Competition in Regulated Industries: Some Reflections

Simon Cowan

Liberalisation, privatisation and independent regulation of utilities were popularised by the UK and Chile in the 1980s. By then the United States already had long experience of private provision and public regulation of utility services, and independent power production had been mandated by the Public Utilities Regulatory Policies Act 1978 (PURPA). Many countries have subsequently adopted some or all of these policies. Competition has been introduced in telecommunications, in electricity production and in the supply of electricity and natural gas to large users. Franchise competition exists in water supply and in some rail systems. Yardstick or comparative competition is used by regulators in industries where there are regional monopolies with national regulation. At the same time direct competition in the provision of network services has usually not been on the agenda, except in the telecommunications industry. Whether there should be direct product market competition in retail electricity and gas supply, in water and sanitation, in rail and post are all controversial issues.

The main message of the paper is that competition can and should be introduced in many parts of utilities’ operations, but that competition is not a panacea and it requires supporting regulation and institutions if it is to work well. Introducing competition into a regulated and monopolised industry requires both imagination and care, and often entails costs which need to be balanced against the expected benefits of competition. It is important for policymakers to remember that competition is not an objective but is an instrument for achieving ultimate goals.

The structure of the paper is as follows. The next section covers the general arguments in favour of competition. The costs of having various types of competition in regulated industries are then covered. This is followed by discussion of the parts of utilities where competition generally works well and where competition is of doubtful value. The interaction between regulation and competition is then considered. The penultimate section assesses the experience of competition in Californian electricity and a number of industries in the UK and conclusions are drawn.

The General Arguments for Competition

The textbook argument for competition is that it is a good instrument for achieving efficiency. In particular the argument for competition is based on the First
Fundamental Theorem of Welfare Economics — see Varian (2003) for a simple
treatment, or Varian (1992) for a more rigorous analysis of the theorems of
welfare economics. Two types of efficiency may usefully be distinguished:
allocative and productive. Allocative efficiency means that resources are allocated
across sectors of the economy in such a way that no-one can be made better off
without making someone else worse-off. In a market economy this requires prices
to equal marginal costs, which is achieved if all markets are perfectly competitive
and there are no missing markets or externalities. Productive efficiency (which is
itself a necessary condition for allocative efficiency) implies that costs are
minimised for given levels of output. Again competition ensures this, because
firms that have excessive costs are driven out of the market or forced to adapt.

Theorists of regulation often suggest that an objective of policy should be the
minimisation of excess profits (see, for example, Baron and Myerson, 1982; and
Baron, 1989). In other words a dollar in the hands of the firm or its shareholders
is worth less than a dollar in the hands of the consumer — unlike the treatment in
standard welfare economics. Fortunately, however, competition ensures in any
case that this outcome is achieved.

There are important additional benefits of competition. Competition can spur
innovation and provide better products and production methods. Competition can
promote choice and variety. Competition encourages radical thinking about cost
structures that might not have been thought of without it. Many forms of
outsourcing, for example in airlines, have probably been considered and
implemented because of the incentives provided by competitive pressure in the
product market. Finally competition might reduce the need for regulation, which
is both expensive and inherently imperfect.

The Costs of Competition in Regulated Industries

What are the costs of having competition in regulated industries? Competition
sometimes entails both set-up and transactions costs that would not be incurred in
a monopoly system. To be concrete consider an electricity system that has been
monopolised. Suppose that the network parts of the business, transmission and
distribution, remain as monopolies, but that both electricity generation and retail
supply are liberalised and competition is introduced. To make competition
feasible a system for balancing electricity supply and demand needs to be
established if there is upstream competition, and a system for switching customers
has to be installed if there is also retail competition. In England and Wales a spot
market for electricity generation was established in 1990 in place of the old
command-and-control ‘merit order’, which determined which generating stations
were used when. The move to full retail competition required an IT system that
was expected to cost £726m (US$1200m) over the first five years (see Newbery,
1999:228). At the time aggregate annual supply costs as a whole were only of the
order of £600m (US$1000m) so this represented a substantial increase in the cost
base. In the competitive electricity markets of England and Wales the average
cost to the supply companies of switching a customer is £30 (OFGEM website).
The customer also bears switching costs if she or he changes supplier. If the customer is required to wait at home to have the meter read on the day that the switch-over happens there is an opportunity cost of the time taken. In the UK in the early days of retail energy competition some vulnerable customers were subjected to dubious high-pressure sales tactics on the doorstep, which in some cases were fraudulent. More innocent mistakes were made with customers who found that they were paying both the new and the old supplier simultaneously, and unwinding these mistakes took time and effort. Sometimes switching costs that consumers bear can be reduced by some simple mechanisms. Price comparison services operating over the Internet let consumers know instantly which retail supplier offers the best deal. In telecommunications number portability reduces the cost (to the customer) of switching supplier. Carrier pre-selection allows the customer easily to use one supplier for, say, broadband access and local calls, and another for international calls.

While competition may reduce the need for regulation, sometimes the opposite is true and competition requires additional regulation. This may take the form of access price regulation. Suppose there is a monopoly running a vertically integrated electricity system. Regulation of the price the final customer pays is required. Now let the incumbent firm remain vertically integrated and suppose that competition is allowed either in generation or in retail supply, or in both. Regulation becomes more complicated: typically both the incumbent’s retail price and its access price will be regulated. This is not the place for a full discussion of the complexities of access pricing, but it suffices to note that regulation of both the access price and the retail price is difficult when the firm is vertically integrated. Integration is likely to give the firm an information advantage relative to the regulator, particularly about the allocation of costs. Allowing competition while retaining vertical integration requires costly regulation. This can be alleviated to some extent by a policy of vertical separation, which will allow competition and reduce the regulatory burden. See Cowan (2001) for a simple model of the trade-offs involved. Vertical separation makes the regulatory problem easier, because it gives the regulator better information, but it also prevents the joint setting of the access and retail prices, which may be more efficient. In addition vertical separation implies that reorganisation costs are incurred and may entail a loss of economies of scope.

Another reason why competition and regulation are not perfect substitutes is that the introduction of competition usually requires strong competition or anti-trust policies. This is particularly so in a utility were there is likely to be a dominant incumbent which might act as a price leader (or price in a predatory way). Similarly horizontal separation of an incumbent into only two or three components might not promote competitive behaviour and instead may facilitate collusion. Thus elimination of sector-specific regulation of, say, retail prices does not mean that regulation is abandoned, and instead extra attention from the general competition authorities is required.

Competition might make it infeasible to retain price structures that are socially valuable. By forcing prices to match costs competition prevents the use
of cross-subsidies which might be thought to be desirable. The classic example is uniform geographical pricing that is used in postal services (combined with a universal service obligation). If new firms are allowed to cherry-pick markets then the original price structure may not be sustainable. In South Africa reform of utilities is constrained by requirements on state-owned utilities to set tariffs with large elements of cross-subsidy. Again this is not a knock-down argument against competition. Either making obligations symmetric, or requiring firms that do not have extra obligations to contribute to the costs that the obligated incumbent bears, will serve to reduce or eliminate this problem.

**Areas where Competition Tends to Work Well**

Competition tends to work well in upstream production of natural gas and electricity, in downstream supply of gas and electricity to large and medium-sized customers and in the retail parts of telecommunications. In addition franchise competition can work when there are local geographical monopolies, such as municipal water supply. As usual, though, there are exceptions and the details of each case matter.

In electricity generation several types of competition are feasible. At a minimum a vertically integrated grid can be required to buy power from independent power producers (IPPs) on reasonable terms, as the 1978 Act (PURPA) required in the USA. Typically such IPPs are small in relation to the overall market size so there is no issue about natural monopoly in generation. In practice the crucial question is whether the terms are reasonable or not. In England and Wales liberalisation to allow independent power producers to supply was enacted in 1983, when the generation industry was still monopolised, integrated with the national grid and state-owned. Because grid access terms were not regulated this liberalisation was entirely ineffective (see Armstrong, Cowan and Vickers, 1994:291). Much depends also on the objectives of the power purchaser. A profit-maximizing private-sector firm would presumably not want to discriminate against efficient independent power producers because they represent a new way to minimise costs. A state-owned firm might have other incentives.

The alternative form of competition in generation is to establish a real-time spot market in wholesale electricity (or indeed in gas), backed up by forward or futures markets. The Pool in England and Wales (which was replaced by the New Electricity Trading Arrangements in 2001), Nord Pool in Scandinavia and the Power Exchange in California are examples. This requires sufficient horizontal competition, which may in turn mean that an existing generator has to be broken up. Establishing a spot market requires that rules are drawn up for bidding, for payment mechanisms and for reserve capacity so that the market works in an orderly fashion and continuous electrical equilibrium is guaranteed. In practice the major problem with the spot markets in England and Wales, and in California, (though not in Nord Pool, which has many more players) is the exercise of market power by a handful of generators who can determine the spot price. For example in England and Wales the rules of the pool allowed the two dominant generators to
declare some capacity unavailable, which then raised the equilibrium spot price and increased the profitability of infra-marginal capacity that was declared available. In a much more competitive market it would not pay to restrict supply in this way. A well-designed spot market requires regulatory monitoring to ensure that market power is not being abused. In natural gas the balancing market in the UK has been less problematic than the electricity market, partly because there are many more producers of gas than power generators. In conclusion it seems that direct competition in generation or production can work if the spot market rules are carefully designed and there is sufficient competition that the abuse of market power is not an issue. In some isolated systems, for example small Caribbean islands with a handful of power stations, competition in generation is hardly going to be worthwhile, and the system would be a natural monopoly.

Often the impetus for reform of electricity supply systems is the demand from large industrial and commercial users for cheaper electricity. Integrated monopoly systems may subsidise domestic customers at the expense of industrial customers, so competition will unwind these cross-subsidies. Alternatively the integrated monopoly may simply earn excess profits on all sales. Large customers sometimes have the option of self-supply. Competition in supplying such customers usually needs a competitive spot market to exist and requires users to have real-time meters. Very often the suppliers will be integrated with generators, but this is not essential.

In many countries the first state-owned network enterprises to be sold off are telecommunications operators. Changes in technology, shifts in demand and recognition of the critical importance of interconnection (or access pricing) have meant that competition is both feasible and desirable in many market segments. In the 1980s the dominant theme of reform in telecommunications was that competition was desirable in long-distance calls because of new technologies and demand growth, but that the ‘local loop’ at either end remained naturally monopolistic. In the UK in the early 1990s policy was revised to encourage entry at the local-loop end and it was expected that cable TV companies would provide this competition. The Telecommunications Reform Act 1996 in the USA also promoted local competition. In the end this facilities-based competition has not proved to be significant. But competition in the provision of retail services over the networks can be, and has been in many cases, successful. The terms of interconnection are critical here, and there is continual pressure for incumbent operators who provide both retail and network services to separate the activities (or for the regulators to split them). In the UK British Telecom (BT) is very dominant in network services but also has a large share of retail services nearly twenty years after competition first started. While retail competition in indirect access telephony services is now coming on stream, competition in the provision of broadband access has probably been constrained by BT’s obstruction of local-loop unbundling (which requires BT to allow broadband providers access to local exchanges).

In the water sector there are few opportunities for direct product market competition. The fact that water businesses are locally organised provides,
however, the opportunity for franchise competition. In France and in many developing and middle-income countries, municipalities franchise out water provision. Competition for the market can reduce excess rents. Contracts, especially long-term ones such as concessions with investment requirements, will typically require on-going regulation because it is usually impossible to write complete contracts covering every contingency. The benefits of franchising are greatest when collusion between bidders is deterred. Regulatory capture is always a danger when municipalities are involved in awarding and monitoring long-term contracts.

Areas where Competition Is of Doubtful Value

In some cases competition is of little value and there is no point in promoting it. In other cases the case for or against competition is less clear. Direct product market competition in water hardly exists, and there are good reasons for this. In the absence of national or even regional water grids there is little prospect for competition via common carriage. Although in the UK such possibilities are being explored, it is fair to say that the policy-makers are being very circumspect about the application of competition. Some methods of introducing competition into water in a small way are feasible and will be discussed in the penultimate section.

Full retail competition for domestic customers of both gas and electricity started in the UK towards the end of the 1990s. While there are many obvious benefits of competition in the markets for larger customers the case for retail competition is not clear (see Giulietti, Waddams Price and Waterson, 2005 for an analysis of UK experience). In the absence of real-time meters multiple suppliers of electricity have to rely on load-profiling of customers to ensure that payments balance usage on average. In a second-best world, where customers face different prices depending on whether they have a real-time meter or not, competition is not axiomatically beneficial. In addition there are the set-up, transactions and switching costs described earlier which are incurred with competition but which are avoided with monopoly supply. While it is true that the introduction of competition lead to price reductions for customers who switched, and may have had some effect on price levels for those who have not switched, the fact that the cost base increased at the same time to facilitate competition makes the effect on social welfare ambiguous. The welfare analysis is very similar to the (reverse of the) classic textbook model of a merger than generates cost reductions but more market power. Competition in retail supply raises costs but reduces profit margins.

Direct competition in network provision is usually absent for the good reason that most networks constitute natural monopolies and network duplication would be wasteful. Telecommunications is the main exception, and even here competition in the local loop is limited. For voice services direct facilities-based competition is essentially a by-product of cable TV provision. In some countries, though, local-loop unbundling for broadband provision has advanced significantly.
In gas, electricity, water, rail and post there is no prospect of stand-alone duplication of the local distribution networks. Depending on the size of the market, there can be competition in the provision of long-distance transportation, for example in natural gas in the USA. Of course even when a local distribution network is a natural monopoly there are some competitive possibilities. Franchise competition has been mentioned already. Yardstick competition, or comparative competition, can be used by a regulator as a surrogate for actual competition.

**Regulation and Competition**

The links between regulation and competition are complex, case-specific and can go either way. Sometimes competition will substitute for regulation. Most obviously this is true in the electricity and gas industries, where regulation of final prices for large customers could be dropped as soon as competition was introduced. In the UK price regulation for domestic customers has also been dropped in those industries, and retail price regulation in telecommunications has over time been reduced in scope. Since regulation is costly the fact that regulation can be dropped when competition has become effective is valuable. In two other ways, that have already been mentioned, the regulatory burden increases when competition comes in. Access price regulation is necessary when there is competition in the provision of services over the network, and if price regulation is abandoned there needs to be increased focus on general competition policy either by an industry-specific regulator or by the general competition authorities.

The UK telecoms regulator, OFTEL, had to act as a competition authority in 1999 in response to an incident of predatory pricing. Scoot plc entered the new market for operator-assisted classified directory advertising, which allowed end-users to phone a single number and be put directly through to a particular business, such as a local plumber or a restaurant. The business model involved end-users paying nothing for their calls, while businesses paid for listings and for each call put through to them. Two weeks after Scoot was the first entrant into this market, BT announced its own service, but with no charges for businesses to be listed and connected for the first year. The industry regulator deemed this to be predatory and had to intervene quickly to prevent BT from building up an unassailable position in this market.

A key issue is the role of regulation in the transition to full competition. A natural policy is to preserve retail price regulation of the dominant firm until competition is firmly established. This requires careful judgement about the optimal level and time path of the price cap. A price cap that is generous to the dominant firm is also generous to entrants and thus may promote competition, but at the same time consumer prices may not fall much. A tough cap will protect consumers but may not encourage much entry.

Often the policies used during the transition to competition are counter-productive (Armstrong and Sappington, 2005). Regulators and policymakers have occasionally acted in a way that suggests that they want to protect competitors rather than to promote efficiency. In the UK telecommunications sector an early
policy was to allow only one new entrant (Mercury) into the long-distance market (and some local markets). In addition the determination in 1985 of interconnection or access charges for Mercury gave the latter access to BT’s networks at what was essentially marginal cost. This did not compensate BT for the opportunity cost of providing access (which would include its lost profits), as the Efficient Component Pricing Rule (ECPR) of Baumol (1983) and Willig (1979) requires. An access price below the level indicated by the ECPR will encourage inefficient entry.

Market share targets for incumbents have been used by regulators in the UK in both the electricity and gas industries. As Armstrong and Sappington (2005: Section 6.3) point out, such policies have two problems. First, market share is not generally a good indicator of market power. A firm with a relatively small market share may still have considerable ability to raise prices above competitive levels. Second, market share targets can induce all firms to price less competitively. For an incumbent the simplest way to cut market share to the target level is to raise the price in a market where products are not perfect substitutes. The rival firms will typically then have incentives to increase their prices as well.

Regulations derived from social rather than purely economic considerations may affect the development of competition. Examples are universal service obligations and requirements to price uniformly irrespective of costs. Again these can be carefully designed to ensure that they do not lead to too much new entry (when entrants do not face similar obligations to incumbents) or too little entry (when incumbents use the existence of these obligations as an argument against any new entry).

The Experience with Competition

In this section the experience of introducing competition into a number of regulated industries is described briefly. The section begins with the infamous case of Californian electricity, and then looks at telecommunications, gas, electricity, water, rail and post in the UK.

Electricity in California

California’s electricity reform of the late 1990s has become notorious. Vertical integration between electricity distributors/suppliers and generators was removed. The latter competed (imperfectly it turned out) in a spot wholesale market. All of the standard methods for sharing risks between different stages of electricity production were banned. Distributors were forbidden from integrating again with the generators. Often a good substitute for vertical integration is a long-term contract. In the interests of maintaining room for new entrants to the generation market, though, long-term contracts were only allowed to cover a fraction of each distributor’s obligations. Distributors sold power to final customers at a capped retail price, and had obligations to supply, whatever the level of wholesale price. An unfortunate combination of individually unlikely events then led to the
wholesale price rising well above the capped retail price until the distributors approached bankruptcy. These included a rise in the natural gas price, a dry spell which reduced the amount of water available for hydro-generation, a rise in the demand for electricity for air-conditioning and possible abuse of market power by the electricity generators.

The partial reform of electricity in California failed. What lessons can be learned from this? First, allowance for robust risk-sharing mechanisms needs to be made when designing a system. Closing off all the traditional routes for dealing with the shocks (vertical integration, long-term contracting, retail price flexibility) was a hostage to fortune. Second, while short-run efficiency is valuable in electricity markets another objective is long-run security of supply, which was not obviously promoted by the new system.

Policy in the UK

Competition in UK telephony began in 1986 with a single competitor, Mercury, being licensed to compete with BT, largely in the long-distance market. The change of policy in 1991 to allow a free for all was also based on the idea that there would be substantial facilities-based competition, but this time the entry was expected to be at the local level and come from cable TV companies. Again this has proved disappointing. BT’s share of voice calls in the residential market was 68.5 per cent in 2004, and its share of lines was 64 per cent, with cable companies having 18 per cent and indirect access providers having the remaining 18 per cent. Retail price regulation of BT remains in existence after twenty years of competition, though the scope of the price cap is smaller than once was the case. For broadband provision ‘local-loop unbundling’, the process of allowing rival internet service providers access to BT’s local exchanges to install their own equipment, has been delayed, and the country lags behind other European countries. A recent regulation requiring BT to cut the wholesale charge it makes for local loop unbundling may go some way to speeding up the process. Over time BT has been forced increasingly to separate its network and retail parts.

In natural gas the privatisation of British Gas in 1986 was accompanied by no discernable effort to promote competition. British Gas remained a vertically integrated monopolist that was also the only buyer of gas from producers. Regulation of retail prices was weak and though de jure competition for larger customers was allowed via common carriage there was no competition for five years because the access terms were not regulated. Subsequent reform concentrated both on the vertically integrated structure of British Gas (eventually the company separated itself) and on the regulation of the terms of access to the grid. The competition authorities, as well as the industry-specific regulator, became involved in monitoring the development of competition via common carriage (including at one stage imposing a maximum market share target on British Gas). Since the late 1990s retail competition in gas supply has been rolled out across the UK, and to date 47 per cent of households have switched supplier (some have switched back to British Gas after a period with a rival company).
The reform of the electricity industry in England and Wales to some extent reflected the lessons learned from the mistakes made in gas. There was both vertical and horizontal separation. Ownership of the national transmission grid was taken away from the monopoly generating company and given to the twelve regional distribution and supply companies. Generation itself was split into three asymmetric companies and a power pool was established. The twelve distribution companies remained as regional monopolies in distribution but not in supply. Supply competition was introduced in three stages, with domestic customers being allowed to choose their supplier from 1999.

The electricity pool had good rules for determining prices but not enough competitors. In the early days of competition in generation the regulator intervened regularly and policies included a temporary pool price cap, bans on vertical mergers with distributors and the requirement that the two major generators sell some of their capacity to third parties. The pool was replaced by the New Electricity Trading Arrangements in 2001, which were designed both to increase flexibility and to reduce transparency so that tacit collusion would be less easy to sustain than under the Pool. Competition in the supply of electricity to domestic customers has resulted in some concentration of the market, with six major companies in a market place where once there were twelve. Just over half the customers had switched supplier by 2004.

The water industry has few opportunities for competition, other than through franchising or through yardstick competition. In England and Wales competition for very large customers was allowed in 1996, which lead to the introduction of ‘large-user tariffs’ with substantial quantity discounts. The result was that no customer switched from their local suppliers. Three main types of competition are feasible: inset appointments (where a new supplier takes one particular customer, perhaps supplied with water from a new sources), competition for customers located at the boundary between two suppliers, and competition via common carriage. While there is a handful of examples of inset appointments the other types of competition have not yet started. There seems to be little prospect of widespread competition.

At the time of railway privatisation in 1994 there was a single company, British Rail, which operated train services, owned and serviced trains and owned and ran the tracks, signalling systems and stations. This was replaced with 84 different companies. There was radical vertical and regional separation. Train operation was franchised, with a small amount of direct competition, train ownership was separated from train operation and the tracks, signalling and stations were separated from train services and put into a single network company, Railtrack. The latter contracted out much of its maintenance work and was then bankrupted by the Government after two high-profile fatal accidents which were caused by inadequate maintenance. The vertical disintegration of the system has increased transactions costs as formal contracts between each layer of the system have to be devised and monitored.

The Post Office in the UK remains state-owned, but recently there have been moves to liberalise the market. Competition in bulk mail is currently allowed for
more than 4,000 items, and full-blown competition will be allowed from 2006. Because competitors are unlikely to want to duplicate the distribution network, Royal Mail is required to offer access to rivals for delivery. Royal Mail is also required to maintain a universal postal service. Royal Mail has announced plans to price by size as well as by weight, but is required to price uniformly across the country.

Conclusions

Competition is valuable in promoting efficiency, but care needs to be taken when introducing it, and policymakers should always bear in mind that competition is not an objective but is an instrument. It can be costly to introduce competition, as set-up, transactions and switching costs are incurred that are not present when there is no competition. For competition to be successful it needs nurturing, but not over-protection. Structural separation of vertically integrated incumbents is often helpful in promoting competition, but has its own costs, and access price regulation and general competition policy need to complement the introduction of competition. When access terms have not been subject to regulation early on (as in the gas industry in the UK) the benefits that competition can provide have been delayed.

References


This is a revised version of a paper presented at the Regulatory Forum of the Australian Centre of Regulatory Economics (ACORE), the Australian National University, Canberra, on 17th October 2005. I am very grateful to Flavio Menezes, Velu Ramasamy, Henry Ergas, Paul Paterson and the other participants at the Forum for very helpful comments.