Abstract

The pricing and investment policies of a public enterprise should be designed to achieve efficiency since equity is better pursued by general policy regarding income distribution. Short-run marginal-cost pricing does not generally lead to long-term deficit, but may involve price and surplus/deficit cycles for the case with lumpy investments and growing demand, where the price increases with demand but is reduced with capacity expansion. Taking account of the extra costs of government revenue collection and the likely average price/cost ratio in the economy, the third-best pricing policy is likely to result in long-term surplus, making the objectives of equity, efficiency and financial viability much more consistent with each other than is generally believed. This is particularly true for water with historically increasing costs of additional sources of supply.

1. Introduction and Summary

In the pricing and investment policies of a public enterprise, the criteria of economic efficiency, equity and financial viability are usually regarded as important. (For example, see Pollett 1985, p. 11.) However, these three objectives are usually regarded as being in conflict with each other so that their simultaneous achievement is impossible. At best, a compromise between the three may be achieved. After examining the conceptual principles relevant for practical pricing policies, this article argues that the three objectives are far more consistent with each other than is generally believed.

Though this article refers mainly to water supply, the principles involved are applicable to many other public enterprises, especially those (for example, power supply) with lumpy investments and growing demand. On the other hand, this article is not concerned with other relevant issues such as managerial motivation in cost minimisation, and external costs and benefits.

Where it is necessary to meet growing demand through periodic lumpy investments (for example, reservoir construction) to expand capacity, it is shown in Section 2 (first in a simplified model, then with relaxation of a number of assumptions) that the simple consideration of economic efficiency in a first-best world results in pricing at levels higher than sufficient to cover full costs. In terms of the rate of return on capital, this would allow the water authority to earn a rate in excess of the market rate of interest.

When consumption is limited by supply capacity, the opportunity cost of water consumption is not the supplier's short-run marginal cost (typically low), but the marginal demand price. This is typically higher than long-run marginal and average costs when demand is high.
relative to capacity — a consideration which seems to have been ignored in a number of recent reports on water pricing and capital structure (debt-equity ratio) which have wrongly concluded that short-run marginal-cost pricing must lead to deficits. With an efficient schedule for the introduction of lumpy facility expansion, short-run marginal-cost pricing should result in low unit-price when capacity is large relative to demand and high unit-price when the reverse is true. The consideration of political feasibility may call for a smaller fluctuation in price or for the adoption of a two-part tariff (a fixed charge plus a unit price) whereby the fixed charge is varied in the opposite direction to the unit price (Subsection 3.3).

The fact that economically efficient prices averaged over time are higher than long-run average cost is due to the fact that water supply faces increasing cost of capacity expansion over time. As demand grows, resort to more and more costly sources is necessary. This point has also been missed by the recent reports which do not distinguish between the microtheoretical cost curves and historical changes in costs.

Since the ratio of average price to marginal cost over the whole economy is larger than one, and since water has no important close complements and substitutes, a third-best policy (roughly one that takes optimal feasible account of second-best complications; see Ng 1977 and 1983, ch. 9) suggests pricing water at above marginal cost (Subsection 3.3). Similarly, the consideration (Subsection 3.2) of the extra costs (excess burden, tax evasion, administration and compliance) of raising government revenue also suggests pricing at above marginal cost. In conjunction with the argument (Section 2) for above full-cost pricing, these considerations suggest substantial surpluses for the water authority, which could be used to pay back existing loans or as a source of government revenue.

Since pricing in accordance with economic efficiency results in surpluses for the water authority, it is consistent with the objective of financial viability. What about equity?

Two concepts of equity are discussed, one between different groups of consumers (as classified by incomes, property values, or types of consumers, that is, residential versus commercial) in the same time period, and the other between present and future generations. For the following reasons (incomplete listing), it is concluded that there is little reason to deviate from the objective of economic efficiency due to the considerations of either concept of equity.

(i) The role of achieving a more equal distribution of income properly belongs to the federal and state governments, not the water authority.
(ii) There is a presumption that redistribution through general taxation is more efficient than specific equality-oriented measures not based on some efficiency considerations (for example, second best, externalities) since the latter involves specific distorting costs on top of the costs of disincentive effects present in both cases (Ng 1984).
(iii) If taxing property is an efficient way to achieve equality, a property tax should be introduced instead of tying it to water consumption.
(iv) The present generation, after paying the above full-cost price, has the option to reduce its general transfer to the future generations.

2. Marginal-Cost Pricing: Short-Run Versus Long-Run

The first-best efficiency case for marginal-cost pricing (MC-pricing) is straightforward. (Second-best complications are discussed in Section 3 and equity considerations in Section 4.) A long-standing problem troubling practical policy-makers is the issue of short-run (S-R) versus long-run (L-R) MC-pricing.

The case for using S-R MC for MC-pricing seems quite compelling. Provided that prices can be varied soon enough, without incurring prohibitive costs, to reflect S-R cost and demand situations, failure to adopt S-R MC-pricing results in allocative inefficiency. In water supply, since the storage capacity is usually fixed for a large number of years, the so-called short-run can be very long. This makes the use of S-R MC even more compelling. However, if S-R MC is below L-R AC (average cost), S-R MC-pricing results in perennial losses. Secondly, if S-R MC is smaller than L-R MC, the use of S-R MC-pricing encourages the excess consumption of water in the long run, necessitating expansion of storage capacity at costs not warranted by consumers’ willingness to pay. It thus seems equally compelling that it is the L-R MC, if not AC, that should be used in determining the price. How is this dilemma between S-R and L-R MC-pricing to be resolved?

2.1 The Static Case of Perfect Divisibility

Most theorists use a static model with perfect divisibility in the size of the fixed (for the short run) facility. In this case, we have the simple solution that S-R and L-R marginal costs are equal if the size of the fixed facility is chosen correctly. This is illustrated in Figure 1. Partly because this is a well-known textbook case (for example, Layard and Wal-
Public-Utility Pricing

For any given static demand curve, the optimal facility is determined by the intersection of the demand curve and the L-R MC curve. For example, if the demand curve is \( D_1 \), it is optimal to produce at \( q_1 \) in the long run. To produce \( q_1 \), it is optimal to build a facility with \( SAC_1 \), since this gives the lowest \( AC \), as \( SAC_1 \) is tangent to \( LAC \) at \( q_1 \).

But why should \( SMC_1 \) also intersect \( D_1 \) at the same point? This must be so since the fact that \( SAC_1 \) is tangent to \( LAC \) at this output level means that \( AC \) and the change in \( AC \) as output changes are both the same in the short run as in the long run at this point. So \( MC \), being the change in \( TC \) (total cost = \( AC \) times output), must also be the same in the short run as in the long run at this output level.\(^1\)

Figure 1 also illustrates the result that a firm which adopts MC-pricing will make profits, break even or make losses, depending on whether its \( AC \) curve is rising, flat, or declining at the output level at which its L-R MC curve cuts the demand curve.

2.2 The Dynamic Case of Growing Demand and Lumpy Investment

2.2.1 A Simple Model

Water supply is characterised by relatively fixed supply (or rather, exogenously given in accordance to the rainfall conditions) for any given storage capacity and by the need to expand this capacity from time to time by building new reservoirs to meet growing demand. Since demand is typically growing rather than static, and since reservoirs are lumpy (minimum sizes are needed for feasibility and cost-effectiveness), the static model of the preceding section cannot be applied. It is little wonder that policy makers find the simple solution of S-R MC = L-R MC of Figure 1 of little help in water pricing.

Apart from lumpiness of investment and growth in demand, water supply is also characterised by expected (for example, seasonal) and unexpected (random) variation in supply (rainfall) and demand. However, to start from a simple model, let us abstract away this uncertainty as well as such complications as non-constant costs, alternative sizes and sites for reservoirs, consumption of water from previous reserves, etc.\(^2\) Also, we concentrate on the water generation activities, ignoring the water distribution system. The costs of the latter could

---

\(^1\) This can be shown more precisely thus. Let total cost = \( C(q) \), \( AC = C(q)/q \). Slope of \( AC \) = \( dAC/dq = (C' - C/q)/q \). Hence, if \( C/q, q \), and slope of \( AC \) are all equal at \( q_1 \), \( C' \) (marginal cost) must also be equal at \( q_1 \) for the short run and the long run.

\(^2\) Thus, among other things, the distinction between capacity level and annual yield of a reservoir is ignored. If desired, the K units of water a reservoir is said to supply below may be taken as the average annual yield if storage level is to be held constant.

The Australian Economic Review 3rd Quarter 1987
form part of the fixed charge in a two- 
part tariff discussed below.

In our hypothetical model, the water 
supply authority is faced with a steadily 
growing demand illustrated by a shift of 
the demand curve from $D^1$ to $D^2$ and so 
on, in Figure 2. As drawn, the demand 
curves are linear and the regular growth 
in demand is arithmetical (that is, growing 
by a constant amount each period at each price) and parallel (that is, the 
same amount of increase at all prices).

Cases of non-linear demand curves and 
proportionate shifts in demand curves 
will be discussed below.

The horizontal line $S-R$ MC represents 
the per-unit supplier cost (assumed con-
stant) given the storage capacity. The 
vertical distance between $L-R$ AC and 
$S-R$ MC is the per-unit cost of storage 
per period, including maintenance of 
storage and interest charges on capital 
construction. The divisor for calculating 
this per-unit cost of storage is not the 
quantity stored at any time, but the num-
ber of units supplied per period. It is 
assumed that storage capacity is main-
tained in a non-depreciating condition to 
last indefinitely and that the cost must 
be paid each period. The rectangle 
HIJK in Figure 2 is the per-period cost 
(assumed constant) of each additional 
lumpy reservoir (assumed fixed in size, 
supplying $K$ units of water each period).

Inter-period transfer of water is ignored.

Starting from a situation where there 
are $N$ (a positive integer) reservoirs in 
operation, the $(N + 1)^{th}$ reservoir is only 
worth its costs when the demand curve 
has grown at least to the position $D^1$ 
where the area of triangle ILM equals 
the area of triangle MJV. (The case 
where $V$ lies below $S-R$ MC will be dis-
cussed below.) This is so since the 
benefit of the extra amount of water 
supply (per period) brought about by the 
$(N + 1)^{th}$ reservoir is measured by the 
area under the demand curve as an acceptable 
approximation for welfare gain.) Assume 
that this $(N + 1)^{th}$ reservoir was duly 
scheduled to commence supplying water 
right at this period, the efficient price 
of water is at $P^1$ where the demand 
curve intersects the vertical capacity line 
at $V$. This involves a price higher than 
the supplier $S-R$ MC. Is this inconsistent 
with $S-R$ MC-pricing?

Where supply is limited by capacity 
before the supplier $S-R$ MC curve cuts 
the demand curve, the appropriate $S-R$ 
MC is no longer the supplier $S-R$ MC. 
In accordance with the correct concept 
of opportunity cost, cost is determined 
by the best alternative forgone.

When consumption is limited under $S-
R$ MC-pricing by the intersection of the 
supplier $S-R$ MC curve with the demand 
curve, the opportunity cost of a marginal 
unit of consumption is measured by the 
supplier $S-R$ MC, since this measures 
the MC of the resources used to pro-
duce this marginal unit of consumption 
($= $ marginal valuation placed on the 
forgone goods elsewhere in the economy 
under the first-best world). However, 
when consumption is limited by the

Figure 2
capacity constraint, no good elsewhere in the economy is forgone by the consumption of a marginal unit of water. No resources are transferred in or out of the water industry by the consumption of a marginal unit of water. The total resource usage (and hence total costs) in supplying this capacity level of water is fixed. This makes the supplier S-R MC irrelevant as a measure of opportunity cost. What then is the correct opportunity cost of a marginal unit of water consumption?

When consumption is limited by capacity, the opportunity cost of a marginal unit of consumption is another marginal unit of consumption. Since the total supply is fixed, if someone consumes an extra unit of water, someone else (or several others in combination) must consume one unit less. Thus, the forgone alternative of a marginal unit of consumption is another marginal unit of consumption which is valued by the height of the demand curve at the capacity level of consumption, that is, the point V in Figure 2 when demand is D^1 and capacity is at Q = (N + 1)K.

As demand gradually grows from D^1 towards D^2, the efficient price based on S-R MC-pricing gradually increases towards the point J. When demand is at D^2, the water authority breaks even for that period. For the time when demand is between D^1 and D^2, the authority accumulates losses. When demand is at D^3, note that the total amount of loss for that period is not just the rectangle RIJV. Assuming non-discriminatory and uniform pricing, S-R MC-pricing at P^1 results in an amount of loss totalling VJ times (N + 1)K.

As demand grows further from D^2 towards D^3, S-R MC-pricing results in surpluses (profits net of interest charges). When demand reaches or exceeds D^3, the (N + 2)th reservoir commences operation and the efficient price drops back to P^1. The whole cycle of deficits and surpluses starts all over again. In our present model of linear demand curves, VJ = JS in distance. With arithmetical growth in demand, it takes exactly the same time for demand to grow from D^1 to D^2 as it takes to grow from D^2 to D^3. Hence, the water authority is in deficit for exactly the same length of time as it is in surplus. Thus, ignoring interest payments/earnings on deficits/surpluses, the water authority exactly breaks even over time. It may be thought that, when interest payments/earnings on deficits/surpluses are taken into account, the authority will be slightly in deficit over time since it starts with deficits before making surpluses. But this is not necessarily so. We started our discussion from D^1 only for convenience. The water authority may start its new pricing policy with its demand curve anywhere between D^1 and D^3.

From the above discussion, we conclude that S-R MC-pricing does not necessarily result in deficit. The usual conclusion that 'the short-run marginal-cost principle cannot be applied' due to deficit is based on inadequate understanding of the principle of opportunity cost in the presence of capacity constraint as discussed above.

2.2.2 Relaxation of Simplifying Assumptions

Our happy conclusion above that even S-R MC-pricing allows the water authority to break even is based on the analysis of a simple model. Does the relaxation of the simplifying assumptions change the conclusion significantly? Let us consider these assumptions in turn.

(a) Convex Demand Curves

Demand curves are more likely to be convex (from below) than linear. A linear demand curve is elastic at high prices and inelastic at low prices and cuts the horizontal axis at a relatively small quantity. This is not consistent with the demand for water which may be presumed to be much steeper at high prices (when water is used mainly for essential purposes of drinking and cleaning) than at low prices (when it is used for gardening).

When we change the demand curves in Figure 2 into convex curves (such as D'^1 without changing other assumptions (arithmetical growth and parallel shift in demand curves in particular), it is not difficult to see that the water authority will be in surplus most of the time and with more surpluses than deficits over time. This is so because D'^7 intersects the vertical line through NK at a higher level than D^1 (that is, L' is above L) and intersects the vertical line through (N + 1)K also at a higher level (that is, V' is above V). This has to be so because the area L'M'N' has to equal the area M'JV' for D^1 to be a critical demand curve just high enough to warrant the (N + 1)th reservoir.

(b) Geometrical Growth and Proportionate Shifts in Demand

Growth in demand over time is more likely to be geometrical (growing by a given percentage per period) than arithmetical. It is not difficult to see that this tends to cause the water authority to be in surplus for a shorter period than
in deficit (in using S-R MC-pricing).

The outward shift in the demand curve is also more likely to be proportionate (the same percentage increase in quantity demanded at various prices) than parallel (same absolute increase). This, too, tends to cause the water authority to earn less surplus.

The two considerations above tend to offset the excess surplus caused by convexity in demand curves discussed in the previous subsection. A more precise analysis is provided in the Mathematical Note in the Appendix. According to this analysis, the water authority, when using S-R MC-pricing, will be in deficit slightly longer than in surplus, as illustrated in Figure 3. This slight deficit may however be offset by the following consideration.

(c) Demand Limited by Supplier S-R MC

In the analysis so far, it has been assumed that consumption is limited by capacity rather than the intersection of the S-R MC curve with the demand curve. In terms of Figure 2, the point V (where the demand curve intersects the vertical line of capacity limit) lies above the S-R MC curve. For cases where demand is inelastic and S-R MC is substantial, it is quite likely that the reverse may apply during the years just after the introduction of a new reservoir. For these years, efficient prices equal supplier S-R MC. From Figure 2, it can easily be seen that deficits in these years would then be less than the maximum surpluses earned during years just before the introduction of a new reservoir. This is also true for the case of non-linear demand curves and geometrical growth in demand discussed in the Mathematical Note.

(d) Non-Constant Costs

The analysis above assumes that the cost of constructing additional reservoirs is constant. This may not be so.

A methodological point may be noted in this connection. In microeconomic theory, cost curves are drawn assuming given technology and input prices. A change in these shifts the whole family of cost curves rather than affecting the shape of an unshifted cost curve. With this method, our assumption of constant cost of constructing additional reservoirs (of given size) must be true by definition. (But the assumption of a constant S-R MC may or may not apply.)

It may be thought that the \( (N + 1) \)th reservoir may have to be constructed at a less favourable site than the \( N \)th reservoir, so that it will cost more at given technology and input prices. However, in the microtheory terminology, this really means that there has been an effective increase in the price of one of the inputs: the price of a good site has increased to infinity so that the water authority has to use an inferior input — the next best site.

Nevertheless, while the methodology of microtheory is appropriate for the choice of reservoir size and optimal pricing in the short run, we have to consult changes in actual costs to see the water authority’s deficit/surplus position over the long run. The building of successive reservoirs is not decided upon in a static framework, but is necessitated by a gradually growing demand over a long period of time during which technology, etc. are bound to change. Moreover, we are interested in the question whether, over this long period, the water authority will or will not break even in using S-R MC-pricing. The change in costs, even if caused by changes in technology and input prices, has to be reckoned with somehow.

Recognising the above distinction between the hypothetical (if output were expanded at unchanged technology and input prices) or microtheoretical cost curves and the actual cost levels incurred in a historical process of meeting growing demand, interesting possibilities arise. For example, we may have decreasing average cost (downward-sloping AC curve) in the microtheoretical sense but increasing AC in the actual sense, or vice versa. If we further recognise cost differences due to different sites, this raises difficult problems in the decision on the timing, choice of sites and sizes for reservoirs. Computer simulation may be the only practical way of obtaining a reasonably accurate solution.

For our present problem as to whether in using S-R MC-pricing (for which the microtheoretical concepts should be used) the water authority will actually be in deficit or in surplus in the long run, we should be using the actual cost curves rather than the microtheoretical cost curves.

Using the simple model of Subsection 2.2.2, it is not difficult to see that, if the cost of constructing a new reservoir falls through time, this tends to bring forward the optimal time for the introduction of a new reservoir, making the water authority earn less surpluses. This leaves the authority in deficit over time.

![Figure 3](image-url)
On the other hand, if the cost of construction increases over time, the authority will be in surplus over time. This is illustrated in Figure 4 where dotted curves refer to the case of an increase in the cost of reservoir construction.

The cost of reservoir construction is affected by two sets of factors. One is technology and the other is input prices and availability, including suitable sites. Technological advances tend to lower the cost of construction but increases in input prices and increasing scarcity of suitable sites tend to raise it. Especially for municipal water supply, water sources have to be found from places further and further away from the city and at less and less favourable sites as demand grows through time. On the other hand, technology in reservoir construction is not an area where spectacular reduction in cost can be expected. Thus, it may reasonably be concluded tentatively that the adoption of S-R MC-pricing to attain economic efficiency will be more likely to result in surplus for the water authority.

A more reliable conclusion together with more precise numerical figures for efficient prices has to be based on a more accurate estimate of current and future cost conditions beyond the scope of the present article. Since this is at the heart of the issues of efficient pricing, investment and capital structure (that is, debt-equity ratio), further studies on this are desirable.

Some previous reports have dealt with these issues but not in a satisfactory way. For one thing, no distinction is made between the microtheoretical and the actual cost curves as discussed above. For another thing, the methodology used in some derivations of cost curves is suspect.

(e) Inter-Period Transfer of Supply and Uncertainty

The analysis so far has abstracted away uncertainties in both supply and demand and the storage of water for future consumption to concentrate on the issue arising from lumpy investment and growing demand. The relaxation of these assumptions complicates the analysis substantially. Nevertheless, the problem can essentially be dealt with by adopting a S-R opportunity MC curve that is upward-sloping at least when the current capacity is approached, as illustrated in Figure 5.

If we take a period to be reasonably lengthy, for example, a year, the capacity level is also somewhat uncertain until the end of the year. This makes the S-R opportunity MC also subject to additional uncertainty. That the S-R opportunity MC should rise sharply as capacity...
is approached and may even rise substantially well before capacity level is based on the following reasoning. The less water is left in the reservoir this year, the more likely it is that a severe shortfall in water supply will be experienced in the following years. Hence, the closer consumption is to the capacity level, the higher will be the opportunity marginal cost of water consumption since the alternative forgone could well be those units that would meet demand with very high marginal valuation.

Once the S-R opportunity MC curve has been estimated, the efficient price is determined by its intersection with the demand curve. The use of this gradually upward-sloping MC curve rather than the vertical capacity line in Figure 2 tends to narrow the difference between the lowest and highest prices.

2.3 S-R MC-Pricing: Concluding Remarks

Our discussion above suggests that the use of S-R MC-pricing in accordance with first-best efficiency consideration need not lead to deficit as commonly feared. In fact, for the case of water supply, it is more likely to result in surpluses.

On purely efficiency grounds, a policy of S-R MC-pricing involves cyclical price changes, with the price gradually increasing until a new lumpy facility (reservoir) is brought into operation, when the price should be lowered by a substantial amount and then gradually increased again as demand gradually expands. If the required fluctuation in price is big, and if the electorate does not understand economics, then a policy of S-R MC-pricing may not be politically feasible. The water authority may then have to adopt a compromise between economic efficiency and political feasibility. Even if the compromise adopted is one that is biased towards price stability, the level of price should still be high enough to cover full costs, if not at a level high enough to generate some surpluses.

3. Second Best and Other Efficiency Considerations: Average-Cost and Above Marginal-Cost Pricing

This section considers complications due to second best and other efficiency considerations which further strengthen the case for above full-cost pricing discussed in Section 2.

3.1 Second-Best Versus Third-Best Policies

The straightforward argument in favour of MC-pricing (price = S-R MC in the short run and price = L-R MC in the long run) is based on a simple first-best world where no distortion exists anywhere in the whole economy, for example, every sector observes the first-best rule of price = MC. But the real economy is riddled with all sorts of distortions. Realising this, theorists have developed a theory of second best, as summarised by Lipsey and Lancaster (1956). According to this theory, even the presence of a single second-best type constraint makes it inefficient, in general, for all sectors of the economy to observe the simple first-best rule. However, the resulting second-best rules are very complicated, making them informationally and administratively infeasible. This is especially so if we realise that the resulting second-best rules have to be followed by all sectors except the constrained sectors. It is no wonder that the theory of second best has had no significant impact on actual economic policy decision making.

In fact, if anything, the theory of second best probably does a disservice to the application of economics to policy making. This is so because those concerned with economic policy formulations and advice are typically roughly familiar with the theory of second best but do not know how to tackle the problem in practice. This is so since it is impossible to follow the complicated second-best rules and the observance of the simple first-best rule may make matters worse according to the theory of second best. Since going to the summit is impossible and what is an uphill direction is unknown, most practitioners are encouraged to ignore the efficiency issue.

The above unhappy situation can be avoided by using the theory of third best (Ng 1977, 1983). Essentially, according to this theory, since the objective function in a first-best world is likely to be concave, if we do not have sufficient information to form a reasonable probabilistic estimate as to how the presence of second-best distortions shifts the objective function, it is still optimal in expected value to follow the first-best rule. On the other hand, if we have enough information to form such a probabilistic judgement, we should revise the simple first-best rule accordingly. Nevertheless, due to substantial informational and administrative costs associated with a detailed second-best rule, the optimal third-best rule is typically far less com-

---

5 The conclusion that, 'at full capacity, short run marginal cost will equal long run marginal cost which will then be the appropriate volumetric price' (Peach 1985, p. 3) is thus misleading. It would lead to a downward bias in the price charged.

6 In the form $MRS = \frac{MRT}{MRT}$ for a pair of goods, where $MRS = \text{marginal rate of substitution}$, $MRT = \text{marginal rate of transformation}$. 
Public-Utility Pricing

Despite more complicated than the second-best rule though more complicated than the simple first-best rule. (See Ng 1977 and 1983, ch. 9 on the details of this argument.)

Typically, a third-best rule is derived by forming an estimate regarding the overall situation together with an estimate or more detailed investigation (depending on the available budget for obtaining information which is itself a variable in a wider problem of the economics of optimal information acquisition) of those sectors that are closely related to the one in question. For example, in determining the pricing policy for the railway, the interaction between railway and road transportation is considered but interactions with other sectors in the economy are ignored.

The case of water pricing is simplified by the fact that water has no important close complements or substitutes, especially not those goods which have important elements of distortion themselves. Thus, in the formulation of a third-best pricing policy for water, an overall estimate of the price/MC ratio for the whole economy is, roughly speaking, sufficient for the purpose.

The average price/MC ratio for the whole economy is almost certainly no less than one. In perfectly competitive industries with no important externalities, it is equal to one. In other industries, it is likely to be slightly larger than one. It is unlikely to be much larger than one except in entrenched monopolies where the monopolistic powers are under little threat from either existing competitors or from potential entrants.

A more reliable estimate of the average price/MC ratio for the economy would be very desirable. Such an estimate would be useful not only for water authorities but also for all public enterprises and all those concerned with public economic policy decisions in general and pricing policies in particular.

On the reasonable assumption that the average price/MC ratio is somewhat larger than one, the theory of third best suggests that efficient pricing policy for a public enterprise supplying a good with no important close complements or substitutes is to price somewhat above marginal cost, that is, to adopt a price/MC ratio equal to the average ratio of the economy. Obviously, this further strengthens our conclusion that pricing in accordance with economic efficiency should lead to surpluses for the water authority.

3.2 The Extra Costs of Financing for Deficits: The Case for AC-Pricing and Above MC-Pricing

Our analysis above suggests that pricing in accordance with simple considerations of economic efficiency should lead to surpluses for the water authority. We must now consider what should be done in cases where that policy leads to long-term deficit, due say to a downward-sloping L-R AC curve (with L-R MC lying below the AC curve) or to falling costs of reservoir construction. Is MC-pricing still appropriate under these conditions?

Deficits have to be financed and this is done mainly through loans or through government subsidies. Loans raised by a public enterprise have to be repaid sooner or later. If present deficits are to be paid by charging higher prices to future consumers, this not only raises intergenerational equity questions (to be discussed in Section 4) but is also likely to cause more inefficiency in the future when prices have to be raised well above MC, unless it is expected that the cost situation is likely to change into one with MC above AC in the future. Financing long-term deficits by loans does not seem to be an acceptable way out.

Government subsidies come from general tax revenue. The need to subsidise deficits of public enterprises thus increases the required amount of taxation, ceteris paribus. An increase in taxation imposes costs additional to the tax revenue. It leads to:

(i) an increase in the excess burden of taxation;

(ii) an increase in the costs associated with tax evasion and policing; and

(iii) (if the increased taxation comes from the introduction of a new tax), costs of administration and compliance.

Conservative estimates put the extra costs of a dollar of government revenue at about 50 cents to 1 dollar. This can hardly be justified by the usually moderate gain of MC-pricing.

As illustrated in Figure 6, the efficiency gain (ignoring second-best complications) associated with MC-pricing

Figure 6

Estimates of just the marginal excess burden alone give figures as high as 50 cents and possibly more than 3 dollars; see Stuart 1984, Ballard, Shoven and Whalley 1985 and Browning 1987.
Public-Utility Pricing

3.3 Two-Part Pricing

A danger of allowing pricing above AC (or earning a rate of return above the market rate) is that some public enterprises may then use it as a pretext in pursuing monopolistic pricing to maximise profit, raising prices beyond levels justified by efficiency considerations. This is an argument justifying a uniform rate of return for all public enterprises (Richardson and Wilson 1983, p. 30). However, while this may be so for certain public enterprises, the case of water supply is quite the reverse. Political pressures have resulted in prices well below the efficient level. A reversal is certainly warranted. It is a long way from having to worry about excessive prices in water supply.

3.3 Two-Part Pricing

It may be better to use two-part pricing (a fixed charge plus a unit price) instead of AMC-pricing. This is particularly likely to be so for water. Since some water consumption is essential for all users and substitutes are almost non-existent, a moderate fixed charge to account for the required amount in excess of MC can be introduced along with MC-pricing without driving out any consumers. This may achieve the same amount of surplus as the AMC-pricing illustrated in Figure 7, without having to incur the efficiency loss of EAH.

Though the elasticities of demand for water are larger than many people believe (since water is used for such things as gardening at the margin rather than drinking), they are still usually in the inelastic zone of about 0.4 - 0.5 in absolute value; see Seidel and Baumann (1957, p. 1541). Due to the importance of gardening usage, the elasticity figures for Australia may well be higher.

8 These considerations seem to be ignored by Aitchison (1983, p. 10) in reaching his conclusion: 'For those public utilities with marginal costs above average costs, and with negligible cross price elasticities ... the correct pricing policy is to set price at marginal cost ... For those public utilities with marginal costs below average costs, estimates are required of the welfare trade off between the move to marginal cost pricing and the recovery of deficits through taxation'.

9 Though the elasticities of demand for water are larger than many people believe (since water is used for such things as gardening at the margin rather than drinking), they are still usually in the inelastic zone of about 0.4 - 0.5 in absolute value; see Seidel and Baumann (1957, p. 1541). Due to the importance of gardening usage, the elasticity figures for Australia may well be higher.

10 In the presence of this effect (some consumers stop being customers altogether to avoid the fixed charge), an optimal two-part tariff (even with a positive budget constraint and with AC > MC) need not involve pricing at MC, contrary to popular belief. Above or even below MC-pricing may be optimal. Crucial factors include the response of consumers to the fixed charge as well as the average consumption of the marginal consumers relative to the average consumption of the average consumers (Ng and Weisser 1974), a point missed out by the discussion of Apps (1984, pp. 22-4) which is above average in its sophistication.

(output Q^M), in comparison to AC-pricing (output Q^A), is measured by ABG, being the area under the demand curve D less the area under the MC curve between Q^A and Q^M. The subsidy required is the rectangle EFGH. For the particular case illustrated, even if the extra costs of government revenue are as low as 20 per cent, they are big enough to offset the efficiency gain. Though one can construct cases where the contrast is not as big as in Figure 6, it is not unreasonable to conclude that, as a rule, where MC-pricing leads to deficits, AC-pricing is a more appropriate policy in view of the substantial extra costs of government subsidies.

It is true that, depending on the extra costs of government subsidy and the amount of subsidy required per unit of efficiency gain from pricing below AC (towards MC), the optimal price may be somewhat below AC, though well above MC. However, this is likely to be offset by the second-best consideration discussed above and by other efficiency considerations such as informational and administrative costs of calculating and effecting the optimal price, and the danger of cost escalation (enterprises not minimising costs) encouraged by subsidy. Taking account of all these considerations, the case for below AC-pricing can only be justified in very isolated extreme cases.

What about cases where MC-pricing results in surpluses or allows the enterprise to break even? Far from suggesting that MC-pricing should be replaced by AC-pricing, the consideration of the extra costs of government revenue suggests that prices of public enterprises should be somewhat higher than marginal costs to further increase the surpluses so as to relieve the need for raising costly government revenue elsewhere.

Using a static framework as in Figure 6, Figure 7 illustrates how AMC-pricing (above MC-pricing) increases the price from P^M to P^MM, the surplus from P^MABC to P^MEFG (about doubled), but only causes a relatively small efficiency loss of EAH. This depends of course on factors such as the fairly inelastic demand curve drawn in Figure 7. However, for water, this may well be approximately the case.

Figure 7
While the above argument in favour of a two-part tariff is valid, it does not dispose of the argument for AMC-pricing based on the second-best and cost-of-revenue considerations discussed in the preceding two subsections. This means that it may be optimal to have the fixed charge in addition to AMC-pricing. However, since this makes the total charge in excess of L-R MC, the size of the fixed charge and the extent of excess price (P-MC) should not be higher than what can be justified by the second-best and cost-of-revenue considerations. Otherwise, the marginal loss from excessive discouragement of water consumption will be higher than the marginal gain.

The use of a two-part tariff may also be a politically acceptable way to approximate the efficient optimal price cycles discussed in Section 2. Thus, a low unit-price with a substantial fixed charge may be used when capacity is abundant relative to demand. The fixed charge can be reduced or even waived completely as the unit price is increased to restrain demand when demand is high relative to capacity. This device may also be used to meet seasonal or other fluctuations.

3.4 Concluding Remarks: The Disposal of Surpluses

The consideration of both second best and costs of government revenue suggests that public enterprises should price their products above marginal costs. This further reinforces our conclusion in Section 2 that economic efficiency requires pricing that generates surpluses for water authorities. What should be done with the surpluses?

Partly due to policies of low prices and of free water entitlement associated with fixed charges (in accordance with property values), many water authorities in Australia have accumulated substantial deficits. Large loans are raised to meet the very high costs of constructing reservoirs that are prematurely required due to the low price policies that encourage excessive consumption of water. With the gradual movement towards the efficient pricing policy recommended in this article, surpluses should be generated that could be used to pay off the existing loans; alternatively, the surplus funds could be transferred to the government.

4. Equity Considerations

So far, we have been concerned only with issues of efficiency. It is now time to discuss distributional or equity considerations. These arise in two ways: equity between different groups of consumers (as classified by incomes, property values, or types of consumers, that is, residential versus commercial) in the same time period, and equity between present and future generations. These are discussed in turn.

4.1 Inter-Group Equity: The Case for Treating a Dollar as a Dollar (Uniform Pricing)

4.1.1 Residential Versus Commercial Users

Are there any acceptable grounds for charging different prices to residential and to commercial users of water? It may be thought that the commercial sector can afford to pay more for water than residential consumers. But there are big and small commercial units that are struggling for survival, and there are residential millionaires. Secondly, even if there were a systematic difference in ability to pay, the argument of Subsection 4.1.2 casts doubts on the desirability of its use for price discrimination.

Thirdly, under competition, the higher price of water charged to commercial users will be passed on to consumers in the form of higher product prices. What if these commercial users (firms) have monopolistic power? In the long run, monopolistic power that can sustain excess profit is of doubtful significance except for monopolistic power created by legislation. Moreover, the water authority is not in a position to discriminate between different firms in accordance to the degree of monopolistic power they enjoy.

If there exists a significant difference in the price elasticity of demand for water between residential and commercial users, it could be used for price discrimination in order to raise a given amount of revenue at minimum cost to the consumers (that is, Ramsay-type optimal pricing). If this is the objective, the distinction between residential and commercial users is unlikely to be a good classification. The water authority really needs statistics on the different elasticities of demand for different separable market groups. Moreover, this is really an efficiency argument, not an equity argument which really calls for a uniform price.

4.1.2 Property Values and Income Groups

The most common base used by water authorities for achieving some form of effective price discrimination and cross-subsidisation across consumer groups is property value. Typically, a fixed charge is levied that is proportional to the value of the property. To make it appear non-discriminatory, some amount of free water entitlement proportional to the fixed charge may be attached. (This was the practice of the Melbourne water authority before 1986.) The effective discrimination and cross-subsidisation comes from the fact that, for consumers...
with high property values, the fixed charges are so high that the free water entitlements are seldom used up. This of course encourages the uneconomical consumption of water.

The use of fixed charges in accordance with property values is usually justified on two different grounds. First, it may be used as a form of two-part tariff discussed in Subsection 3.3. However, for the case of water, it has been argued above that the unit price should really be more than sufficient to cover full costs. It may be desirable to use fixed charges to increase the surplus further, but it is inefficient to use it to lower the unit price to a level lower than MC as this encourages uneconomical consumption of water.

The second argument is that property value is a good proxy for the income and/or wealth level of the owner. Making people with high property values subsidise people with low property values is perceived to contribute to distributional objectives. However, this is a dubious and inefficient way of achieving distributional objectives.

First, distributional objectives should properly be a function of the federal and state government, not of water authorities. Unless there is some delegation of this function to water authorities, the adoption of the distributional role by the latter is of questionable legitimacy.

Secondly, if the government wants to use property values as a base for distributional purposes, it should be the total value of all properties owned by a taxpayer rather than the value of a water-consuming property. Some states already have property taxes. To achieve distributional objectives by levying fixed charges in accordance with property values can be said to involve double taxation of property.

Thirdly, charging for water in accordance with property values is not without efficiency costs. It is true that, because water is an essential good without close substitutes, few consumers will be ‘driven away’. However, a unit price below MC is inefficient. On the other hand, if the unit price is already at a level satisfying efficiency considerations, the additional fixed charges for distribution purposes will have some distortive effect by modifying people’s choice of properties.

It is true that the use of other policy instruments such as progressive income taxation also distorts people’s choice between income and leisure. It is because of this consideration that many economists accept the use of other inefficient instruments to achieve equality. However, unless these instruments are justified on the complicated second-best consideration (with very demanding informational requirements) or some other efficiency considerations (for example, external effects), they actually involve double distortion costs. This is so because attempts to make real incomes more equal through equality-oriented policies (for example, distributional weights in cost-benefit analysis, rationing, rent control, etc.) also distort people’s choices between income and leisure. In addition, they distort people’s choice with respect to consumption and production. Thus, the following proposition has been proved recently (Ng 1984).

Proposition A (A Dollar is a Dollar): For any alternative using a system of purely equality-oriented preferential treatment between the rich and the poor, there exists another alternative which does not use preferential treatment, that makes no one worse off, achieves the same degree of equality (of real income, or utility) and raises more government revenue, which could be used to make everyone better off.

From the above proposition and the associated discussion in Ng (1984), it follows that the water authority (or any other public enterprise, cost-benefit analyst, or public policy decision maker) shall generally conduct its business purely in accordance with the principle of economic efficiency, leaving the distributional issues to the government. This general presumption does not hold only if all the following hold.

(i) It is not illegitimate for the water authority to take over the function of redistribution.

(ii) It is more efficient (less distortion caused) to redistribute by incorporating distributional components into water pricing (for example, based on property values) than through income taxation.

(iii) The government fails to use the more efficient property taxation and cannot be persuaded to do so.

(iv) The resulting equity gain offsets the associated administrative and distortive costs.

Doubting that all the above hold simultaneously, we recommend that water authorities concentrate on the efficiency issues in their pricing policies.

4.2 *Inter-Generational Equity*

Inter-generational equity is a more important issue in water supply than in most other public enterprises because of the extreme lumpiness and long lifespan of water reservoirs, which account for a major proportion of the total costs of water supply. How should the substantial costs of reservoir construction be spread over time to achieve inter-generational equity?

It may seem that inter-generational equity can be achieved simply by charging consumers of each period the full costs of water supply in that period. (The
costs of long-lasting items such as reservoirs must of course be appropriately amortised.) However, this simple solution could be in conflict with economic efficiency which may require pricing at above or below the full-cost level. This conflict is unlikely to be as important as is widely believed.

Due to the second-best and cost-of-revenue considerations discussed above, it is unlikely that economic efficiency requires less than full-cost pricing for many public enterprises. The railway is a possible exception; it supplies a close substitute to road transportation which is subjected to congestion. A low price for rail travel and transportation may thus help to relieve road congestion. But even here, a better solution may be tackling road congestion directly for example, by imposing a congestion tax on petrol (whose consumption also produces other external costs, that is, noise and air pollution).

In the case of water supply, it is argued above that economic efficiency requires more than full-cost pricing. Does this also create a conflict with inter-generational equity? Should the price be reduced somewhat to strike a compromise between economic efficiency and inter-generational equity? Our answer is largely in the negative. This is so because the present generation has the option to vary the amount of transfers to the future generations in the forms of bequests, capital accumulation, borrowings, etc. If the economically efficient price is above the full-cost level, it is better to charge that price, and let individuals and governments make adjustments to their decisions affecting general inter-generational transfers if they so desire.

It may be thought that the above argument is not applicable if the economically efficient price is below the full-cost level on the ground that the future generations cannot decide to have more transfers from the present generation. This asymmetry raises interesting questions regarding the extent to which the above argument is applicable. Obviously, it is not altogether irrelevant since the present generation will perceive its improved position if the price is below the full-cost level and may increase its general transfer to the future. However, since we are dealing with a case where the efficient price is above the full-cost level, we need not be concerned with this question here.\footnote{I hope to pursue other issues related to inter-temporal equity such as the conventional versus Buchanan view on the burden of debt-financing and the relevance of the Ricardo-Barro equivalence theorem on another occasion.}

First version received June 1987; final version accepted August 1987 (Ed.).

Appendix

Mathematical Note: Optimal Lumpy Investment and S-R MC-Pricing with Growing Demand

Let each lumpy investment (for example, construction of a reservoir of a given size) cost an amount A per period and provide a capacity supply K. In addition, there is a constant per-unit cost of supply b. The total cost of supplying $Q_t$ units is thus

$$N_tA + bQ_t; \quad Q_t \leq N_tK$$

where $N_t$ is a positive integer indicating the number of reservoirs in operation at time $t$.

Quantity demanded $Q_t$ at each $t$ is a function of price $P_t$, but growing at a given rate $g$,

$$Q_t = e^{gt}f(P_t)$$

from which

$$P_t = F(Q_t/e^{gt})$$

The problem is to choose a time profile for $Q_t$ (which also implies the choice of $N_t$ and $P_t$) to maximise the discounted present value of the net economic surplus through time:

$$\max_{Q_t} \sum_{t=0}^{\infty} e^{-rt} \left[ Q_t \int_0^t F(z/e^{gt}) dz - (N_tA + bQ_t)\right] dt$$

subject to $Q_t \leq N_tK$.

By construction, the problem is time-independent since, (i) possible time-interdependency in demand and/or supply is ignored, and (ii) demand is postulated to grow at a positive rate such that no constructed reservoir has ever to be wasted. Hence, the maximisation problem can be decomposed into the
maximisation of the economic surplus of each period

\[ \text{Max } Q \int_0^Q F(z/e^{gt})dz - (N_tA + bQt) . \] (5)

Assuming that, at the optimal choice of \( N_t \), \( F(Q_t/e^{gt}) \) is no smaller than \( b \) (that is, consumer willingness to pay at the margin is no smaller than the supplier S-R MC), it is clear that optimal \( Q_t = N_tK \), that is, all built reservoirs should be operated to capacity. (This is due to our ignoring of time-interdependency in supply through meeting future demand by this period’s rainfall.) The choice of optimal \( Q_t \) then boils down to the choice of \( N_t \), the number of reservoirs. It follows that \( N_t \) should be increased (a new reservoir scheduled to commence operation) if and only if

\[ (N + 1)K \int_0^{(N + 1)K} F(z/e^{gt})dz \geq A + bK . \] (6)

If the demand function is known, the exact solution can be obtained by integration. The text deals with a simple case of linear demand curves. Let us now consider the more acceptable benchmark case of a constant elasticity demand function

\[ Q = e^{gt} p^{-1/\alpha} \] (7)

where \( \alpha > 0 \) and \( 1/\alpha \) is the constant (absolute) elasticity of demand. From (7),

\[ P = e^{gt} Q^{-\alpha} . \] (8)

The problem is thus

\[ \text{Max } Q \int_0^Q e^{gt} Q^{-\alpha}dQ - (NA + bQ) \] (9)

subject to \( Q \leq NK \).

The critical values of \( t \) when a new reservoir is just introduced is found by integration which yields, for \( \alpha \neq 1 \),

\[ \{(N + 1)^{1-\alpha} - N^{1-\alpha}\} e^{gt} K^{1-\alpha} = (1 - \alpha)(A + bK) . \] (10)

Equation (10) gives an implicit solution for \( t \) as a function of \( A, b, K, \alpha, g \), and \( N \) (which assumes positive, integer values only). For \( N = 1, 2, \ldots \), (10) solves for the optimal time for the introduction of the 2nd, 3rd, ..., reservoirs respectively. As an illustration, consider the case where \( \alpha = 2 \). From (10),

\[ e^{2gt} = \frac{-(A + bK)K}{(N + 1)^{-1} - N^{-1}} = \frac{(N + N^2)(AK + bK^2)}{(N + N^2)(AK + bK^2)} . \] (11)

For each successive value of \( N = 1, 2, 3, \ldots \), (11) gives the value of \( e^{2gt} \) as \( 2\beta, 6\beta, 12\beta, 20\beta, 30\beta, 42\beta \) respectively, where \( \beta = AK + bK^2 \). These values of \( e^{2gt} \) determine the optimal time for the introduction of the 2nd, 3rd, ..., reservoir respectively.

We turn now to examine the critical times at which the water authority just breaks even, that is, a position like \( D_2 \) in Figure 2. These are determined by the solution of the following equation,

\[ F(Q_t/e^{gt}) = b + A/K . \] (12)

For different integral values \( N_t \), the substitution of \( Q_t = N_tK \) into (12) gives different values of the break-even time corresponding to the operation of different numbers of reservoirs. For the demand function given in (7), we have,

\[ e^{2gt}(NK)^{-\alpha} = b + A/K . \] (13)

For the specific case of \( \alpha = 2 \), (13) gives the values of \( e^{2gt} \) for successive values of \( n = 1, 2, 3, \ldots \), as \( \beta, 4\beta, 9\beta, 16\beta, 25\beta, \ldots \), respectively. Comparing this timetable for the (just) break-even periods with the previous timetable for the introduction of new reservoirs, we obtain Table 1.

From Table 1, it can be seen that the value of \( e^{2gt} \) for the break-even time is exactly halfway between that for the introduction of the existing marginal reservoir and that for the introduction of the next reservoir. Thus, the water authority will only be in deficit for a slightly (depending on the growth rate of demand \( g \)) longer time than in surplus, as illustrated in Figure 3.

The maximum surplus occurs just

| Table 1 Cycles of Deficits/Surpluses Associated with the Introduction of New Reservoirs |
| --- | --- | --- |
| Value of \( e^{2gt} \) at | Introduction of new reservoir | Just break-even time |
| 1st reservoir | 2\beta | 2\beta |
| 2nd reservoir | 2\beta | 4\beta |
| 3rd reservoir | 6\beta | 9\beta |
| 4th reservoir | 12\beta | 16\beta |
| 5th reservoir | 20\beta | 25\beta |
before the introduction of a new \((N + 1)\)th reservoir when the per-unit (of water consumption) surplus is given by

\[
e^{2q} N^{-2}K^{-2} - (b + A/K) = \left( N + N^2 \right)(AK + bK^2)N^{-2}K^{-2} - (b + A/K) = (b + A/K)/N .
\] (14)

The maximum deficit occurs just after the introduction of a new reservoir when the per-unit deficit is given by

\[
e^{2q} N^{-2}K^{-2} - (b + NK) = \left( N - 1 \right)^{\frac{1}{2}} + \left( N - 1 \right)^{\frac{3}{2}}
\]

\[
(AK + bK^2)N^{-2}K^{-2} - (b + A/K) = -(b + A/K)/N .
\] (15)

Thus, the maximum surplus and the maximum deficit are equal in absolute values.

References


