Valuation and costing issues in access pricing with specific applications to telecommunications.

Henry Ergas, University of Auckland, New Zealand.

Normatively, regulation can be viewed as an implicit or explicit contract between the regulatory authority, consumers and the regulated supplier(s). The essence of this contract is that the authority, acting as a consumer agent, commits to setting prices on a basis which recoups the long-run costs of efficient supply, where efficient supply is defined with reference to that which would prevail were the market at issue contestable. This, in turn, implies that the supplier derives a stream of residual income dependent on the costs it incurs relative to the prices which would be established under contestability.

In practice, the regulator cannot know the prices and costs associated with a purely hypothetical contestable market. In these circumstances, a pure cost-plus version of the regulatory contract, such as that involved in rate of return regulation without regulatory lag, would provide little incentive for cost-minimisation by the regulated supplier. Hence, if such incentives are to be provided, the expected residual income to the regulated supplier (and the expected path of prices to consumers) must — if opportunities to increase productivity are taken up — exceed that which would accrue to the supplier in the perfectly contestable counterfactual. The regulatory contract will therefore centre on an agreement between the regulator and the supplier about a residual income formula which balances the allocative efficiency costs of supra-normal margins with the gains in technical and dynamic efficiency which the prospect of supra-normal earnings can bring.

The fact that the regulated output is an input into a more-or-less competitive downstream market does not materially alter this conclusion, although it does make it more difficult to discern the ultimate welfare consequences of particular paths of prices. There is, in particular, no case for manipulating the upstream price so as to artificially encourage downstream entry. It is true that the regulator, were it an omniscient social planner armed with perfect information, might balance incentive effects in the upstream market with competitive effects in the market downstream, and carry out the type of second-best analysis conventionally associated with infant industry protection. However, it is readily shown that the knowledge required to carry out this kind of analysis is very great; and since regulators do not have anywhere near

---

2 Economic analysis shows that welfare losses arise when taxes are imposed in upstream rather than downstream markets — that is, when intermediate inputs, rather than final outputs, are taxed. Even putting aside any consideration of competition effects, this complicates the setting of regulated prices for a supplier of intermediate inputs. (This is taken up again in the discussion below). The point being made in the text, however, is a different one and relates to whether the upstream price should be manipulated in such a way as to artificially encourage downstream entry.
3 See, for example, Damania, 1996.
perfect information, and nor do their proximate and ultimate supervisors, this kind of social engineering, like infant industry protection generally, is a policy that ‘invites ultimate disappointment and (what might be worse than the mortification of non-arrival) misery on the way’

Putting the temptations of social engineering aside, three issues need to be addressed in defining the expectation of residual income to the regulated supplier. These are: the definition and measurement of the residual income stream; the allocation issues associated with joint and common costs; and the basis for evaluating the resulting rate of return in terms of market benchmarks for the cost of capital.

Subsequent sections of this paper deal with each of these. A final section summarises and concludes.

**Asset valuation and the definition and measurement of the residual income stream**

I take it as given that the relevant income concept is that of Hicksian, standard stream income — that is, the maximum amount which could be taken out of the regulated enterprise by the owners in a given period without impairing their ability to take the same amount in all future periods of equal length. Residual income, is in other words, defined in terms of the level of distribution to owners consistent with capital maintenance. This, essentially economic, approach to income determination is consistent with the fundamental requirement of company law (Corporations Law, s. 201(1)) that ‘No dividend shall be payable to the shareholder of any company except out of profits’.

Operationalising the Hicksian approach requires a definition of capital maintenance. Attempts to do so in terms of the maintenance of physical operating capability suffer from serious, in my view fatal, conceptual weaknesses: they are, in particular, of no use when the asset base of the enterprise whose income is being measured is changing. Rather, as Hicks himself emphasises, an economic concept of income can only be given substance in terms of maintaining the value of owners’ equity. Thus defined, the

---

4 Oakeshott, 1996, at p. 31.
5 Scott, 1976, at page 2. The generally accepted ex-post counterpart to Hicks’ definition is ‘the amount the corporation can distribute to the owners of equity in the corporation and be as well-off at the end of the year as at the beginning’; Alexander, 1950, at p. 15.
6 As Lemke notes: ‘When the firm is in the process of changing the nature or composition of its physical capital (operating capability), it is nonsensical to talk about maintaining it’, Lemke, 1982, at p. 320, emphasis in original.
7 Hicks contrasted his own views, which he considered consistently ‘fundist’ (i.e. concerned with maintaining ‘the fund’ advanced by the purchasers of residual income rights), with those of Pigou, who most clearly expressed what Hicks termed the ‘materialist’ approach to capital maintenance. See Hicks, 1974. The essence of the materialist view is that changes in asset prices, to the extent to which they do not force changes in physical operating capability or alter the physical service potential of assets, are not treated as forming part of current income but rather flow directly through to an adjustment reserve in the equity section of the balance sheet. Defences of the physical maintenance concept can be found in Revsine, 1973, and — in a much modified and qualified form — in Scott, 1989. Strong, and in my view convincing, criticisms are Sterling, 1982, at pp 3–58 and Carsberg, 1982, at pp 58–74.
capital maintenance standard determines both the measurement of the income flow and the resulting changes in the balance sheet.

In particular, since the decision to put an asset into operation must take account of anticipated changes in the asset’s value, these changes are properly charged as depreciation against income from the asset in the period in which they are expected to occur. Economic income, as Hicks put it, is then obtained by defining ‘depreciation of the original stock of capital as the difference between the total value of the goods comprising that original stock as it is at the end of the year \( (C_1) \) and the value \( (C_0) \) which would have been put upon the initial stock at the beginning of the year if the events of the year had been correctly foreseen, including among those events the capital value \( C_1 \) at the end of the year’. The Profit and Loss statement should, in other words, record a charge against operating income for correctly anticipated changes in asset values, leaving a Balance Sheet value of assets which reflects their replacement cost.

Basing the charge against operating income on the period-on-period change in replacement cost raises the issue of how ‘replacement cost’ is to be assessed when the asset which should be put in place at some future date differs from that currently used. In particular, it is sometimes argued that the replacement value should reflect the level of outlays required by the set of techniques yielding a service flow equivalent to those of in-use assets at the lowest monetary charge. This corresponds to valuation on a Modern Engineering Equivalent Replacement Asset (MEERA) basis.

Where in-use assets are still being traded — for example, because a programme of continued investment in those assets is in train — the prices at which those assets are obtained will provide the best reflection of their value to the entity. In particular, those prices will embody both the efficiency gains obtainable from the next generation assets and the adjustment costs the entity would incur in shifting from this generation to the next generation. In contrast, optimised or MEERA prices may overstate the value of

---

8 In terms of Alchian’s approach to the definition of costs, the firm incurs an exposure to the change in asset prices as a result of beginning the production programme at that time rather than postponing it to a future date (when the changes in asset prices will have occurred). The consequences of that change are therefore a charge correctly attributable to consumption in the period leading up to the change in prices. See Alchian, 1977, at pp 273–300.

9 Hicks, 1942, at p. 177, emphasis in original. Hicks repeats this definition in Hicks, 1973, at pages 164–6 and emphasises that it ‘is rigorously forward-looking’. Indeed, he doubts whether any other rigorous definition of depreciation can be given consistent with his definition of income. A similar definition, cast in terms of social income, can be found, along with a useful contrast of capital maintenance concepts, in Triplett, 1996, at pp 93–115.

10 Were prices higher than this amount, the enterprise would simply shift to the next generation. In practice, adjustment costs, and hence the optimal lag in changing from current to next generation technology, are likely to be significant, as is apparent from the extent of technical inefficiency (in the sense of gaps between the production possibility frontier and actual factor usage) in even very competitive markets subject to rapid technological change; see for example, Caves, 1992.
next generation assets by excluding the capital costs involved in shifting from one vintage to the next\textsuperscript{11}.

Moreover, from a practical point of view, when optimised or MEERA values are used to value in-use assets, consistency between the Profit and Loss statement and the Balance Sheet can only be maintained by imputing notional adjustments to operating outlays. In particular, any savings in operating costs that would be achieved using the next generation assets need to be added back to actual (in-use technology) cash operating inflows\textsuperscript{12}, as is obvious when the next generation involves a lower (higher) initial cost but a higher (lower) per-unit cost\textsuperscript{13}. It is, to my mind, questionable whether this can be done to an acceptable level of accuracy in complex systems with rapidly changing technologies\textsuperscript{14}.

In short, (1) assets should be valued at replacement cost, using, whenever possible, entry prices for in-use assets\textsuperscript{15}; these will incorporate the effect of gains in efficiency associated with the arrival of new vintages; and (2) the period charge for depreciation should reflect the anticipated period-on-period change in replacement cost, adjusted for decay and deterioration\textsuperscript{16,17}.

\begin{itemize}
\item See Revsine, 1979, at pp 306–322; Brooks, 1993, at pp 246–257; and Hulten and Wykoff, 1996, at pp 10–23. MEERA or optimised prices will overstate economic value unless the supplier of the next generation assets bears the adjustment costs involved in shifting users of current generation assets to the next generation assets. In contrast, the efficiency gains, netted against the transition costs, set a ceiling on the price of current generation assets. Current prices for the current generation assets will therefore provide a conservative basis for their valuation.
\item This is because the use of ‘optimised’ values involves deducting depreciation on the basis of the next generation asset. A consistent income statement can then only be obtained if current outlays are adjusted to reflect the level which would prevail were that asset in place. Revsine, 1979, at pp 306–322 provides worked examples.
\item In accounting terms, the initial cost is a non-current cost while the per-unit cost is a current cost. Adjusting one without adjusting the other would make little sense.
\item One can but sympathise with the view of the Bell System technical staff who wrote, of optimised asset valuation, that ‘... for a large utility like a telephone company, even the visualisation of a complete property built of the most modern facilities would be a staggering job.’ American Telephone and Telegraph Company, 1957, at p 121.
\item By convention, ‘entry’ prices are the prices at which goods can be purchased, while ‘exit’ prices are the prices at which they can be sold. Consistently with common sense, increments in asset prices should be valued using exit prices (the highest price the firm could obtain for its now more valuable assets) while decrements should be valued at entry prices (the lowest price at which the firm could obtain the now less valuable assets).
\item Decay is the decline in expected useful life; deterioration is the decline in the value of the services achieved at each point in remaining life. (The classic ‘one-hoss-shay’ depreciation of simple growth models arises when assets decay but do not deteriorate before exhaustion (retirement); with a zero discount rate, this gives rise to straight-line economic depreciation.)
\item This is subject to the following conditions: (1) No asset should be valued at more than the net realisable value associated with its flow of services. This condition, although not of practical significance in a growing network, means that in a declining network assets which will ultimately not be replaced should be valued at the higher of economic value and scrap value. (2) For all assets, differences between anticipated and actual changes in replacement values are taken directly to shareholders’ equity.
\end{itemize}
This approach is consistent with observed valuation behaviour in capital markets. Although studies find mixed results with respect to the market response to current cost accounting information, there is growing evidence of the predictive ability of the ‘clean surplus’ equation, which defines corporate income in terms of the sum of this -period earnings and changes in asset values18. Investors, in other words, appear to assess returns on the basis of Hicksian income as defined above, and could be expected to cast a forward-looking contract, be it with a firm’s managers or with its regulators, on that basis19.

It remains to contrast the valuation approach set out above with three others: that based on historical costs; Optimised Deprival Value; and valuation based on exit costs.

Historical costs, as George Stigler put it, ‘have powerful sway over untutored minds’20. Despite this, income flows derived from historical cost accounting will coincide with economic income only under extremely restrictive circumstances. It is possible to manipulate historical cost data so as to make the accounting rate of return on a project equal to its economic rate of return; however, this can only be done by taking account of changes in asset prices over the life-time of the project (and hence no longer corresponds to historical cost accounting as conventionally defined)21. Prices based on historical asset costs, and which reflect historical cost depreciation22, will, as a general matter, bear no relation to those which would prevail in an unregulated market, since decisions in such a market will be made looking to the future, rather than to the past pattern of outlays23.

---


19 The theoretical case for the role of changes in asset prices and of depreciation in valuation is set out in Feltham and Ohlson, 1996.


22 Under historical cost accounting, depreciation is not a valuation process but rather one of allocation of costs previously incurred on the basis of some concept of ‘matching’ the flow of services from those costs (which are then said to ‘expire’) with that of current revenues. This is inevitable as ‘in an historical-cost based accounting system, a valuation approach to depreciation would be inconsistent with the measurement of non-current assets. Adoption of an allocation approach to depreciation is unavoidable given the decision to measure non-current assets at historical cost and to maintain intact money capital.’ Peirson and Ramsey, 1994, at p. 16.

23 This point was apparent to the classical economists. Thus, Wicksteed noted that cost of production ‘in the sense of the historical and irrevocable fact that resources have been directed to this or that special purpose, has no influence on the value of the thing produced. (What affects supply is anticipated cost) in the sense of the alternatives still open which now be relinquished in order to produce this specific article’ since it is this, and only this, which ‘influences the craftsman in determining whether he shall produce it or not’; Wicksteed, 1910, at p. 380. The best elaboration of this point remains Buchanan, 1969, which also contains a detailed discussion of the development of economists’ view of costs.
Two claims are nonetheless made on behalf of historical cost accounting.

The first is that only the use of a historical cost standard can guarantee capital maintenance — i.e. can ensure that the value of investors’ initial outlays is preserved\(^\text{24}\). In principle, regulators, given a free hand, could set prices so as to secure virtually any concept of capital maintenance. From an economic point of view, however, the question is whether the capital maintenance goal selected has any normative justification. The replacement cost approach set out above values assets at current entry prices — and hence reflects the asset valuations consistent with the output prices which would prevail in a contestable market — while achieving financial capital maintenance through the depreciation charge to period income. In contrast, maintenance of historical capital, achieved through a backward-looking depreciation charge, will generate a balance sheet which serves merely as a ‘repository of unamortised costs’\(^\text{25}\), and hence reflects no more than ‘a mélange of variously dated amounts, of qualitatively disparate significance’\(^\text{26}\). It is difficult to find an efficiency justification for setting prices on this basis.

The second, somewhat stronger claim, points to practicality. Historical costs, it is said, are what firms do; they are ‘objective’ and easy to ascertain. In contrast, replacement valuation relies on judgements which are ‘subjective’, outside of firms’ normal experience, and inevitably contentious\(^\text{27}\). Two points need to be made in this respect.

To begin with, even in principle, historical cost accounting involves a broad range of inherently arbitrary judgements and allocations. For example, there is no clear basis in historical cost accounting for determining the choice of the depreciation schedule; rather, this is done with reference to a notion of ‘matching’ (or ‘cost attachment’) which is broadly incapable of falsification\(^\text{28}\). There is, in other words, no analytical basis in historical cost accounting for the capital consumption charge imposed against revenue in any particular period, making the apparent ‘objectivity’ of this charge entirely illusory.

\(24\) Professor Hausman, has argued, for example, that: ‘Even if actual historical network investment decisions were always completely efficient at the time they were made, improvements in technology will always guarantee that a totally new, hypothetical network will have a theoretical lower cost than the actual network in place (or otherwise the older technology could be used in the hypothetical network). Thus, basing cost on the current most efficient technology will impart a downward bias on estimates of actual network costs, causing an economic loss to the LECs which made the historical investment’ Hausman (Affidavit), at note 4 on p. 7.

\(25\) Coase, 1973, at p. 120.

\(26\) Chambers and Wolnizer, 1990, at p. 360.

\(27\) See, for a recent expression of these views, King, 1996.

\(28\) The classic demonstration of the arbitrary nature of these allocations is in Thomas, 1969.
Second, precisely because historical cost accounting is so arbitrary, it is not in fact what firms do or are required to do. Thus, entities subject to the Corporations Law are required to account in conformity with Australian accounting standards. AASB 1010 requires that non-current assets not be shown at above recoverable amount. In conjunction with s. 294(4) of the Corporations Law, this provision requires Directors to review the replacement value of non-current assets relative to their earnings ability, and report accordingly. Consistent with the comprehensive income definition embodied in AASB 1018, AASB 1010 also requires that downward revaluations flow directly to the Income Statement, where they are treated as an expense.

With the number of Australian companies revaluing their assets increasing substantially in recent years, these provisions have been of growing significance. The requirement to maintain accounts on a basis consistent with current asset valuations is even clearer in respect of government business entities and forms a central part of National Performance Monitoring. Viewed together with mark-to-market valuation for current assets, and for particular classes of reporting entities (notably for insurance funds, including those involved in superannuation), it seems reasonable to conclude that in Australian accounting ‘too many instances of the use of market prices as the basis of valuation currently exist for anyone to say that market prices do not constitute sufficient objective evidence’.

Optimal Deprival Value (ODV) valuation, also known as valuation on the basis of ‘value to the owner’, differs from the replacement cost approach set out above in two respects.

To begin with, ODV, at least in its implementations to date, rests on a concept of physical rather than financial capital maintenance, and hence corresponds to the entity, as against proprietorship, approach to financial accounting. For the reasons indicated

---

29 The interaction of AASB 1010 with s. 294(4) is set out in ASC Practice Note 21: Value of Non-Current Assets. See also Perkins, 1996, at 74–75. Note also that AASB 1021 requires annual reviews of likely useful life, and that the effects of resulting changes must be reported in the Profit and Loss statement in accordance with AASB 1018.

30 Presumably on the basis of conservatism, under Australian accounting standards, asset revaluations do not flow to the P&L but rather must go directly to an asset revaluation reserve, which is available for distribution on realisation.


33 Under AASB 1023, these are now required to report all assets at market values on balance sheet date.

34 Godfrey, Hodgson, Holmes and Kam, 1994, at p. 476. The authors go on to note that ‘it may be true that in some cases they are not reliable, but we cannot demand that in all cases a firm must be a direct participant in the transaction before revenue or gain can be recognised’.

35 It is therefore curious that the Public Sector Accounting Centre of Excellence of the Australian Society of CPA’s, in its useful study of asset valuation approaches for GTE’s (Asset Valuation by Government Trading Enterprises, March 1996), argues that ODV accounts should record holding gains on debt. Even if there were such gains (which assumes that creditors do not fully price protect), they would be transfers among stakeholders in the entity, and hence would not be viewed as income in an entity (as against proprietorship) accounting approach.

Infrastructure regulation and market reform 97
above, this approach breaks down when the composition of the assets held and services
provided by the firm changes substantially over time.

The second difference lies in the role and nature of optimisation. In theory, ODV bases
valuation not solely on MEERA but also on re-engineering the system being valued to
eliminate redundant assets. It is a matter of some debate as to how far this is actually
taken in ODV implementation\textsuperscript{36}. Nonetheless, two points need to be made.

First, the areas where ODV has been applied — primarily high-voltage electricity
transmission — have highly characterised technologies, in the technical sense of having
a clearly defined best-practice framed by tight physical design constraints.
Technological change tends to be relatively slow and incremental, rather than radical,
in nature, and relative input prices and implied service prices are broadly predictable.
As a result, there is a high level of agreement as to the ‘right approach’ to be adopted to
service provisioning in particular instances, facilitating the kind of optimisation
envisaged in ODV\textsuperscript{37}.

Second, even within these systems, optimisation has very largely concentrated on
stranded assets — that is, on assets which clearly will not be replaced, and where the
optimisation involves bringing forward in time the recognition of the asset’s writing-off.
This is a particularly simple version of system redesign, and one that raises issues
which are primarily commercial rather than technical.

In contrast, in telecommunications, where technological change is faster and more
radical, there is far greater uncertainty as to the network designs which will ultimately
prove successful, increasing the error associated with hypothetical efforts at
optimisation. Moreover, when there is uncertainty as to the ‘best bet’, yet investment
decisions are irreversible, the appropriate metric for determining the efficient path of
investment is not the Net Present Value of expected cash flows (or even less of future
gross costs) but rather the option value of each of the alternative paths\textsuperscript{38}. However,
even the most sophisticated of the current network optimisation models are far from

\textsuperscript{36} It has, for example, been suggested that the stream of service costs associated with the asset
values generated by the ODV valuation of Transpower in New Zealand are too high to be
consistent with effective optimisation; see Cox, Spiller and Teece, 1994.

\textsuperscript{37} Equally, the fact that these systems involve many assets which are extremely long-lived, and for
which reproduction costs may not be available or meaningful, makes it more sensible to use
MEERA for class-of-asset valuation than it would otherwise be.

\textsuperscript{38} Kester shows that project selection on the basis of the NPV metric may yield extremely
misleading results under these circumstances: see the examples given in Kester, 1993, pp 187–
207; see also Trigeorgis, 1988.
being capable of carrying out this kind of analysis, and hence cannot provide a reliable indication for asset valuation purposes of the gap between the ‘network as it is’ and the ‘network as it ought to be’. Lastly, the fact that telecommunications networks generally require two-way, any-to-any, compatibility means that the adjustment costs involved in shifting vintages can be considerable. Efficiency, even when narrowly defined as cost-effectiveness, may therefore require a network design which seems quite at odds with that of a ‘green-field’ alternative; but this too is poorly captured by drawing-board contrasts between actual and (hypothetical) best-practice.

As a result, ODV-style optimisation cannot readily be implemented in the circumstances of the telecommunications industry. Rather, as suggested above, current price information should be used to value assets onto a replacement basis, with anticipated reductions in asset values being treated as part of period costs.

A final approach is that of exit price valuation, which is often referred to as Continuously Contemporary Accounting (CoCoA). The criticisms of CoCoA are well-known: not the least of these is the point that if the economic value of firms simply equalled the sum of the current resale prices of the individual assets under their control, firms as such would cease to exist. Further, CoCoA seems to be based on a confusion as to the nature of opportunity costs: there is no sense in which the opportunity cost of using an asset for any one purpose can generally be taken to be equal to the salvage value of the asset; rather, it is equal to the value of the asset in the next most-highly valued use, which may well be its deployment for some other purpose within the firm. Moreover, even if the disposal stand-point is adopted, exit prices are in practice highly sensitive to the level of aggregation (the ‘bundles’ of assets being priced), to the timing of sale, and to the information asymmetries familiar from the ‘lemons’ problem in the market for used cars; as a result, exit prices are difficult to define and calculate. Because of these issues and others, a positive profit under CoCoA valuation merely indicates that the firm should retain its current assets, but says nothing as to whether it should replace these assets over the longer term.

39 Some of the analytical requirements for even rather simple models of network evolution under uncertainty are set out in Adjali, Fernandez-Villacanas and Gell, 1994, making the gap between these requirements and current implementations all the starker.

40 Exit price accounting is primarily associated with Chambers, 1966, and Sterling, 1970.

41 After all, the resale (exit) prices are presumed to be certain, while any income to the residual claimant is not. As a result, all assets would be sold.

42 King, 1996, suggests the use of exit prices (salvage values) as an alternative to ODV (which he considers impractical), but does not consider the standard issues raised in this respect in accounting theory. See, for example, Godfrey, Hodgson, Holmes and Kam, 1994, at pp 181–202; Whittred, Zimmer and Taylor, 1996, at pp 577–580, who conclude that while exit price accounting has some attractions, it is impractical and hence not used in contracts; and Thomas, 1969, at p. 67 and following (although it is fair to note that Thomas, despite his earlier strong criticisms of exit prices, later gave highly qualified support to the use of CoCoA as being less harmful than conventional historical cost accounting — see American Accounting Association, Florida, 1974).
Despite these weaknesses, the main feature of CoCoA of relevance here is its treatment of highly specialised, ‘nonvendible’ assets (such as ducts). In CoCoA, these are subject to immediate expensing, on the grounds that the firm, in acquiring these assets, forecloses the opportunity of adapting to changing circumstances through either the postponement of decision or (subsequent to purchase) the resale of the assets in question. In this, CoCoA parallels the practice common for many years in railroad valuation, and which has re-emerged in recent years as ‘infrastructure accounting’.

Two points need to be made in this respect.

First, infrastructure accounting offers only a very partial valuation framework, because it deals mainly with the nonvendible component of the firm’s asset base. The valuation method to be used in respect of the remaining assets is not determined within the infrastructure accounting approach.

Second and most important, the immediate expensing of nonvendible asset outlays, if reflected in prices, entails a form of short-run marginal cost pricing, in which prices rise steeply as capacity constraints are approached. Simulation studies suggest that such pricing can provide for welfare gains when compared to smoothed, or Boiteaux, prices based on approximations to long-run incremental costs. However, maintaining these welfare gains when the access service is required to break-even will generally entail a lump sum charge which will be the higher, the greater the lumpiness of capacity, the difference between capacity goods in the size of efficiently-scaled lumps, and the extent of economies of scale in capacity addition. The effect of these lump-sum charges on the cost structure of entrants, combined with the inherent variability of congestion-related charges, may make them unacceptable. Moreover, the calculation of short-run marginal costs in a network environment will generally require imputing shadow capacity prices at each node in the network — a task of very great complexity.

---

43 See, for example, Office of Water Services, 1994.
44 See especially Ng, 1987. Somewhat unusually, Hausman says that ‘short-run marginal costs do not account for the cost of capital at all’ (Professor Jerry A Hausman In the Matter of Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, Federal Communications Commission, CC Docket No. 96–98, at p. 8). This is not true with goods subject to congestion, in which short-run cost rises to the lesser of the value of lost demand, or the full cost of the next lump of capacity, as supply constraints are approached. As Hotelling put it ‘when a train is completely filled, and has all the cars it can haul, the marginal cost of carrying an extra passenger is the cost of running another train.’ (Hotelling, 1938, at p. 264). In switched telecommunications networks, where the blocking probability rises rapidly as the ratio of the offered load to capacity rises, the congestion cost element in short-run marginal cost can be triggered even at what seem to be low levels of capacity utilisation.
45 Boiteaux prices involve smoothing the path of costs over time; they are obtained by pricing as if (lumpy) capacity were perfectly divisible — that is, pricing at the average incremental cost of capacity plus variable costs. Since Boiteaux prices are defined as the minimum of the cost function, Boiteaux pricing requires that capacity be fully utilised, which in turn requires that unit charges are calculated at a level of output defined in terms of integer multiples of the asset with the largest minimum efficient lump size. As a result, at output levels other than that corresponding to full utilisation, Boiteaux prices will fall short of recouping costs.
in fully-interconnected star or mesh network structures. As a result, estimates of these costs can be subject to considerable error.

Overall, three conclusions flow from the analysis set out above.

First, asset values should be set on the basis of replacement costs, subject to the constraint that no asset be valued in excess of its recoverable amount. So as to maintain consistency between the balance sheet and the income statement, replacement costs should be calculated by reference to the current purchase prices of the best in-use assets. Since these prices will reflect expectations of efficiency gains available from next-period technology, the use of these prices provides the most reliable means of effecting optimisation.

Second, anticipated changes in replacement cost should be included in the period depreciation charge, and hence netted from income in the calculation of economic profit.

Third, assets with indefinite lives may be treated separately and expensed in their period of acquisition. However, this entails acceptance of charges which may be highly variable over time, and may differ substantially from node to node in the network. Moreover, as congestion is reached, this approach may give rise to contentious, and fundamentally irresolvable issues with respect to cost responsibility, as traffic-generating units seek to be treated as infra-marginal so as to escape the charge bearing on marginal traffic.

The treatment of joint and common costs

However complex it may be, asset valuation is obviously only a part of the process of defining service costs. In practice, asset services flow to a range of products, both at any one point in time and across points in time, and cost definition requires some means for identifying these flows. These issues, classically dealt with in the setting of cost recognition rules, will not be canvassed here; rather, I will take as given ‘best practice’ analysis of cost causation, and focus on the problems associated with the recovery of those costs which, despite such analysis, cannot be unambiguously allocated to particular service flows.

The problems posed by these costs, which can loosely be referred to as common costs, are clearly of considerable practical significance. Thus, Professor Hausman states that ‘fixed and common costs are typically estimated at about 50 per cent or more of total Local Exchange Carrier (LEC) costs.’ Professor Hausman’s approach is unusual in grouping ‘fixed’ costs (which may well be product-specific) with ‘common’ costs (which by definition are not). Moreover, in the long run, all the costs are variable, but not all are product-specific. Professor Hausman then states, at paragraph 10, that TSLRIC will not recover ‘fixed and common costs which arise from network economies of scale and scope’. Either this statement is wrong, or Professor Hausman is actually referring to joint and common costs. Fixed costs, and those associated with economies of scale, are included in TSLRIC.
costs which are truly common is rising over time, as the transition to multi-service, integrated networks increases the economies of scope attainable in telecommunications networks.

Further issues of common cost recovery arise from the costs of non-commercial service obligations. These costs are not caused solely by the requirement to provide designated services throughout Australia at geographically averaged prices; they also arise from the constraints which the current price caps impose on price rebalancing as between the provision of access on the one hand and of usage on the other. There seems little doubt that the latter costs are very much greater than the former; yet it is only the former that are captured in the explicit funding mechanisms designed to recoup the costs which non-commercial obligations impose. The unfunded liability this creates can be viewed as a cost which is common to, and must be recouped across, the range of services provided.

In short, common costs are likely to account for a substantial part of the total resources deployed in the telecommunications network. Barring direct subsidies (which are clearly not a feasible political option at this stage), these costs must be recovered if the access provider is to break-even.

Efficient recovery of these costs will require a mark-up over the attributable long-run costs of each service, including access. This is for two reasons.

First, recovery of the entirety of these costs at the final service (retail) layer would only be efficient if the access service provider could directly or indirectly tax competitors’ revenues at that layer by an amount just sufficient to recoup the cost shortfall. A tax only applied to the access service provider’s final output would distort choice toward competitors’ offerings, and could result in substantial inefficiency. Since such a competitively neutral tax on final output is not feasible under the regulatory arrangements, some part of the relevant contribution must be obtained through the price of access.

Second, if access is priced only at Total Service Long Run Incremental Costs (TSLRIC), access providers will, under plausible assumptions, lack incentives to fully

---

48 The increased ability to exploit scope economies is emphasised (although without being referred to as such) in, for example, the analysis of ATM networks in Linberg, 1995; of optical transport in Gerstel, 1996; and of intelligent networks in Magendanz and Popescu-Zeletin, 1997.

49 See, for example, the estimates of inherited price distortions presented in Industry Commission, 1997, at p. 99 and at pp 149–156. These confirm the earlier results presented in Ergas, Ralph and Sivakumar, 1990, derived from reworking the data set used in the landmark study by the Bureau of Transport and Communications Economics, 1989. The Ergas-Ralph-Sivakumar estimates, when combined with those from BTCE, imply that transfers between services are nearly ten times greater than those between geographical areas.
achieve economies of scope in the provision of access services. Rather, they may opt for technologies which increase total costs, so long as a larger share of these costs is directly attributable to the access service. Given that the regulatory arrangements may create other incentives which run in the same direction, it would seem undesirable to compound the efficiency-reducing effects through the system of access price regulation.

Given that a mark-up over long-run attributable costs is required, the allowable nature of this mark-up needs to be determined.

In principle, the mark-up should be sensitive to demand conditions, with the wedge between prices and attributable costs being set so as to minimise the mark-up’s effect in suppressing demand. The practical implementation of this principle needs to be modified to take account of the consequences of the constraints on service price rebalancing, which mean that loss of contribution in the (relatively elastic) trunk services cannot be recouped through raising charges for rentals and local calls. As a result, although the unconstrained second-best mark-up would be greatest on inputs to the local access service, and lowest on those to the trunk service, the opposite pattern may need to prevail.

Within each service, the mark-up across customers should also be such as to minimise the efficiency costs of the price distortion. This, in turn, suggests that the contribution

---

50 Hausman put this proposition as follows: ‘If the LECs do not recover all of their joint and common costs, they have an incentive to use technology with reduced economies of scale and scope but higher per-unit (in case of LRIC) or per-service (in case of TSLRIC) costs, because the latter costs will be more fully captured by prices set at incremental cost. This action may be rational for the firm but it raises social costs and deprives society of productive efficiencies.’ Professor Jerry A Hausman In the Matter of Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, Federal Communications Commission, CC Docket No. 96–98, at p. 5.

51 For example, assume that there are two services, ‘access’ and ‘usage’; that all costs are fixed; that there is only one provider of access but that usage is contestable; and that there are two alternative technologies, one which allows substantial economies of scope (and lower total costs (TC)), and one which does not. Let TC1 and TC2 refer to total costs for the first and second technology respectively, and similarly for LRIC(access). If access prices are set at LRIC, and TC1 - LRIC1 (access) > LRIC2 (usage) then, noting that competitors downstream cannot be assumed to be operating at more than LRIC2(usage), the access provider will choose technology two (since its ‘costs’ in the downstream market otherwise exceed those of competitors).

52 Thus, incentives to increase the attributable costs of access may arise if access prices are effectively subject to some form of lagged rate of return regulation, and access demand has a low firm- and product-elasticity of demand, while retail prices are subject to incentive regulation and are more elastic.

53 As Baumol and Sidak succinctly put it ‘where economies or diseconomies of scale are present, both the state of demand and the structure of costs must be taken into account in the setting of efficient prices’; Baumol and Sidak, 1994, at p. 50.

54 A useful survey of research on demand elasticities for telecommunications services is provided in Industry Commission, 1997, at pp 31–42 and 133–148.
to common costs should be recovered to at least some extent through a higher charge on infra-marginal units. For example, a non-linear pricing scheme, in which fixed fees played a part, could well provide a less distorting means of pricing the access service than would a uniform price. The effects of such schemes are admittedly complex when the number of consumers is endogenous, and the complexities are all the greater when the consumers of the service are competitors in a downstream market. However, assuming that the downstream service is at least reasonably potentially competitive, a multi-part tariff which is available to all purchasers of the input is unlikely to distort competitive conditions, so long as the purchase price at the top end of the discount scale does not fall below attributable cost\textsuperscript{55}.

Experience in competitive markets is relevant here. In effect, studies find (1) that firms, in setting internal transfer prices, take account of the need to recover common costs, and hence mark-up over attributable costs; (2) that complex, non-linear charging arrangements are used to minimise the efficiency costs of the mark-ups; and (3) that these patterns of charging are consistent with the need to align the incentives of managers at decentralised levels within the firm with overall value maximisation, notably in the taking of investment decisions\textsuperscript{56}. At the same time, similarly complex, discriminatory pricing arrangements are widely used in respect of third parties, including in markets for intermediate inputs where the supply of these inputs is characterised by economies of scale and scope\textsuperscript{57}. All of this merely reinforces the conclusion that efficient pricing of access services will require prices which depart significantly from uniformity.

**Determination of the cost of capital**

The last issue to be considered here is the determination of the cost of capital. In a relatively capital-intensive activity such as telecommunications, the approach taken to determining the cost of capital can have a major impact on estimated service costs. This makes it all the more important that the approach adopted to be firmly based.

The Capital Asset Pricing Model (CAPM) provides the primary point of reference in this respect. The CAPM has figured in AUSTEL’s regulatory price determinations; and it is put forward for special consideration in the Australian Competition and Consumer Commission’s guidelines for Part IIIA undertakings, as well as for access pricing in telecommunications.

The prominence of the CAPM sits uneasily with its performance. It has been known for some years that the CAPM does not perform well in explaining behaviour in capital markets, although it is admittedly difficult to frame models which discriminate between testing for the CAPM and testing for capital market efficiency. Further, it is also

\begin{footnotes}
\item[55] The relevant tests are set out in Baumol and Sidak, 1994, at p. 73; and in Baumol, 1996.
\item[56] For empirical studies see, for example, Fremgen and Liao, 1981; Joye and Blayne, 1990; and Shih, 1996, 178–195. The relevant theory is surveyed in Hemmer, 1996.
\item[57] The results set out in Dolan and Simon, 1996, provide a wealth of examples in this respect.
\end{footnotes}
known that the CAPM has difficulty in explaining relatively widespread patterns of corporate behaviour, for example the taking of insurance against what appear to be diversifiable risks. Last but not least, the analytical bases of the CAPM have come under sustained criticism in recent years, and the results of that criticism suggest that the CAPM may be especially inappropriate in telecommunications.

In essence, in the CAPM, uncertainty only affects investment through the expected covariance of a firm’s earnings with those of the market (which is defined as the collection of assets which would figure in a fully diversified portfolio). An increase in the covariance increases the riskiness of investment, increasing the required rate of return. Conversely, a change which leaves this covariance unchanged will not alter the firm’s incentive to invest or its desired level of the capital stock even if it increases the expected variance of the firm’s earnings.

However, this view of the relation between the expected pattern of future earnings and the required rate of return does not hold when investment is irreversible. At its simplest, this is because irreversibilities make returns to investment asymmetric. If outcomes fall short of expectations (say because competitors’ entry costs fall more rapidly than anticipated), investors are stuck with the low returns. Conversely, if outcomes exceed expectations, the incentive is to invest more, thereby limiting the rise in the marginal product of capital. As in the CAPM, an increase in uncertainty will, under these circumstances, lead to a rise in the required rate of return and a reduction in the desired capital stock. However, in contrast to the CAPM, so long as the change increases the expected variance of the firm’s earnings, the discouraging effect on investment will occur independently of any change in the correlation between the uncertainty bearing on the firm and that bearing on the market as a whole.

Empirical tests confirm this result, showing that it is the variance of earnings, rather than the covariance with market returns, which determines required rates of return on investment. The results of these tests suggest that CAPM-based estimates of capital costs will be inaccurate, and may be seriously so in an industry as capital-intensive as telecommunications.

Two further points are important in this regard.

---

58 See, for example, Stedman and Kim, 1993, at pp 120–139.
59 See generally Dixit, 1989; and Pindyck, 1988.
60 Even if a firm can resell its capital stock, the same market conditions reducing the attractiveness of those assets to the firm may well reduce their attractiveness to potential buyers. As a result, even industry-specific, but not firm-specific, investment may be irreversible in the face of industry-specific shocks. See Abel et al, 1996.
61 This simple model is set out in Leahy and Whited, 1996.
62 See for example Howe and Vogt, 1996; Hurn and Wright, 1994; Leahy and Whited, 1996.
First, greater competition, to the extent to which it increases the variance of expected earnings, will increase the required rate of return on investment\(^{63}\).

Second, the extent of these effects is likely to rise as advances in telecommunications technology increase the minimum efficient scale of capacity increments. While new technologies do offer some scope for modularity, the trend, at least in the fixed network, is to provision capacity in ever larger lumps\(^{64}\).

Given these considerations, the determination of capital costs will need to draw on an option value framework, rather than on the CAPM. While analytically challenging, considerable experience has by now accumulated in this type of analysis, making it ripe for use in the costing of access services.

**Conclusions**

This paper has touched on only a few of the many issues which arise in the costing of access services. The main points made can be summarised as follows.

First, asset valuation should be based on replacement costs, with anticipated reductions in asset values being charged against period revenues. Rather than system optimisation of the kind envisaged in ODV, the effects of improvements in technology should, where reasonably possible, be addressed by pricing assets at entry prices for the best in-use vintage. The major constraint here lies in the possible error involved in estimating the relevant prices. Unfortunately, the higher the rate of technological change — and hence the greater the gap between historical and current costs — the greater may also be the difficulty and error estimating replacement costs entails. This can severely limit the additional informational the analyst — be that analyst a regulator or investor — can derive from replacement cost valuations.

Second, access prices will need to contribute to the recovery of joint and common costs. The contribution sought for these costs from access services will need to reflect regulatory price distortions, notably so as to avoid inefficient entry. To minimise the resulting economic costs, access prices should be structured in such a way as to secure the greatest contribution from infra-marginal traffic. This will entail a move away from uniform prices.

Third, the cost of capital benchmarks used to evaluate the reasonableness of access charges need to reflect the effect of irreversibilities in investment. Cost of capital

---

\(^{63}\) Thus, Caballero and Pindyck, 1996, find, using U.S. data, that a doubling of industry-wide uncertainty raises the required rate of return on new capital by about 20 per cent.

\(^{64}\) Fibre optic transmission, for example, is being provisioned in cables with up to 288 fibres, with single-fibre-pair WDM multiplexers available capable of providing 40 Gb/s over 600 kilometres. Assuming that each fibre supports 20 wavelengths, and each wavelength is modulated at 2.48 Gb/s, a node joining the cable to others would handle over 6000 wavelengths and have an offered load of some 15 Tb/s. Optical servers capable of handling this throughput would dwarf even the largest current exchanges.
estimates derived from the CAPM are likely to understate these effects, and hence to underestimate required rates of return.

Ultimately, the regulatory contract is by its nature incomplete65. Preserving the ability to respond to change requires that the parties retain a degree of flexibility, rather than seeking to anticipate, at the outset, all possible contingencies. In turn, the discretion which contractual incompleteness vests in the regulator, and the resulting scope for the regulator to act opportunistically66, inevitably creates an added element of uncertainty. Reducing uncertainty is not a sensible goal in and of itself67; but nor is any purpose served by introducing more uncertainty than is needed68. A clear commitment to addressing the issues set out above in a manner consistent with economic efficiency could only help to achieve a desirable balance in this respect.

Bibliography


---

65 See Crocker and Masten, 1996.

66 A common form of opportunistic conduct by regulators is the attempted expropriation of sunk assets, which (by reducing the return on these assets below that corresponding to long-term replacement valuation) takes from the supplier the quasi-rents it would otherwise secure. Extreme forms of this have at times prevailed, notably but not solely in developing countries, often under the pressure of political populism.

67 See for example Raz, 1977, stressing the trade-offs between certainty, both as a goal in itself and as a means to other goals, and other objectives of public policy.

68 Thus, the evidence suggests that unstable and opportunistic behaviour by U.S. regulators, notably in disallowing the recovery of investments, has been associated with increases in the cost of capital to regulated utilities and reductions in their investments and service quality. See, for example, MacAvoy, 1992, at pp 53–57 (on the experience of the electric utilities); and Loudder, Khurana and Boatsman, 1996.


