice no government will allow this to happen, and no firm would believe that it would be allowed to get away with rents that last in perpetuity. In other words such a regime would be incredible. This is why formal price reviews after fixed periods are built into the system. Interestingly in the UK the length of time between formal price reviews – four or five years – is the same as the usual length of a Parliamentary term, and the length of time that price caps in the USA last (typically two years) is the same as the gap between Congressional elections. At a price review the regulator takes into account actual and predicted cost changes, and fixes the next set of X values so that a firm that was as efficient as expected would earn a normal profit. Thus all of the factors that come into a review under rate-of-return regulation (cost of capital, asset valuation, efficiency analysis) are also an integral part of a price cap review. The critical factor that provides extra incentives for cost reduction under a price cap is not the way that price limits are set but the fact that prices in the future are predetermined for a relatively long time. If the regulator, or the government, has discretion to intervene in pricing in between formal reviews then the incentives in the scheme are reduced. There was some criticism in the early 1990s that the water regulator was using his discretion to cut prices below the levels that had been

previously agreed. In the mid-1990s there was much discussion in the UK about moving towards formal schemes where profits are shared more rapidly with customers and thus the incentives for efficiency are reduced. In the end the regime of price capping remained in place, but we should always remember that there is a trade-off between providing incentives (which requires a long gap between reviews) and the credibility of the system (which might require shorter lags).

Two procedures that improve the regulator's trade-off between providing incentives for efficiency and reducing rents are yardstick competition and franchising. Yardstick competition has been used in the UK to help to determine price levels for the regional monopolies in water supply and electricity distribution. The existence of similar firms who face cost shocks that are correlated effectively reduces the information asymmetry. In theory it is possible for the regulator to use yardstick competition to maximize efficiency and eliminate rents when the cost shocks are perfectly correlated (Shleifer, 1985). In practice yardstick competition has not proved to be straightforward. The linear programming technique of Data Envelopment Analysis (DEA) and the econometric method of stochastic frontier analysis have both been used by regulators to develop comparative information about cost functions, but the precise way in which comparative cost information is used in price reviews remains unclear and one suspects that comparisons are to some extent subjective.

Franchising provides, in theory, a similar way of improving the trade-off. If there is competition for the market then the rents will be competed away while the incentives for efficiency will remain. The UK has made little use of franchising within the major public utilities, though many local public services, such as rubbish collection, are franchised out. This is probably because the privatised firms have had large investment requirements and franchising is easier to apply to the management of assets than to their expansion. The main example of franchising in the UK is the train operating companies. In developing countries, though, franchising and long-term concessions are much used. The nature of asset markets will partly determine whether full asset sales are feasible, as in the UK. If domestic financial markets are thin then concessions and franchising are probably the only way forward.

The discussion so far has been about price levels and how prices should change when costs change. Another issue in regulation concerns relative prices when a firm operates

in several markets or produces many products. The problem of access pricing is an example. This applies when a firm with a monopoly of network provision supplies a product or service in competition with rivals who must use the integrated firm's network as an input. Examples are interconnection between BT and its rivals, where the latter need access to BT's local network to complete calls, and suppliers of electricity who must use the wires of a company that both distributes and supplies power. Under some restrictive conditions (see Armstrong *et al.*, 1996) the optimal access price is given by the difference between the network firm's retail price and its cost in the retail segment, so the Efficient Component Pricing Rule of Baumol (1983) and Willig (1979) applies. With this rule a rival contributes just as much per unit of sales to the network firm's fixed costs as the network firm's own retail business would, and the network firm is thus fully compensated for lost retail profits through the access price. But if conditions are different this rule needs to be adapted, for example to allow for product differentiation, the possibility of network bypass, and the fact that the network firm's retail price might be incorrectly set. The main lesson to be learned here is that access pricing is difficult. If the access price is too low then there might be excessive entry of rivals – some argue that Mercury was inefficiently subsidized by the telecommunications regulator's 1985 that interconnection should be at short-run marginal cost. If the access price is too high then there will be too little competition and potential cost efficiency and product differentiation benefits from the entry of rivals will be lost. We shall return to the issue of access pricing in Section 3, where it is argued that the difficulty of access price regulation provides a rationale for vertical separation.

When entry exists or is threatened in retail markets, and the dominant firm faces a retail price cap there is a role for regulation of price levels to prevent predation. This issue applied in telecoms where BT faced a price cap based on an index of long-distance and local call charges. Without any further constraints BT was able to cut prices in the long-distance market and recoup the revenue through higher prices in the captive local markets. Overtime the industry regulator has recognized the danger of predation inherent in an average price control and has imposed subsidiary rules that restrict prices to be in between incremental and stand-alone costs (Vickers, 1997). Again the interaction of competition and regulation requires careful regulation and constant vigilance on the part of the regulatory authorities.

Of course consumers do not only care about the prices that public utilities charge. Quality of service is also of great importance, and this is often linked to capacity. Early on BT's regulator realized that the incentives to cut costs provided by a price cap were inducing BT to cut spending on quality enhancement and maintenance, and quality regulation were quickly imposed. In all the industries various aspects of quality of service are measured and prices are cut at the next review if quality targets are not met. The water industry was privatized so that the private sector could finance a large programme of investment required by the European Union to improve drinking water quality and the treatment of wastewater. Real prices have risen accordingly. Again in this industry the regulator has to take a detailed look at investment plans to check that quality targets will be met, but without encouraging gold-plating. Although the water regulator emphasizes that it is regulation of outputs that matter rather than inputs, it is inevitable that this regime starts resemble the type of detailed intervention in enterprises' decisions that was characteristic of both the situation before privatization and a rate-of-return regime. Perhaps, though a virtue can be made of this. A standard fault of rate-of-return regulation is thought to be the incentive to over-invest in capital. When quality is endogenous, however, extra capital investment which produces higher quality and reliability levels might be just what is required. Again the general less is that the addition of quality targets, and maybe investment targets, to the basic framework of price regulation makes things much more complicated than was initially anticipated (Littlechild, 1983).

Spot markets have been created in the electricity and gas sectors in the UK. California, Chile and Norway have undertaken similar reforms. Markets for instant delivery are necessary to ensure that energy systems balance. In electricity in particular there must be close coordination between the transmission company and the power generators. The first example of such a spot market in the UK was the Electricity Pool, a centralized arrangement run by the transmission company that determined which generators were called on to run. Rival generators announced their capacities and prices for each period of the next day, and the Pool constructed a supply curve. The price paid to all generators was the bid of the marginal producer. Generators also received "capacity payments" which were higher the more likely it was that there would be a power failure and the higher the estimated cost that customers would bear in such an event. With sufficient competition price in the daily auction

would equal marginal cost and the capacity payments would provide long-run incentives for investment.

In practice the Pool worked less efficiently. With only a few bidders, each with a portfolio of plants with different marginal costs, the temptation was to restrict capacity or to raise the price of their marginal plants. This reduced the likelihood that the marginal plants were called on to generate, but increased the profits of their base-load plants (i.e. ones with lower marginal costs). In an early analysis of the operation of the Pool, Green and Newbery (1992) modelled the game as one where the competing generators offered supply functions, and showed that the duopolistic structure of the generating market was likely to lead to inefficient outcomes.

In early 2001 new trading arrangements in electricity were introduced. These are designed to promote more flexibility by allowing trading to take place outside the Pool, by encouraging the growth of futures and forward markets and by allowing demand-side bidding, while maintaining a short-run balancing market. In the balancing market bidders will receive the amount of their bid rather than the price of the marginal bidder. Changing the rules of the game will alter bidding behaviour and the hope of the regulator is that collusion will be less feasible than under the old Pool with its very transparent market price. It is too early to say what effect these reforms have had. The point to note here is that, as usual, firms will act in a strategic manner when faced with a given set of incentives, and that it is competition that is critical for economic efficiency.

3. Industry structure and entry conditions

BT and the water companies were not restructured vertically at privatization but the electricity and railway industries were, and British Gas broke itself up a decade after privatization. A standard argument is that regulation is easier when the natural monopoly parts of a firm are separated from the competitive parts. In the usual case the natural monopoly business is a network that rivals must use in order to reach final customers, as in the access pricing case considered in Section 2, so vertical separation is called for. If the regulator can fix access prices appropriately, however, the outcome with vertical separation would be the same as with vertical separation.

The difficulty, of course, is that with vertical integration the regulator is likely to suffer from a severe information asymmetry, and access pricing will not be straightforward. The firm would have every incentive to claim that its marginal retail cost is lower than the true value so that the allowed access charge is higher than necessary. Competition would thus be restricted and customer choice would be limited while the network firm would earn unnecessary information rents. Vertical separation would reduce the regulator's information asymmetry and thus allow efficient and non-discriminatory pricing of network services.

The trade-off is between improving regulation of the network through vertical separation that bans the network firm from also supplying retail services, and the loss of economies of scope and the existence of break-up costs that structural separation might entail. Optimal solutions are likely to differ across industries and across countries. The UK solution of vertical separation for energy industries seems sensible. Competitive entry was allowed in electricity and gas from the early 1980s, but with the incumbents being vertically integrated no competitors entered until a decade later. Competition in gas and electricity supply, especially for large industrial and commercial customers, is now intense. The key has been to separate the transmission grid from the power stations in electricity and in gas to separate transportation from supply. Note, though, that the electricity industry in Scotland has remained vertically integrated and given the smaller size of the market in Scotland than in England and Wales this is appropriate.

In telecommunications vertical separation might not be worthwhile given the rapid changes in technology. In the USA the 1982 solution of structural separation between long-distance and local services is being unwound after the Telecommunications Act of 1996, and in continental Europe the dominant players have remained integrated. BT was left alone, though it has to keep accounts separate between its retail and network businesses. There is now the possibility that BT will divest itself of the local loop of its own accord, and this might ease the problem of allowing rival broadband Internet service suppliers access to the local loop. As far as the railway industry is concerned it can be argued that vertical separation has gone too far, and the crisis that this industry has been in over the last year goes some way to support this claim. It is not clear that in a complex and time-dependent industry such as railways the most efficient form of organization involves explicit contracts between each

stage of production, especially when there is little genuine product market competition between train operators.

A final lesson to be learned from the UK experience of vertical separation is that decisions about separation are much better made at the time of privatization (as with electricity) than after the asset sale (as was the case with gas). Once assets have been sold it is much more difficult to restructure. It took eleven years of regulatory intervention for the gas industry to move to the structure that many argued was appropriate at the time of privatization. At privatization shareholders acquire property rights, and subsequent regulatory action to alter the structure can be thought of as a breach of an implicit contract made at the time of the sale.

Two other aspects of industry structure have also been of concern to regulators. In the electricity industry the UK Government decided to alter horizontal structure at the generation stage. The Central Electricity Generating Board was split into three companies, a nuclear company with 20 percent of capacity, and two fossil-fuel generators with 50 percent and 30 percent of capacity. The nuclear company played no role in determining prices in the Electricity Pool since it bid zero to ensure that its plants always ran, so effectively a duopoly was created. Naturally this created excellent conditions for collusion. The number of players was low, the product is homogenous, frequent purchases are made by customers, and the firms had spare capacities so prices could be driven down in the punishment phase. Since privatization in 1990/91 the regulatory authorities have had to intervene constantly in the generation market to create a more competitive structure. Again horizontal separation should create enough independent competitors to avoid collusion, and should take place at the time of privatization.

The ability to apply yardstick competition is dependent on regional separation of firms. In the water and electricity distribution industries this regional separation existed in any case before privatization. There might be a trade-off between keeping firms regionally separated to provide comparative information and allowing them to merge to gain the benefits of economies of scale and scope. In the water industry all non-trivial mergers have to be referred to the Competition Commission, which must take account of the impact of a proposed merger on the regulator's ability to make comparisons.

The final aspect of structural regulation is the licensing policy of the regulatory authorities. Should the regulator allow free entry or should the rate of entry be controlled? Initially in telecommunications UK regulators opted for slow liberalization. Only one competitor, Mercury, was licensed to compete with BT, and simple resale of capacity leased from BT at wholesale rates was not allowed. Following a review of the duopoly policy in 1991 entry was fully liberalized. In electricity and gas entry into the market for supply to large customers was liberalized early on in the process but was not effective initially because of the lack of attention to access prices and conditions. Now, though, there is full retail competition for all customers. Although it is too early to say what equilibrium market structures will look like in energy supply, it seems safe to conclude that full competition should be allowed as early as technically feasible in market segments where there are no natural monopoly conditions.

4. Regulatory Institutions

Regulation of network industries is complex and requires appropriate resources and institutions. Since large sunk costs are involved and asset lives are long it is especially important that the regulatory regime has credibility. Private sector firms will be reluctant to invest if regulators or governments have discretion to change the terms of implicit or explicit contracts after capital has been sunk, since such discretion might be used to deny firms returns on existing investments. Levy and Spiller (1996) present a useful analysis of both the trade-off between allowing flexibility and achieving commitment in regulation, and of the role of a country's institutional endowment in fostering credible regulation.

The UK has a favourable institutional endowment for regulation. A crucial factor has been the independence of the regulators and the consequent lack of day-to-day involvement of government departments in utility pricing decisions. The argument for independent regulators is essentially the same as that for independent central banks and depends on a certain distrust of politicians. Regulators for each sector were established at privatization, with duties laid down in the relevant Acts of Parliament. They are accountable to Parliament and must act in accordance with their legal duties but these are specified in very general terms so in practice there is substantial autonomy. The regulatory agencies are financed through licence fees paid by the regulated firms themselves. Decisions by the regulator can

be appealed to the Competition Commission, the main competition policy authority in the UK, and ultimately to the courts. As markets and industries have developed regulators have been merged to enhance coordination – for example the separate gas and electricity regulators were merged into OFGEM, the Office of Gas and Electricity Markets and telecommunications regulation is to be merged with broadcasting. There is a good supply of appropriate professionals to staff the regulatory agencies.

It should be remarked that the UK's position is perhaps unusual, and other countries with less favourable endowments are unlikely to be able to mimic the same regulatory institutions. If governments are unstable or the judiciary is not independent then the credibility that is necessary for a favourable climate for private investment is likely to be absent.

5. Conclusion

In this paper I have reviewed the lessons to be learned from the UK's experience, as a pioneer, of regulatory reform. A general theme has been that regulation is complicated and necessarily imperfect, although structural separation can ease the burden on regulators somewhat. Competition, where feasible, offers the best solution to the regulatory problem, but there remains the problem of regulating networks that retain natural monopoly characteristics. Policymakers should be wary, though, of importing wholesale the methods and institutions that are used in the UK.

References

- Armstrong, M., S. Cowan, and J. Vickers, 1994, *Regulatory Reform: Economic Analysis and British Experience*, Cambridge, MA: MIT Press.
- Armstrong, M., C. Doyle, and J. Vickers, 1996, 'The Access Pricing Problem: A Synthesis', *Journal of Industrial Economics*, 44, 131-50.
- Baumol, W., 1983, 'Some Subtle Issues in Railway Regulation', *International Journal of Transport Economics*, 10, 341-55.
- Cowan, S., 2001, "Developments in Regulatory Principles: The UK Experience", in Henry, C., M. Matheu and Alain Jeunemaître, eds., *Regulation of Network Utilities: The European Experience*, Oxford: Oxford University Press, 39-53. Available at <http://www.oup.co.uk/pdf/0-19-924415-4.pdf>.
- Green, R. and Newbery, D., 1992, 'Competition in the British Electricity Spot Market', *Journal of Political Economy*, 100, 1089-99.
- Helm, D. and T. Jenkinson, eds., 1998, *Competition in Regulated Industries*, Oxford, Oxford University Press.
- Kay, J. and Vickers, J., 1988, 'Regulatory Reform in Britain', *Economic Policy*, 7, 285-343.
- Levy, B. and Spiller, P.T., 1996, 'A Framework for Resolving the Regulatory Problem', in Levy, B. and Spiller, P.T. (eds.), *Regulation, Institutions and Commitment*, Cambridge: Cambridge University Press.
- Littlechild S.C., 1983, *Regulation of British Telecommunications' Profitability*, London: Department of Industry, HMSO.
- Newbery, D., 1999, *Privatization, Restructuring, and Regulation of Network Utilities*, Cambridge, MA, MIT Press.
- Shleifer A., 1985, 'A Theory of Yardstick Competition', *RAND Journal of Economics*, 16, 319-27.
- Vickers, J.S., 1997, 'Regulation, competition and the structure of prices', *Oxford Review of Economic Policy*, 13, 15-26.
- Willig, R., 1979, 'The Theory of Network Access Pricing', in H.M. Trebing (ed.), *Issues in Public Utility Regulation*, Michigan State University Public Utilities Papers.