Urban Water Policy: In Need of Economics

Terence Dwyer

The thesis of this article is that Australian urban water supply and pricing policies are not well grounded in rational economic thinking. It questions the policy consensus that water is necessarily so scarce in Australia as to require the sorts of penal charges, water restrictions and use regulations which are being imposed or being considered for imposition upon urban users. It also questions the policy consensus on vetoing new dams for urban water supply, a consensus from which Queensland is partly defecting. (Rural water use is a separate story, though it seems that policy failures there have been used to justify the punishment of urban water users.) This article does not question the function of the price system nor does it suggest capital for water infrastructure is a free good, but it does argue that the current ‘water policy consensus’ which involves deliberate non-augmentation of supply, discriminatory pricing, rationing and the creation and exploitation of monopoly rents by treasuries is hardly an economic optimum.

The thesis is illustrated by pointing out apparent logical problems with Australian Capital Territory (ACT) water supply and pricing policies. The ACT’s policies are more or less paralleled by other jurisdictions and its urban water prices have been driven higher than most. It is therefore a representative example of an increasingly dominant policy consensus.

At the outset, one readily acknowledges that when a resource is genuinely scarce, rising prices do serve to ration demand. However, although Australia is a dry continent, it has a relatively small population and a wide variety of regional climates. Water is not necessarily scarce in North Queensland, for example, and a policy which forbids Cairns from building a second dam when six times the volume of the existing dam flows over its spillway in the wet season may be quite irrational. Few Australians live in areas with very low rainfall.

Basic Propositions

In making this critique of Australian urban water supply and pricing policy one may start with some basic and well-established propositions. One assumes that readers who are economists understand marginal cost pricing, sunk costs and external benefits to serviced land; and that they understand the difference between a true cost of supply and a monopoly rent. One also assumes that, like Böhm-Bawerk (1894-95; 1907-08), they understand the difference between ‘capital’ as a

Terence Dwyer is Visiting Fellow, Asia-Pacific School of Economics and Government, The Australian National University.
factor of production and ‘capitalised value’ representing the present worth of monopoly taxing power. In particular, one trusts readers are not so beguiled by the mysteries of accounting as to confuse ‘profitability’ with ‘economic efficiency’.

Readers will also understand how the optimality rule for marginal cost pricing and minimising deadweight loss in public utility pricing are parallel to the problem of deadweight losses of explicit sales or income taxes (for an overview of these issues from a practical perspective see Larkin and Dwyer, 1995). Economists understand that administered and arbitrary utility prices are really disguised taxes which impose similar deadweight losses on the community.

Proposition A: The environment is not an absolute value.

By this is meant that human activity is allowed to change the environment or even degrade it in some sense if it serves human interests. Were it otherwise, Sydney would be bush today, London would be a marshy swamp and Rome, seven hills. The point is that environmental costs must be quantified and set against environmental and other benefits. For example, it is absurd to count the costs of dams in stopping natural environmental flows if one does not count the benefit of dams in helping to sustain such environmental flows when there would have otherwise been no flow during a drought. Similarly, if one counts the lack of water flow at the Murray mouth as a cost, one has to balance this against having amenity in the city of Adelaide.

Proposition B: If there is no scarcity of water then short-run marginal cost pricing for water usage represents an economic optimum.

We assume here that capital has been invested to collect, store and deliver water and that, once done, available water supply exceeds demand. The only things to be paid for are (i) the current supply cost of water usage, and (ii) the fixed costs of water storage and supply infrastructure. This proposition that short-run marginal cost pricing (for water usage) is optimal can be traced back to Dupuit (see Hotelling, 1938) and is ably defended by Vickrey (1948). It involves the corollary that pricing water usage so as to recover sunk or fixed costs or to generate a rate of return on the replacement cost of immoveable assets is not economically efficient (on objections to marginal cost pricing and answers, see Coase, 1946 and Vickrey, 1948; 1987).

The only real objection to short-run marginal cost pricing is the financing problem. The objection is on the lines that: marginal cost is a useful concept and defines a lower price for producers facing a distressed market situation. If suppliers can find ways where sunk costs are paid by a lump sum, marginal cost also becomes a useful pricing approach. But to require it as a condition of investment or to impose it after firms have sunk their capital, means foregone investment and gradual impoverishment.

In other words, private sector suppliers have to recover sunk costs or they will not invest and for governments to use tax revenues to cover fixed costs of
infrastructure usually creates distortions elsewhere. But as Hotelling (1938) realised, financing infrastructure through land value rates or taxes does not involve deadweight losses (he was in error in thinking deadweight losses were minimal in relation to income taxes). In effect, short run marginal cost pricing for water usage combined with land value rating to cover fixed costs represents a non-distorting two-part tariff.

**Proposition C:** If water does have a scarcity value, then the price of that scarcity should be determined by competitive markets and all users should pay the same scarcity price as a resource rent (with any adjustments for transmission wastage). The ‘law of one price’ should prevail.

Here we assume that water demand exceeds supply. This means that in a river basin, urban users in the same basin as rural irrigators should be charged the same scarcity rent per kilolitre and vice versa. Further, if 30 per cent of water is lost in transmission downstream then a rural irrigator downstream should pay 30 per cent more than an urban or rural user water upstream. In other words, there is a natural comparative advantage to using water where it falls *ex caelis*, which should not be eroded by cross subsidisation through prescribing identical water abstraction charges.

**Proposition D:** Investment in water storage and reticulation infrastructure should be undertaken when social benefits exceed social costs.

Here again, we assume that water demand exceeds supply, given the existing storage and supply infrastructure but that the storage and supply infrastructure can be augmented. In deciding whether to augment, one should look at both public and private costs and benefits. For example, if one attaches a value to downstream or estuary fish habitats and a dam might adversely affect those habitats, the adverse effects have to be weighed against potential beneficial effects such as the dam supplying water downstream in a drought, and avoiding the cost to urban and rural users of a lack of water destroying or damaging parks, gardens, sports fields, crops, cattle *et cetera*. Whether such augmentation comes from new dams, from ground water or buying storage in existing dams elsewhere, are all supply options to be considered but one notes that ‘water trading’ cannot push water cheaply upriver or over mountains from a rural to an urban basin.

**Proposition E:** Land values reflect the value of access to infrastructure.

Land which is serviced by access to water supply will have a higher value even if no water is actually being drawn from the infrastructure. If water is available when you want it (at whatever the marginal cost of supply plus a scarcity rent, if applicable) then your block of land is more valuable than one without access. The recognition of this simple fact was the basis for the Australian system of rating ‘unimproved’ land values (that is, unimproved by the landholder as
opposed to society) which sought to recoup the cost of public works from the land benefited. Insofar as land value rating recoups external benefits conferred on land by infrastructure, it is not a tax nor does it result in deviations of per litre usage prices from optimal short run marginal cost. This may be contrasted with systems of infrastructure pricing which attempt to recover all fixed cost of infrastructure from ‘user pays’ water usage charges levied only on immediate users of what is carried over or through the infrastructure.

**Proposition F:** Land values also reflect environmental costs and benefits of human activity generally.

Environmental degradation lowers land values (for example, salination of pastures). By contrast, human efforts to conserve water and make it available for use may add value to lands. That is to say, there may be trade-offs. For example, is the increased value of cotton land upstream worth the degradation of pastures and estuary fishing areas downstream?

**A Critique of Water Restrictions**

We are now in a position to critique Australian urban water policy. The urban ‘water policy consensus’ which is emerging is to:

1. block construction of new dams (although the Queensland Government shows signs of abandoning the new political and economic ‘correctness’ on this point);
2. rely on detailed regulations and restrictions to limit and control of urban water use (for example, no sprinkling at certain hours and no hosing of hard surfaces or windows);
3. increase water prices generally;
4. impose discriminatory block pricing aimed at punishing ‘large’ household water users; and
5. use water utilities as treasury cash cows through dividends, thanks to the rising cash flows generated by 3 and 4.

None of these policies can be justified by orthodox economic reasoning.

*The veto against new dams*

This violates Propositions A and D. What is apparently going on is that there is an implicit acceptance by policymakers (or policy dictators?) that the environment would be catastrophically destroyed by new dams in all cases.

Perhaps some will argue that new dams are not available as a policy option because of negative environmental externalities. But this does not necessarily apply everywhere. Strangling Sydney does nothing to lessen Murray-Darling salinity problems. Claimed environmental costs should be spelled out in full cost-benefit studies in each case and weighed in alongside urban users’ valuations of
their living amenities, parks and gardens. Where are the cost-benefit studies proving people prefer rivers go into the sea rather than water urban parks and gardens? And is it true that Canberra’s use of water takes water out of the Murray-Darling when domestic water is recycled back into the river anyway? Economics is about the revealed preferences of human beings — environmental preferences are only part of those preferences.

A nice example of the peculiar modern tendency to place other species as an absolute value ahead of human beings was displayed when Environment ACT defended ACT ‘environmental flows’ (The Canberra Times 10 May 2005:9). The argument was that a few endangered fish needed water let out of ACT dams to survive as a species. While most would not wish to exterminate fish, the idea that a few native fish need more water than some 370,000 human beings in Canberra and Queanbeyan must surely rank as bizarre. Australia’s native fresh water fish have evolved to survive the cycles of drought and flood, so why do they need artificial ‘environmental flows’ from dams to survive? Even if one disregards evolution, one must question how any rational cost-benefit analysis could justify a policy of providing 41 gigalitres (ACT Government, 2004:43) in so-called environmental flows for a few fish while limiting the use of 370,000 human beings to a net 17 gigalitres.

If someone thinks fish are more important than human beings he is entitled to that view, but other human beings are not bound to respect it as a moral absolute which must force them to die of thirst, no more than Western societies must accept the ideas that cows are sacred and not to be eaten. To put the perceived needs of a few fish ahead of the conservation of trees, ovals, parks and gardens seems merely an arbitrary exercise in placing the interests of one environment (the fish one) above another environment (the one inhabited by humans). Further, if there were a social consensus in favour of doing so, it would follow that the costs should be borne by all taxpayers, not off-loaded onto water users alone. Presumably the fish cannot pay for their share of the dams (without which there would be zero flow in a drought anyway). But if Environment ACT (as their human proxy) wants water for fish, its budgetary allocation should be charged with that cost and water users given a refund so that they can fund a new dam for human use.

A veto on augmentation of urban water storage is an irrational dogma if there is no attempt to weigh up the social and economic costs and benefits in each case (including weighing up competing environmental claims). A veto on dams may be due to ideology, pantheist views on Nature as Deity, political pressure groups or the vested financial interests of monopoly infrastructure ‘owners’ such as treasuries but a dogma it remains. An economist is duty bound to point out the implied social costs of such a dogma.

Taking the ACT as an example, if the social costs of cutting consumption by quantitative water restrictions are a recurring $71 million per annum, then at a discount rate of 10 per cent, it would pay to spend up to $710 million to build sufficient water storage infrastructure to enable abolition of the restrictions.

The costs of water restrictions in the most recent drought have been estimated by the ACT water corporation (The Canberra Times, 29 April 2005:1) at $71
million and the cost of reducing consumption by a Government-mandated further 25 per cent to 2023 (it has already been cut by 20 per cent per capita since 1993-94 pricing changes) is estimated at $323 million. As the cost of a new Tennant Dam is estimated at $238 million (curiously increased from prior estimates) it makes sense, even on the official figures, to build the new dam.

We thus see, in the case of the ACT a violation of Proposition D. The admitted costs of reducing consumption and imposing water restrictions are greater than the cost of a new dam.

A refusal to build a new dam seems to depend on a denial of Proposition A, an implicit assumption that the environment is an absolute value and nothing further should be done by human hands to alter it. Unless someone can explain what are the negative costs of a new dam, and show they outweigh the costs of not building a new dam, then the violation of Proposition A stands.

When prices rise, they should do more than merely ration existing supply — they call forth new supplies. A rising water price tilts the cost-benefit ratio more and more in favour of investing in a new dam. Basic economics would dictate a new dam be built in response to shortages — unless someone can show that 100,000 water tanks and double plumbing are more cost effective (ACTEW Corporation, 2004a:10, Figure 3-2, shows even the most expensive dam option is cheaper than rainwater tank rebates).

Reliance on regulations and restrictions

Water restrictions necessarily violate Proposition C. They deny the legitimacy of markets in allocating resource use and claim superior economic wisdom for the policy or regulation maker. Rationing obscures, rather than reveals, preferences, just as petrol rationing did in the 1940s. Normally, economists denounce command economy measures to control consumption outside the price signal mechanism. It is therefore surprising there has not been universal condemnation by economists of proposals to regulate in minute detail what water users are allowed to do with the water they pay for.

For example, if you choose to wash down your patio with a hose because it does a better and cleaner job than sweeping (and runs off to water the lawn in the process), who is some officious State bureaucrat to second guess your economic judgment? If you consider that watering your garden for a good soaking 8 hours every week is better than watering it for 3 hours every second night, why should you be fined? You may actually even be using less water in toto!

Water restrictions, like other forms of rationing, are an inherently second-best method of dealing with shortages of supply.

Increasing water prices generally to reflect ‘scarcity’

Perhaps the most intuitively appealing argument for water restrictions has been the dogma that water is ‘scarce’.

But water is not necessarily scarce — contrast rainfall in Cairns with Adelaide (and even Adelaide can draw water from the Murray). Yet some
economists insist water has a scarcity value which must be paid. Fine, then where are the competitive markets to establish true scarcity prices? How else can one say water is necessarily scarce? Or is water supply being constrained by governments blocking the construction of long planned dams? (For an overview of the interaction between environmental flows and supply needs in the ACT, see McIntyre et al, 2003.) How can anyone talk about scarcity prices without examining the causes of scarcity? Does ‘scarcity’ mean OPEC should charge $100 or $200 per barrel of oil? All prices reflect scarcity but are also held in check in a free market by the costs of new supply, so how can the reality of water supply issues be ignored?

If embargos were imposed on power stations, there would be rising scarcity prices for electricity. But no serious economist would accept these were genuine scarcity rents. Economists would quantify the economic costs of distorting markets and forcing business and households to buy their own generators or use candles instead of light globes.

Saying urban water is ‘scarce’ is like saying that office rental space in Sydney would be ‘scarce’ if building codes vetoed any building over one storey high. Like office space, the abundance of ‘water’ as a delivered commodity depends on supply as well as demand — as much on capital investment as natural availability of dam sites or aquifers. Office rents per square metre can fall with new supply even as ground rents and land prices rise. If political correctness or environmental dogma says dams are an inherent evil and should not be built, water may indeed become ‘scarce’ and rationed or expensive. But, in a free market, as the price of office space or delivered water rises, new buildings are erected or old ones refurbished and new water storages are built or water recycled. Both new storage construction and recycling are natural economic responses to scarcity — neither is ‘right’ or ‘wrong’: it all depends on the cost-benefit figures for each response in each case.

When it comes to explicit scarcity pricing for water, the ACT has, in fact, purported to impose such a ‘scarcity price’ (see Independent Competition and Regulatory Commission, 2003a; 2003b; and 2003c). It is embedded in the ‘water abstraction charge’ (WAC) on water taken from rivers or dam storages.

But is urban water necessarily scarce? While one must agree that water is scarce for irrigators drawing water from over-allocated rivers in the Murray-Darling Basin, it does not follow that water is scarce everywhere. Coastal cities storing water which would otherwise flow unused into the sea might question what is meant by scarcity.

Taking the example of the ACT, the most basic criticism of the WAC is that water is not scarce in the ACT. The ACT water utility, ACTEW, states that the ACT has enough water for a million people, after meeting environmental requirements, and is only using (after recycling) about one-sixth of the amount of water available for human use. ACTEW states that some 94 per cent of ACT water goes downstream free of charge to New South Wales, which sells much of it for irrigation use. Even after allowing for mandated ‘environmental flows’ (which
claim over half of the ACT’s water) gross usage in the ACT is only one-third of the water available for human use (ACTEW Corporation, 2004a).

The criticism here is not that genuine water scarcity might lead to water price increases. The criticism is that water prices increases are being allowed to be dictated solely by so-called ‘demand management’ imperatives where there is no genuine economic scarcity. Prices are not rising in a way which is consistent with a free market. They are being set arbitrarily by governments and regulators. In a free market, prices are allowed to respond to scarcity, both to reduce lower-valued usage and to draw forth new investment in water storage or recycling infrastructure.

So-called ‘demand management’ strategies are a euphemism for rationing. It is true that prices should rise when water becomes scarce but that is only the start of the economic story of demand calling forth supply and new equilibrium.

**Adopting discriminatory block pricing to punish large household users**

The intellectual dishonesty of appeals to water price increase as being justified by increasing scarcity becomes apparent when we examine the precise structure of price increases in urban areas. They are often notably non-uniform. For example, in an unsurprisingly titled media release, the ACT’s Independent Competition and Regulatory Commission (ICRC, 2005) has announced its approval of discriminatory increases in water prices from 1 July 2005. The fixed annual (access) charge drops from $125 to $75. Consumers pay 58 cents per kilolitre (kl) for the first 100 kls (previously 51.5 cents); $1.135 per kl for the next 200 kls (previously $1); and $1.53 per kl for annual consumption over 300 kls (previously $1.35). These charges are separate from the ACT Government’s water abstraction charge (which is supposed to represent the scarcity value of water as a resource and is also increased from 20 to 25 cents per kl from 1 July 2005).

Now an economist should ask some basic questions.

- Aren’t fixed charges preferable as ‘lump sum taxes’ to excises on output when one is looking at excess burdens of taxation? Wouldn’t even second-best Ramsey pricing dictate that the fixed charge should be as high as possible to minimise the ‘access deficit’ or that there be declining, not rising, block tariffs, as used to occur with gas pricing? (Berg, 1998). In fact, the ideal two part tariff charge is an access price based on a ‘lump sum’ rating of the value of the serviced land, as used to be the case in Australia, plus a uniform charge solely representing marginal cost, which would include a uniform scarcity rent (if there is real scarcity).

- If we are increasing prices on large users because of scarcity and the supposed ‘need’ to reduce consumption, why is that not reflected in the water abstraction charge for all users rather than in the tariff which generates a return on capital and recurrent costs for the utility which provides infrastructure? Is it really socially costlier to supply 10 large families consuming 500 kls each than 50 smaller households consuming 100 kls each?
One might have thought economies of scale operated in infrastructure supply — installing, maintaining and reading 10 connections is surely cheaper than for 50. And what about economies in consumption? A household of 6 people may use less water per capita than 6 households of one person.

- If the justification for price increases is that water is scarce (as ICRC appears to argue), then why aren’t all users to be charged the same price per kilolitre for all amounts used. Is marginal water used by a small high-income household less precious or scarce than water used by a large struggling family?

- Why should ACT households pay up to $1.53 plus a tax of 25 cents per kl when ACT water is being sent downstream free of charge to irrigators paying far less? It is curious environmental policy to turn a garden city into a desert so that rice can be grown downstream in a desert. In a free market the law of one price would prevail. Only monopolists can sell the same product at different prices to different buyers.

Thus discriminatory block pricing violates Proposition C. It also violates Proposition B, as the scarcity value of water should enter as a uniform marginal cost.

Using water utilities as cash cows

When confronted with the new (and heretical) public finance fashion that all public services should be provided on the basis of ‘user pays’, most consumers, in their naiveté, assume that ‘user pays’ means:

- that they are charged the cost of what is supplied; and
- that they will get what they have paid for.

The use of utilities as cash cows defeats these humble expectations. What has been blessed (by ignorance or wilful negligence or treasury subornation on the part of pricing regulators) is multiple charging for public works.

Once upon a time, dams were constructed and financed by the perfectly sensible system of rating land values to amortise Water Board loans incurred to service the land. Land developers often had to contribute infrastructure free of charge to the water authorities. It made, and still makes, economic sense for the fixed costs of network infrastructure to be charged against the land values being benefited. Then the flow of water through the system can be charged at its low marginal cost, ensuring optimal use.

Thus, suppose a town was built over 90 years and dams and water works and pipes were financed by levies on ratepayers. A normal person, and a genuinely rational economist, might think that these capital works had been paid for and therefore should not be charged for again.

Yet ‘reform’ has meant water users are being asked to pay again for costs already recovered or never borne by State Treasuries. To take just one example,
the Queensland Government has sought to charge a rate of return on the Burdekin Dam which was paid for by the Federal taxpayer. As Professor Bob Walker (1993) noted some years ago, there are not many businesses where you can get your assets given to you by taxpayers or consumers and then turn around and demand a return on money you never had to outlay.

The conventional wisdom underlying the Council of Australian Governments and National Competition Policy water reforms is:

(i) Water authorities in Australia were unprofitable and delivered a low return on the ‘investment’ governments had made in them.
(ii) Hence, they should be made to deliver a rate of return to their government ‘owners’ equal to the rate of return on private sector equity investments.

Those assertions, constantly repeated, have assumed the status of urban legend, if not the status of self-evident truths. Both assertions are wrong.

Leaving aside whether water schemes were undertaken as ‘investments’ or were created as public utilities for public benefit, as Walker and Walker (2000) and Walker (1993) have pointed out, public sector accounting for water authorities has often employed techniques which result in misleading figures showing low profitability when the reality is that water authorities have been more profitable than most listed industrial companies.

Walker and Walker (2000:87-88) point out that

many of those GTEs [government trading enterprises] were highly profitable by private sector standards, but had only reported low levels of profitability. This occurred because Australian GTEs were required to adopt radical methods of accounting — methods not used anywhere else in the world ... there are some important differences in the way private sector and public sector entities value their assets. It is not widely recognised that since the late 1980s, Australian GTEs have used a system of accounting which is radically different from that used in private sector accounting. This has produced radically different financial results. (emphasis in original)

Walker and Walker (p. 97) go on to observe that

the major omission [of the Steering Committee Guidelines on accounting policy for valuation of assets of Government Trading Enterprises] was a failure to recognise that recent advocacy of the use of current replacement prices was that for consistency, the amounts by which asset values were increased should be brought to account in the operating statement as revenues (or unrealised gains). Adoption of this model of ‘clean surplus’ accounting would have meant that poor rates of return would have been converted into good rates of return. As it happened, the end result of the Working Party’s ill-informed activities was that Australian GTEs were told to adopt a system of accounting
which produces figures for ‘profit’ and ‘rate of return’ which differ substantially from the figures which would be produced by private sector firms using private sector accounting methods. ... The Steering Committee had set out to ensure that the accounting methods used by GTEs would enable comparisons to be made between the government-owned businesses and ‘comparable’ private sector firms. They ended up promoting a system that ensured exactly the opposite’

It was also forgotten that (Walker and Walker, 2000:104-05):

A significant proportion of the assets of water authorities has been acquired through compulsory ‘donations’ from property developers. ... From an accounting perspective, the receipt of these ‘donations’ meant that water authorities had to record an increase in their assets — but most authorities recorded the receipt of these assets as an increase in ‘reserves’ rather than a source of revenues — the treatment indicated by the accounting profession’s statements of accounting concepts. Having recorded increases in assets, water authorities then wrote-off those assets through depreciation charges (which were treated as expenses, and hence reduced reported profits). The combination of these treatments meant that the more donated assets received by those GTEs, the lower their reported profits.

The net result is that, after adjusting for revaluations and placing accounts on a common historic cost basis ‘water authorities were far more profitable than listed industrial companies.’ (Walker and Walker 2000:106). Yet, increasingly, a mantra of the need for ‘forward-looking’ prices is invoked to justify charging consumers on the basis of what it would cost to replace these assets in their current state. Hence consumers should be charged on the basis of ‘depreciated optimised replacement cost’ (DORC). (For other critiques of DORC, see Johnstone and Gaffkin, 1996; Johnstone and Wells, 1998; King, 1996.)

DORC may lack economic merit but it has wonderful features for infrastructure owners, such as State and Territory treasuries. One can blithely ignore the past history of who financed and contributed to public works and proceed to write up their value and claim a required rate of return on the capitalised value. To reap where he never sowed is the dream of every landlord, as Adam Smith observed. To charge for use of an asset you never paid for must rank alongside it as the dream of treasury officials and infrastructure monopolists.

The effect of such artificial and contrived accounting is to enable water utilities to generate astounding cash surpluses on cash actually invested while showing apparently low rates of return (which are used to justify claims for ever-higher prices). The net effect is the State and Territory treasuries can look forward to higher ‘dividends’ and pretend that this is ‘user pays’ financing ‘reform’ rather than the elaborate monopoly tax farming which it really is.

Worse still, unlike normal competitive industries, none of the new-found ‘profitability’ of water authorities is leading to a rush of new investment in much-
needed water storage infrastructure. Quite the contrary, State governments are trying to find reasons for forcing a massive and uneconomic duplication of water infrastructure, as each home has to fund its own mini-reservoir. (It may be socially more efficient to build a new dam and save the community the potential health and financial costs of millions of tanks, but high private costs are not borne by State Budgets.)

If none of the excess profits are ever ploughed back into additional infrastructure, of course water prices must rise towards infinity. But this is the result of State government policy. It should not be blamed on Nature — ‘Accuse not Nature, she hath done her part; Do thou but thine’ (Milton, Paradise Lost VIII:561).

Conclusion

There is no economic justification for spurious water pricing ‘reforms’ which strip excess profits from urban users, block increased supply, discriminate arbitrarily and push up business and living costs in the process.

Urban water pricing in the ACT and other places is not being determined by genuine economics. It is a form of thinly-disguised excise taxation enforced by a blockade of new supplies. Urban water pricing reflects abuse of monopoly power and lucrative monopoly rents, rather than true resource scarcity rents. It is as trivially beside the point to say that urban water pricing should reflect scarcity as to say the price of bread in Berlin in 1918 reflected scarcity rather than the Allied blockade.

The real hidden issues in the new urban water ‘political correctness’ are the blockade of new supply and arbitrary rationing of existing supply. Those willing to defend this situation are at odds with orthodox economics. While one can understand the revenue desires of treasuries, the self-interest of would-be private water infrastructure owners, the ideology of environmental Puritanism and the love of politicians for redistributing income, none of these amount to a valid economic consideration in pricing urban water. Urban water supply and prices should not be set by an Australian ‘water OPEC’ of State governments acting like the monopolists justly despised by Adam Smith (1776-1784) when he wrote:

The monopolists, by keeping the market constantly under-stocked, by never fully supplying the effectual demand, sell their commodities much above the natural price, and raise their emoluments, whether they consist in wages or profit, greatly above their natural rate.

References

Urban Water Policy 15


Engineers Australia (2004), ‘Engineers Debunk the Water Myths’, Canberra.


The author acknowledges the help of Ross McLeod who thought the subject worthwhile as well as the comments of two anonymous referees who, without necessarily agreeing with the paper, felt some points were obvious. I hope they are correct and urban water policy undergoes a rethink. A fuller annotated copy of the paper is available on request from the author.
Urban Water Pricing

Hugh Sibly

A water crisis looming in some Australian cities. In Sydney it is currently feared that without significant rainfall the city will run out of water within a few years. In the short term water restrictions have been implemented. Public attention has focussed mainly on technological solutions to the shortage. In Sydney there is debate over whether the water network should be expanded using either recycling or desalination plants. Similar issues are arising in other cities. The use of efficient pricing to ration available water resources, however, has received little attention. In contrast a number of water authorities have championed the use of increasing block tariff (IBT) as a means of conserving water. This paper considers the merits of these pricing reforms.

In Sydney, as in most jurisdictions, the method of retail pricing of water has varied over time. Prior to 1993-94 the then Sydney Water Board adopted a form of IBT. Under IBTs the water provider sets a number of tiers of consumption levels. A low volumetric rate is set in the lowest consumption tier (tier 1) with the rate increasing in subsequent tiers. Prior to 1988-89, Sydney households were given a free allocation of water (that is, tier 1 had a zero volumetric rate). This was replaced by a three tier IBT from 1990-91 until 1992-93 (Sydney Water Board, 1994) and by a two-part tariff in 1993-94 which continued to apply until June 2005. A two-part tariff consists a fixed annual access fee (the fixed charge) and a single volumetric charge which is applied to all consumption. The nominal volumetric rate was 65c/kl in 1993-94 and had increased to 1.013$/kl by 2004-05. An IBT was re-introduced for 2005-06 and is to apply until at least June 2009.

Urban water is priced differently to other goods. At times of shortages — notably droughts — restrictions have been used to control consumption. In contrast, consumers are accustomed to significant variation in the price of virtually all other goods in response to supply side shocks. The recent history of oil prices is a good example. It is often argued that water is unique and should not be priced like other goods. This, as discussed below, is the underlying argument for the use of IBTs — a pricing mechanism rarely used outside the urban water industry.

Implementation of IBTs is possible only because the use of a fixed network to deliver water allows water authorities to tag customers’ consumption. IBTs cannot be used without tagging of customers. For example, suppose service stations sought to implement an IBT, say charge $1 per litre of petrol up to 20 litres (tier 1) and then $2 for each additional litre (tier 2). The likely customer response would be to purchase no more than 20 litres at a time. Thus, the IBT would be ineffective as no customer would purchase petrol in tier 2.

A network is used to distribute water to urban residents because it is the least cost method of doing so. However, should the concomitant capability of providers

Hugh Sibly is a senior lecturer in the School of Economics University of Tasmania
to measure customers’ consumption levels be used in implementing IBTs? The economics literature demonstrates that a two-part tariff is the efficient way to price goods delivered by a network (see Carlton and Perloff, 2005:664-9; Sibly, 2005). The key requirement for economic (allocative) efficiency is that the volumetric charge should equal the opportunity cost of water. The volumetric charge has to be flexible in the face of changing cost and demand conditions. In particular, the charge should rise in periods when water is relatively scarce, or when it is expected to be relatively scarce. This paper compares the implications of adopting an efficient two-part tariff with those of adopting IBTs and argues that an efficient two-part tariff is better than IBTs in terms of fairness, efficiency and conservation.

Recent developments in Sydney’s urban water supply provide a useful setting for the discussion. Although the states use different methods to regulate and manage their urban water supplies, Sydney’s problems are the best documented and, at the time of writing, seem the most acute. Other states are referred to when appropriate. Edwards (2005) considers related issues with the implementation of IBTs in Melbourne and is critical of their introduction in that jurisdiction.

The data on storage levels shows that in 1994-95 Sydney storages were, on average, at 62 per cent of capacity, while in 2000-01 they were on average at 90 per cent of capacity. The seeds of the current water crisis can be traced to 2002. Storage was approximately 80 per cent in late 2001, but had fallen to 60 per cent by early 2003. On 5 June 2005 storage was at 39 per cent of capacity. The current concern on storage levels arises because Sydney’s storages hold, when full, less than 4 years supply at current consumption rates.

**Figure 1: Sydney Urban Water**

![Figure 1: Sydney Urban Water](image)

Source: Data provided by Sydney Water, Sydney Catchment Authority.

Figure 1 shows the annual percentage change of available water as a fraction of total system capacity over the last 11 years. The striking feature of this series is its wide variation across years, which reflects the underlying variation in the average annual storage levels of the Sydney water catchments. Figure 1 also shows the growth rate in Sydney’s average metered residential consumption. Up
to 2003 changes in consumption appear to move in the opposite direction to the changes in storage. This pattern is consistent with the demand for water, particularly for outdoor use, being inversely related to rainfall. The fall in average metered residential consumption since 2003 reflects the imposition of increasingly strict restrictions affecting outdoor water use.

The market response to a reduction in supply would be an increase in the price per unit, which in the case of urban water is the volumetric rate set by water regulators. In NSW the regulator is the Independent Pricing and Regulatory Tribunal of NSW (IPART). The growth in the real volumetric rate is also shown in Figure 1. There is no apparent relationship between changes in available storage and the volumetric rate suggesting that the rate has not played a material role in managing demand at times of water shortage. The volumetric rate increases reflect IPART’s move away from property based fixed charges toward the implementation of a two-part tariff in the wake of National Competition Policy (NCP) guidelines for ‘consumption based pricing’ of urban water.

Following the NCP guidelines, NSW, Queensland, Tasmania and Victoria adopted two-part tariffs. However, concerns over increased water scarcity led NSW and Victoria to abandon two-part tariffs with a single volumetric rate and adopt IBTs. This move is presented as a fair and popular response to the water crisis. In NSW, an IPART survey found that 63 per cent of respondents believed IBTs were fairer than the (then) current two-part tariff. The Victorian Government’s white paper states that IBTs are ‘widely regarded as the fairest and most effective way to price water for conservation’. Both SA and WA have had IBTs in place for some time. The SA Government Committee Water Proofing Adelaide (2004:32) stated that a ‘tiered pricing system is considered by some as a fair and effective way of reinforcing the need for conservation…’.

The Council of Australian Governments (COAG, 2004) National Water Initiative (NWI) calls for best practice in water pricing and ‘improved pricing for metropolitan water’. Although there is no clear statement of what is best practice in urban water pricing, it is argued in this paper that IBTs are inferior to two-part tariffs. Furthermore, they are inflexible in the face of the erratic nature of water availability in Australia. (Boland and Whittington, 2000, de-bunk the arguments in support of IBTs in an international context.) The apparent support for IBTs arises both from a confusion on the appropriate level of the volumetric rate in a two-part tariff and from a political response against raising the cost of water.

This paper uses economic analysis to clarify these issues and argue that two-part tariffs are the best pricing mechanism. However careful attention must be paid to the setting of the volumetric rate. Specifically, in contrast to the setting of IBTs, the volumetric rate needs to be adjusted to current and expected water availability. Section 2 of this paper discusses efficient pricing methodology, with emphasis on the efficient volumetric rate. Section 3 considers how the methodology to determine the present volumetric rate is different to that required to achieve efficiency. Section 4 considers the impact on efficiency of the use of IBTs. Section 5 discusses implications of the paper’s analysis.
Economic Efficiency and Pricing

This section uses a simplified economic framework to derive efficient pricing by urban water authority. The framework is not intended to capture every detail of the operations of an urban water network. It is, however, a useful representation of the salient features of a hypothetical water network that identifies underlying efficient pricing principles, particularly the determinants of the efficient volumetric rate. The model assumes a network of fixed capacity. It also assumes there is no alternative use (for example agricultural use) for water flowing into the network and that waste water is released into the sea (so it has no further use). These assumptions avoid introducing complications of non-urban uses of water, which have no material impact on the pricing principles. This scenario is not necessarily unrealistic. For instance, it might broadly represent Sydney if no capacity augmentation (for example, recycling or desalination) is pursued.

Figure 2: Efficient Pricing Over Two Periods

A two period analysis is used to describe the principles for efficient pricing (see Figure 2). Period 1 represents the present and period two the ‘future’. Water volume is shown on the horizontal axis and the volumetric rate on the vertical axis. The maximum capacity of the network in both periods is $Q_{\text{max}}$. MC is the short run marginal cost curve at output levels below $Q_{\text{max}}$ and represents the cost to the water provider of supplying an additional kl of water at each production level. Demand in period 1 is $D_1$, while expected demand in period 2 is $D_2$. 
It is assumed that there are no capacity constraints in the network. The only constraint is availability of water. Specifically the network has the capability of delivering more water than is available. It is assumed that in period 1 there is an existing water storage level of $S_1$ and rainfall causes an inflow of $I_1$. The total quantity of water available in period 1 is thus $S_1 + I_1$. Expected inflow in period 2 is $I_2$. It is increasingly recognised that a certain fraction of available water in a catchment must be retained in order to maintain the integrity of the ecosystem. Assume that water authorities decide that a volume of water, $E_i$, is retained in each period to support the environment. Furthermore, a volume, $R_1$, is retained in period 1 for use in period 2. That leaves an amount $Q_1$ to be consumed in period 1. Similar allocations must be made in period 2 (although it is assumed no water is retained in period 2 as it is the last period). Of the water retained in period 1, there is assumed to be a fractional loss of $1 - \delta$, so that $S_2 = \delta R_1$.

The available water is allocated across periods on the basis of that available in period 1 and that expected to be available in period 2. The expected volume in period 2 is equal to the sum of storage and the expected inflow, less environmental flows ($S_2 + I_2 - E_2$). A formal derivation of the efficient allocation of quantity across periods (or equivalently the storage level) is beyond the scope of this paper. However, it can be assumed that the usual influences on efficient inter-temporal allocation apply. Thus it is assumed that the efficient allocation satisfies $Q_2 = \rho Q_1$, where $\rho$ depends on the rate of time preference, the rate of population growth, technology efficiency improvements in water use and the rate of water lost in storage $\delta$. For example, an increase in the rate of population growth would see an increase in $\rho$, and thus the size of $Q_2$ relative to $Q_1$. Similarly an improvement in the technology of water use, such as dual flush toilets, reduces $\rho$.

The efficient volumetric rate generates an efficient inter-temporal allocation. In particular, the efficient volumetric rate is the greater of the market clearing rate in that period or $MC$. This ensures that the volumetric rate is equal to the opportunity cost of water. As shown in Figure 1, when water availability is low, the efficient volumetric rate clears the market. The gap between the volumetric rate and $MC$ represents the scarcity value of water. Having made this observation, the efficient response to a change in environmental conditions can be determined. For example, a decrease in current rainfall reduces the current inflow — this has the effect of decreasing the availability of water, thus it is efficient for both $Q_1$ and $Q_2$ to fall. This results in an increase in the efficient volumetric rate in both periods. Similarly the anticipation of a drought in period 2 reduces $I_1$ and thus water availability in period 2. The efficient response increases storage in period 1, reducing $Q_1$. Again, the efficient volumetric rate in both periods increases.

The key conclusion from this analysis is that the efficient volumetric rate should reflect water current and expected availability of urban water. With improvements in long range weather forecasting, meteorological analysis can provide increasingly accurate predictions of seasonal rainfall patterns. Hydrologists can use this information to predict future water availability, which should then be reflected in the volumetric rate. Such a scheme has been suggested as a way of increasing crop yields (Naylor et al., 2002). An increased probability
of drought (or continuation of drought) should see a rise in the current volumetric rate. Similarly an increase in the environmental allocation should see an increase in the volumetric rate. In cities that are growing, the volumetric rate should also reflect the increased demand caused by population growth.

National competition policy requires government-owned businesses to satisfy competitive neutrality (gain no special advantages from government ownership). Specifically, water authorities should not be subsidised from general government revenue. This implies that they should be regulated to give a normal economic rate of return. They must cover all their costs, including the cost of capital. Because infrastructure is the most significant cost to urban water authorities the revenue raised by the efficient volumetric rate is unlikely to be sufficient to cover the operating costs. However a two-part tariff, consisting of a fixed access charge and the efficient volumetric rate, can be used to achieve both economic efficiency and satisfy this revenue requirement. This is possible because the fixed access charge does not affect water consumption nor does it significantly affect disposable income. The appropriate fixed charge in the two-part tariff is thus set to ensure the water provider’s revenue is just sufficient to cover its costs.

Equity considerations are often used to oppose two-part tariffs, particularly if the volumetric rate is perceived to be high. While efficiency requires a volumetric rate equal to the opportunity cost of water, it does not require that the fixed charge be equal across consumers. Thus equity issues can be addressed by varying the fixed charge levied on different classes of consumers. Disadvantaged consumers (as determined by government social policy) could pay a lower fixed charge with water authorities receiving commensurate compensation from the government either in direct payments, or by the government (as owner) accepting a lower than market rate of return from the authority.

Pricing Methodology

The implementation of NCP has seen a movement toward consumption based pricing. IPART(2004:6) considers that ideally the volumetric rate should be set efficiently. How the efficient volumetric rate should be determined has been controversial even at times of plentiful supply. In Figure 2, plentiful supply of water would occur when \( S_1+I_1 \) and \( S_2+I_2 \) are both greater than \( Q_{\text{max}} \). Many academic writers, including Hirschliefer et al (1960), Darr et al (1976), Ng (1987), and King and Maddock (1996) have argued for ‘short run marginal cost (SRMC) pricing’ when water is plentiful. SRMC is the measure of marginal cost when capacity is fixed. Under SRMC pricing the volumetric rate should be set to the level where the demand curve intersects the MC curve. However, IPART and other Australian water regulators argue that long run marginal cost (LRMC) of supply is the efficient volumetric rate. This is consistent with some overseas thinking. The consulting firm London Economics (1997), and subsequently the UK’s Office for Water Services, have argued that LRMC is the efficient volumetric rate. The American Water Works Association (2000:120) states that
‘economic theory suggests that water rates be set equal to long run marginal cost to ensure efficient allocation of water services’.

In economic texts, LRMC is the measure of marginal cost when capacity can be varied incrementally. It is the cost of an extra litre of water when all factors of production (including capital) are varied optimally (Call and Holahan, 1983:258-62). LRMC pricing is efficient in those circumstances. In other words, LRMC is the efficient price only when capital is adjusted to its optimal level. But as textbooks note in those circumstances, short- and long-run marginal cost coincide.

Divergences of opinion between academics and regulators on efficient rates seem curious, given that the textbook analysis of marginal cost is well understood by the economics profession. This controversy appears to stem from observations that investment decisions of water utilities are ‘lumpy’ rather than incremental (Hirschliefer et al:1960) and subsequent interpretation of the analysis of marginal cost by Turvey (1971; 1976) who argues that investment decisions by public enterprise are best thought of in a dynamic context. An incremental increase in the growth of output can be accommodated by incrementally advancing the utility’s investment plans. Thus, at any one time the availability of infrastructure may be thought of as varying incrementally in a textbook ‘long run’ fashion. Pricing and investment decisions should go hand in hand, with the efficient outcome being the combination of a short run pricing rule and optimal investment rule (1971:74). In Turvey’s work the timing of investment in infrastructure is varied so that a volumetric rate equal to LRMC is sufficient to utilise existing capacity. Thus in each period, under the optimal investment plan, $MC=LRMC$.

Turvey’s argument is both ingenious and subtle. However there are two problems with its application to current Australian circumstances. It assumes that the optimal investment rule is in place, and capacity may be fully utilised when required. Neither of these requirements is usually true for Australian water utilities. First, many prospective infrastructure developments face significant uncertainty and political controversy. With such planning difficulties, projects are likely to be delayed or even cancelled. Even if this were not so, the investment decisions and pricing decisions are made by different authorities. In Sydney’s case, the Sydney Catchment Authority is charged with infrastructure development, while IPART regulates retail prices. This discourages coordination of pricing and investment decisions. It is hard to see how the optimal investment rule could in practice be implemented in a manner envisaged by Turvey. Second, and most notably, the erratic nature of rainfall patterns in Australia means that it is quite common for water utilities to be unable to fully utilise their network’s capacity.

The impact of water authorities imposing LRMC pricing in current (water shortage) circumstances can be seen using Figure 2. An increase in capital in period 2 increases $I_2$. (Assume the expansion is a desalination plant that does not rely on rainfall to operate.) Capital should be increased in period 2 until the (market clearing) volumetric rate is equal to LRMC. Thus, with this expansion in capital, the volumetric rate would be lower than it would otherwise have been. However, if for some reason the expansion does not take place, the efficient volumetric rate is $p_2$, not LRMC. In this case LRMC underprices water.
Observe that an expansion of capacity in period 2 enables less water to be stored in period 1. Thus $R_1$ could be decreased and the volumetric rate in period 1 lowered. However, such a strategy requires that the additional capacity be available in period 2, otherwise a shortage in period 2 ensues. Again, the reliance on LRMC pricing is likely to lead to an inefficiently low volumetric rate if for some reason new capacity is delayed or not installed.

Under current Australian institutional arrangements, it is thus reasonable to assume that investment plans are independent of the pricing decision. In this case LRMC does not give a good approximation to the efficient volumetric rate, even with stable availability of water and demand growth. With erratic availability of water, the price stability provided by LRMC pricing works against efficient water allocation. In periods of drought LRMC significantly underprices water because it fails to reflect opportunity cost. At times of high storage levels and low demand water is overpriced (price is greater than MC). In addition, calculation of LRMC is difficult in practice. It is therefore simpler and more efficient to determine the volumetric rate using the methodology discussed in the previous section.

**Effect of IBTs**

In 2005 IPART moved to a two tier IBT to regulate the volumetric rate charged by Sydney Water. The regulated volumetric rates have been set until June 2009. These are shown in Table 1. IPART (2004) sets out the reasoning for adopting this particular two tier IBT. The first tier of the volumetric rate is intended to equal LRMC. Thus the volumetric rate in tier 1 is roughly the existing volumetric rate adjusted for inflation. The volumetric rate in tier 2 is intended to ‘send strong conservation signals’, and thus, as shown in Table 1, is moving over time towards being 1.5 times the tier 1 volumetric rate. Tier 1 applies to consumption up to 400kl per annum. The setting of the tier 1 boundary is to allow ‘non-discretionary use’ to be paid for at the tier 1 volumetric rate. It should be noted that the average consumption of a 5 occupant household is 370 kl per annum and 6 occupant house 408kl per annum (see Table 2).

<table>
<thead>
<tr>
<th></th>
<th>to 30/06/2006</th>
<th>1/07/06 – 30/06/07</th>
<th>1/07/07 – 30/06/08</th>
<th>1/07/08 – 30/06/09</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1</td>
<td>1.20</td>
<td>1.23 x (1+ΔCPI1)</td>
<td>1.26 x (1+ΔCPI2)</td>
<td>1.31 x (1+ΔCPI3)</td>
</tr>
<tr>
<td>Tier 2</td>
<td>1.48</td>
<td>1.59 x (1+ΔCPI1)</td>
<td>1.72 x (1+ΔCPI2)</td>
<td>1.85 x (1+ΔCPI3)</td>
</tr>
<tr>
<td>Tier 2/Tier 1</td>
<td>1.23</td>
<td>1.30</td>
<td>1.37</td>
<td>1.41</td>
</tr>
</tbody>
</table>

Source: IPART (2005)

The efficiency of IBTs is assessed using Figure 3. Under an IBT each connection pays a volumetric rate $t_1$ per kl for consumption up to $T$ kl, then a rate of $t_2$ per kl for consumption beyond that level. Suppose, for simplicity, there are only two types of customers. One type has the low demand $D_L$ and the other has
the high demand \( D_H \). Under the IBT low demand consumers consume \( q_L \) and high demand consumers consume \( q_H \).

**Table 2: Average Annual Household Consumption**

<table>
<thead>
<tr>
<th>Household size (no. of occupants)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption per occupant (kl)</td>
<td>142</td>
<td>114</td>
<td>89</td>
<td>76</td>
<td>74</td>
<td>68</td>
</tr>
</tbody>
</table>

Source: IPART (2004:15)

**Figure 3: Increasing Block Tariffs**

Figure 3 shows that the inefficiency of an IBT depends on the efficient volumetric rate in that period. The efficient rate represents the opportunity cost of water. The loss of economic efficiency of IBTs can be measured by the deadweight loss (DWL), which represents the loss of potential net benefit of an inefficient price. When water is abundant, suppose the efficient volumetric rate is \( p_1 \). If the volumetric rate is \( p_1 \) the high demand consumer uses \( q_1 \). The loss of net benefit (or DWL) is measured by the difference between the height of the demand curve and the opportunity cost of water \( p_1 \). In this case DWL associated with the consumption of high demand consumer is \( e+d \). Similarly low demand consumers have a DWL equal to the area \( a \). If water becomes less available, the efficient volumetric rate would rise. Consider the efficient volumetric rate \( p_2 \), which lies between \( t_1 \) and \( t_2 \). In this case the DWL associated with a low demand consumer is \( b \) while that of high demand consumer is \( e \). Finally if the efficient volumetric
rate rises to \( p_3 \) the DWL from low demand consumers is \( b+c \) and that from high consumers is \( f \). This analysis indicates that an IBT can never be efficient when there are heterogeneous customer demands. At every efficient level of the volumetric rate there is a DWL created by an IBT.

A common argument is that IBTs do not create inefficiencies because the IBT reflects the increasing marginal costs of the water provider (or increasing scarcity of water). This argument is spurious because each individual user has a negligible effect on total consumption and therefore on the marginal cost of the provider. However, a variant of this argument is valid when there are common demands and thus demand curves shift simultaneously. For example, consider seasonal shifts in demand. Figure 3 can be adapted to consider this circumstance. Suppose now DL represents the common winter water demand and DH the common summer water demand. Suppose that \( t_1 \) and \( t_2 \) are the efficient volumetric rate in winter and summer respectively. In this case IBTs result in an efficient allocation of urban water. However, for this result to hold demand must be homogeneous and there must be no annual variation in the efficient seasonal volumetric rate. Neither of these conditions holds in practice. Furthermore, the same outcome could be achieved more simply with a two-part tariff with an efficient volumetric rate.

The focus in the literature on seasonal prices derives from European and North American applications, where water utilities face a capacity constraint at peak times. Thus seasonal prices are directly related to peak load pricing theory issues. However this literature is not of central importance to current Australian applications where the concern is erratic variation in water availability. IBTs (or any form of peak load pricing) do nothing to address shortages that arise because of this variation, primarily because they do not vary with availability. An IBT that fully utilises capacity at times of high availability does not also ration water at times of low availability. In spite of popular perception, IBTs are not flexible enough to solve the water crisis and also be efficient when the drought breaks.

Water authorities have noted the inefficiency of IBT. IPART (2004:20) notes that in its proposal ‘tier 2 charges would not be chosen for economic efficiency reasons … there could be merit in using it to send an additional signal to those residential customers who use a high volume of water, to encourage them to reduce their discretionary use of water’. This ignores economic reasoning, in particular that water consumption, discretionary or otherwise, confers a benefit on households. There is no reason to believe that the 401th kl confers less benefit on a 7 person household than the 300th kl confers on a one person household. In fact, this example demonstrates how an IBT leaves unexploited gains to trade between households. If the two households were next door to each other, they could gain by running a hose from the 1 person household to the 7 person household!

Penalising high use consumers under IBTs is thus justified on equity and conservation, rather than efficiency grounds. However, the assumption that high users are necessarily wasteful is contradicted by the data in Table 2 on average consumption per occupant for households in the Sydney, Blue Mountains and Illawarra regions for 2003. The figures show that consumption per occupant declines as household size increases. This is consistent with economies of scale in
water use. Indeed the consumption per occupant for the 6 person household is less than half that of the single person household. Yet IBTs will cause the larger household to face a higher marginal rate than that of the smaller one. This is neither fair, nor serves conservation.

**Discussion**

Economists teach first year undergraduate students that monopolies set inefficiently high prices, and that price controls can mitigate this exercise of monopoly power. The NWI (COAG, 2004:66) calls for upper bound (maximum) pricing to avoid water providers making monopoly rents. However this regulatory focus on the natural monopoly aspect of urban water authorities has deflected attention from the efficient price of water itself. In major metropolitan areas of Australia the problem is not that water is overpriced because of monopoly power, but rather that its volumetric rate is systematically underpriced.

The sources of the current underpricing are manifold. In its implementation NCP has focussed more on cost recovery and technical efficiency than allocative efficiency. This reflected the concerns of providers and regulators, who have focussed most of their attentions on cost recovery, and how its achievement affects various household types. In fairness, regulators have also concerned themselves with allocative efficiency issues, but the methodology used to determine the efficient volumetric rate is not appropriate for Australian conditions.

However, as is often the case with government business enterprises, the problem in achieving allocative efficiency is the political dimension to water pricing. Both the LRMC and IBT methods are relatively politically appealing. As argued above, a volumetric rate based on the LRMC methodology is both stable and, in effect, low. Boland and Whittington (2000:234) argue that IBT allows water utilities ‘to deliver cheap water to the middle- and upper-income groups while appearing to serve the poor’. In Australian, IBTs have the effect of insulating typical consumers from facing the cost of decreased water availability.

IBTs are easy to justify politically, because their implementation can be blamed on water hogs, who are punished by their introduction. But the real culprit for the water shortage is climatic variability, and the way in which the country’s water authorities plan for, and respond to, it. The implementation of IBTs is an attempt to allow water users to avoid confronting the implications of variability in water availability, and it allows many users escape the cost of their actions.

IBTs thus seem to have an irresistible appeal to regulators. Providing a ‘free allocation’ to consumers has been used frequently in the past by water providers. Sydney and Melbourne are now returning to the use of IBTs. IPART (2004:24) recommends IBT instead of two-part tariff because:

1. vulnerable customers face a higher bill under two-part tariffs;
2. two-part tariffs send a weaker conservation signal to high water users; and
3. two-part tariffs do not distinguish between discretionary and non-discretionary use.
These arguments are commonly cited to support the use of IBTs, and are misleading for the following reasons:

1. IBTs do not provide targeted welfare, and thus do not necessarily deliver low bills to vulnerable consumers. For example, large, low income families are likely to face the higher tier 2 tariff notwithstanding that on average they consume less water per person than smaller households. From a welfare and efficiency point of view it would be superior to introduce a two-part tariff with an efficient volumetric rate. Then a rebate of the fixed charge can be paid to vulnerable consumers. Such a rebate can be well targeted and flexible; for example, it could vary with a consumer’s income and number of dependents.

2. Two-part tariffs send only an inefficiently weak conservation signal when the volumetric rate is set below the efficient level. IBTs necessarily send the wrong signals to some consumers.

3. Non-discretionary use is a poorly defined concept. For example, the use of toilets would normally be classed as non-discretionary. Yet replacement of a single flush toilet with a dual flush system can significantly reduce this ‘non-discretionary’ use. However defined, the annual non-discretionary use is well below the cut-off point for current tier 1 tariffs. Under these IBTs the volumetric rate for most ‘non-discretionary’ use is the tier 1 rate.

Invoking the concept of non-discretionary demand as a means of allocating water is problematic from an economic perspective. The distinction between discretionary and non-discretionary demand presumably relates to differences in the willingness to pay for different uses of water. For example, water used for showers (non-discretionary) would have a higher marginal benefit than water used to wash the car (discretionary). However, only consumers can identify their demand for the duration and frequency of given activities. An efficient volumetric rate allows consumers to make this choice. Water authorities targeting, and then restricting, activities declared ‘non-discretionary’ removes this choice from households, and thus may lead to an inefficient allocation of water across uses.

The concept of discretionary use has enabled water authorities to avoid raising the volumetric rate to the efficient level. Rather they have introduced restrictions on certain types of use and introduced IBTs. By attempting to restrict supply this way, they have attempted to smooth the volumetric rate rather than (as is efficient) water quantity. Such an approach is politically appealing, as it gives the appearance of reducing real water bills at the same time as restricting usage. However it is inefficient, and will be ineffective if available water is not increased.

IBTs not only send the wrong price signals to consumers, but to the water providers themselves. IBTs (usually) undervalue the water sold in the first tier, and thus provide insufficient incentive to expand supply. Without identifying the market clearing volumetric rate, it is unclear whether an expansion of capacity is warranted. This issue will become increasingly important if, as foreshadowed by
NCP, it is decided that some water networks will allow private providers to supply water. An underpriced volumetric rate set by a government owned provider would appear to contravene the NCP requirement of competitive neutrality.

Adopting a volumetric rate that reflects current and expected water availability will incur some additional costs. For example, frequent changes in the volumetric rate may require water meters to be read more frequently. It would also be necessary to advertise changes in the volumetric rate. However, these costs are likely to be small when compared to the benefits. The adoption of an efficient two-part tariff would see storage levels drop more slowly at times of drought. This would see a reduced need for additional storage capacity to cover consumption over periods of low rainfall. Furthermore, there would be less need to rely on water restrictions at times of drought (see Sibly 2005). Indeed the reduced requirement to advertise water restrictions would, to some extent, offset the costs of advertising changes to the volumetric rate.

As water shortages grow in Australia’s metropolitan areas, the importance of using efficient pricing increases. The key reform required is the implementation of an efficient volumetric rate, which takes account of the erratic Australian rainfall patterns. It is the rate that clears the market for the available water and, importantly, reflects the expected availability of water. The fixed charge satisfies the residual revenue requirement. Equity concerns can be addressed by rebates of the fixed charge to disadvantaged groups. Unfortunately, many water authorities have abandoned the implementation of two-part tariffs and are proposing IBTs, which are not flexible in the face of changing availability of water. Rather they give the illusion of delivering efficiency, equity and conservation. This paper demonstrates that, in reality, they provide these things imperfectly.

References


*I would like to thank two anonymous referees and editors Ross McLeod, Franco Papandrea and Graeme Wells for very useful suggestions on an earlier draft of this paper. All errors remain my responsibility.*
This article provides the first Australian estimates of income mobility transitions among Australian children over a three year period. The results suggest that there is considerable income mobility but that, in the vast majority of cases, the shift is not dramatic. Transitions from the top to the bottom of the income spectrum, and vice versa, are very rare. Using a financial disadvantage line set at the cut-off point for the bottom quintile, the study also showed that about twelve per cent of all Australian children were in persistent financial disadvantage, experiencing financial disadvantage in all of the three years. An estimated 28 per cent of Australian children experienced financial disadvantage in at least one of the three years — about 40 per cent more than the proportion experiencing financial disadvantage in annual ‘snapshots’. Comparison with overseas studies suggests that Australia does not appear to have greater mobility, perhaps contrary to the Australian perception of the ‘fair go’. Analysis of the factors precipitating transitions into and out of financial disadvantage suggests that the emphasis by the Federal Government on requiring welfare recipients to work is an appropriate policy response, which may help to reduce the proportion of children who remain disadvantaged for extended periods of time.

The extent of income mobility and financial disadvantage among children is widely regarded as an important social issue. In Australia, the then Prime Minister, Bob Hawke, focussed attention on this issue in the 1980s, when promising that ‘by 1990 no child shall live in poverty’. In Europe, the goal of promoting social inclusion has led to regular assessment of financial disadvantage and in such countries as the UK and Ireland there are strong anti-poverty initiatives in place (Department of Work and Pensions, 2002; Atkinson et al, 2001). A related key issue is whether or not there is considerable income mobility among children over time. As Bradbury, Jenkins and Micklewright observe, a child poverty rate of 10 per cent ‘could mean that every tenth child is in poverty all the time or, at the other extreme, it could mean that all children are poor for one month in every ten’ (2001:1). Knowing the extent to which the answer for a particular country tends towards the former situation rather than the latter is of crucial importance to policy makers, given the very different likely policy responses required.

One reason these issues attract so much concern is the growing body of evidence suggesting that adverse childhood experiences have a profound impact.
upon the rest of a child’s life. Evidence shows that children who experience multiple forms of social and economic disadvantage are more likely to have poor health and well-being, with these adverse effects persisting throughout their lives and including lower subsequent rates of educational and labour market achievement and a greater likelihood of becoming teenage parents and serving prison sentences (Duncan et al, 1993, Poulton et al, 2002; UNICEF, 2000; Moore et al, 2002; Annie E Casey Foundation, 2003).

In Australia, there have been numerous studies of the extent of financial disadvantage among children at a particular point in time (Bradbury, 2003; Harding, Lloyd and Greenwell, 2001; Harding and Szukalska, 2000). But there has not been analysis of patterns of income mobility and financial disadvantage among Australian children over a longer time period. To undertake such analysis requires that we move beyond the weekly or yearly snapshot captured in the ubiquitous cross-section surveys of the population to longitudinal or panel data (where the same people and children are surveyed in each time period).

Until relatively recently, such longitudinal data have not been available for representative samples of the whole population. Full longitudinal data have previously been available for certain sub-groups of the population, including youth (through the Australian Youth Survey and its predecessor the Australian Longitudinal Survey — Bell, Rimmer and Rimmer, 1992), and recent immigrants (covered by the Longitudinal Study of Immigrants to Australia — see www.immi.gov.au/research/publications/index.htm for a list of most recent project reports). In addition, since our research was undertaken, the Australian Department of Family and Community Services has commissioned the Melbourne Institute to produce the HILDA longitudinal data (see www.melbourneinstitute.com for further information). Four waves of this survey are now available for analysis and the first papers on poverty dynamics using this data source are now emerging (Woodena and Headey, 2005; Marks, 2005). The Wooden and Headey study suggests that the majority of children who become poor do not remain poor for very long - and, while this is what we also find, this study is not directly comparable to ours because it uses a different definition of financial disadvantage (namely, half median equivalent disposable income).

This study examines income mobility and financial disadvantage among children using data from the Survey of Employment and Unemployment Patterns (SEUP), conducted by the Australian Bureau of Statistics (ABS, 2005). The SEUP was a longitudinal survey with information collected from the same individuals over three annual waves of interviews conducted in 1995, 1996 and 1997.

Internationally, the evidence suggests: that there is some income mobility among children, but that most children do not move all that far up or down the income spectrum; that the proportion of children who experience a spell of low income increases substantially as the time period extends beyond one year; and that most children do not remain persistently poor or financially disadvantaged for long periods of time (although sufficient variation between countries in the proportion of persistently poor children is revealed in these studies to underline
Income Mobility and Financial Disadvantage

the importance of looking at income dynamics) (Bradbury, Jenkins and Micklewright, 2001). For example, Gottschalk and Danziger (2001:142), who examine income mobility among children in the US over 10 years in the 1990s, echo the findings of other studies when they conclude that ‘while there is some income mobility during childhood, children who started at the bottom of the distribution tended to remain there … and children who started at the top of the distribution seldom fell very far’. Similarly, like other studies showing that the proportion of children facing a brush with low income is much higher than the proportion facing persistent poverty, Schützer (2001:166) found that 23 per cent of West German children experienced at least one year of low income during a five year period, but that only four per cent of West German children experienced low income in every year during the five years.

Do similar results emerge for Australia? Our study starts by looking at income dynamics, to establish the extent of overall income mobility among children within Australia. We then shift the focus to the lower end of the spectrum, to the persistence of financial disadvantage. Our financial disadvantage measures were chosen to be consistent as possible with results from Germany, Hungary, Ireland, Britain and the US (Bradbury, Jenkins and Micklewright, 2001).

Section 2 of this paper describes the SEUP data and key methodological features of our study, while Section 3 examines patterns of income mobility among Australian children. Section 4 analyses how persistent financial disadvantage is, and the characteristics of those in persistent financial disadvantage, while Section 5 concludes.

Data and Methodology

The SEUP sample comprises three subgroups, and our study utilises the population reference group (PRG) sample, which is a random sample of the population aged 15-59 years (with a sample size of 2311 people in the first wave, of which some 14 per cent was lost between wave 1 and wave 3). The sample used for our study was considerably smaller than this for three reasons. First, a substantial proportion of respondents to the SEUP survey did not provide estimates of their incomes and, following concerns about the way in which incomes had been imputed by the ABS for this group, we dropped them from our study (see Abello and Harding, 2004 for full details of this and other aspects of the methodology). The SEUP includes both ‘current weekly income’ (income in the week immediately preceding the survey) and annual income from the previous financial year. We focused on annual rather than current income, as the former is generally regarded as a more reliable measure of economic well-being. Second, more than half of those captured in the SEUP sample did not have children and so were deleted from our analysis. Third, the SEUP collected income data only for the respondent to the survey and their spouse (if any). As a result, in cases where the respondent was a dependent student, information about the income of their parents was not collected. If we left these dependent students in the sample then we would have had extraordinarily high rates of financial disadvantage.
Accordingly, we deleted those records where the respondent was a dependent student (approximately five per cent of records in wave 1). After these three amendments, we were left with a sample of 607 records in wave 1, 593 in wave 2, and 667 in wave 3 for all dependent children, or with 542, 512 and 574 respectively for children aged under 15 years only.

The SEUP dataset comes with a set of weights that indicates the number of persons in the whole population represented by each observation in the sample survey — that is, the weights are applied to the data to make the survey results more representative of the whole population and are also the means by which researchers ‘gross up’ the sample results to estimates for all of Australia. The ABS has calculated longitudinal weights for the SEUP data that take into account the representativeness of each respondent as well as the issue of sample attrition. We used three sets of weights, depending on the nature of the analysis. For analysis of data on a per wave basis, the ABS weights associated with each wave were used for each wave’s data. For analysis of transitions through all three waves a common set of weights across the waves was required, so the wave 3 longitudinal weights calculated by the ABS were used. As the ABS only calculated the weight of the responding person to the SEUP survey, we had to assume that the weight of the income unit (and thus of any children) was the same as the weight of the responding person. All of our results are for persons, rather than for families.

We used a family’s annual cash income before income tax as the basic indicator of their standard of living — that is, gross or total income in the preceding financial year. Most studies of financial disadvantage use disposable (after income tax) income as the measure of resources. However, the SEUP did not collect data on post-tax income and it was not possible for us to accurately impute income tax onto the survey. Income is defined as ‘regular cash receipts’ and includes wages and salaries, business and investment income, and government cash transfers such as pensions and family allowance. There were some extreme values in the SEUP data which would have unduly influenced the results (particularly means), so any negative value was set to zero and any weekly income greater than $5000 was set to $5000 (see Greenwell, Lloyd and Harding (2001:16-17), for options regarding treatment of negative incomes).

A couple with three children and a gross income of $25,000 will not be as affluent as a couple with one child and a gross income of $25,000 — because, in the former case, the income is supporting five people while, in the latter case, it is only supporting three people. Researchers typically use equivalence scales to adjust income, so as to take account of such differences in family size and composition. There is no agreement, either in Australia or internationally, about which is the ‘right’ equivalence scale to use (ABS, 2004:52). We have used the original OECD equivalence scale which assumes, for example, that couples with two children require 59 per cent more income than couples without children to reach the same standard of living. The OECD equivalence scale carries a weight of one for the first adult in the unit, 0.7 for any other adult and 0.5 for each child.
The ‘income unit’ is the group among whom income is assumed to be shared equally. Our study employs the ABS definition of the income unit, which means that an income unit is defined as either a couple with dependent children, a couple without dependent children, a sole parent with dependent children, or a single person. A dependent child is defined as a child aged less than 15 years or a 15–24 year old in full-time study and still living in the parental home. Following convention, throughout the paper we have referred to the income unit as ‘the family’.

Finally, we use two financial disadvantage thresholds. These were set at the boundaries for the lowest quintile and lowest decile of equivalent family gross income of all Australian children in each wave. That is, there are 20 per cent of the weighted sample of children in each quintile in each wave, and 10 per cent in each decile.

Income Mobility

This section begins with a description of the extent and pattern of income dynamics among children. Every child in the sample was classified into a family income decile group in each of the three years (that is, 10 per cent of all children were in each decile), and the resulting group classification in one year was cross-tabulated with the group classification at another year to reveal the pattern of change in income groups over the period. The one-year transition or movement is based on the pattern of movement for two one-yearly intervals, that is, movements starting in year 1 and ending in year 2, and movements starting in year 2 and ending in year 3. The two-year movements are those between years 1 and 3.

Over one year, 36 per cent of all children remained in the same decile group they started out in, whether it was the first, second or tenth decile (Table 1). Over two years the percentage declined to 34 per cent. Thus, about two-thirds of all children switched income deciles over the course of just one year, suggesting considerable mobility. If only children under 15 years of age are considered, the percentages are 37 and 29 per cent respectively.

Table 1 also shows the proportion of children in the sample who were in the same or neighbouring (one higher or one lower) income group. It shows that most movements in one or two years are over a short range. For example, although only 36 per cent of the sample stayed in the same decile group over a one-year interval, more than double this proportion (76 per cent) remained in the same decile or moved to a neighbouring decile group. Similarly, while 61 per cent of all children remained in the same quintile group after one year, 92 per cent were in the same or an adjacent quintile group.
Table 1: Overall Mobility Between Income Groups: One- and Two-year Intervals, 1994-95 to 1996-97

<table>
<thead>
<tr>
<th></th>
<th>All children</th>
<th>Children under 15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 yr(^a)</td>
<td>2 yr(^b)</td>
</tr>
<tr>
<td>Proportion remaining in the same:</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Decile group</td>
<td>36</td>
<td>34</td>
</tr>
<tr>
<td>Quintile group</td>
<td>61</td>
<td>58</td>
</tr>
<tr>
<td>Proportion remaining in the same or adjacent:</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Decile group</td>
<td>76</td>
<td>70</td>
</tr>
<tr>
<td>Quintile group</td>
<td>92</td>
<td>89</td>
</tr>
</tbody>
</table>

Notes: \(a\) = Average of transitions between waves 1 and 2 and waves 2 and 3.  
\(b\) = Transition between waves 1 and 3.

Low Income and High Income Persistence

The data showing the degree of mobility between income groups can also be used to show persistence in low income and high income groups over the short term (Table 2).

The degree of short-term low-income persistence depends on the low-income threshold chosen. For example, if the cut-off is the lowest decile, then about one-quarter of children who were in the bottom decile in either years 1 or 2 were still there one year later (and thus about three-quarters moved out of extreme financial disadvantage). If the threshold is raised to the lowest quintile, the profile is very different, with almost three-quarters remaining in the same bottom quintile group and only about one-quarter escaping to higher income quintiles.

The degree of short-term high-income persistence also depends on the definition of the income threshold. If we focus on the top income decile, then less than seven out of every 10 children in the top decile in one year have moved out of that decile by the following year. This is a much lower degree of mobility than is apparent for children in the lowest income decile. However, once the income measure is extended to the top quintile, there is relatively little difference in mobility between the top quintile and the bottom quintile. As Table 2 shows, in each case 72 per cent of children who were in either the top or bottom quintile in one year remained in that quintile in the following year.
Table 2: Low Income and High Income Persistence: One-year Interval, 1994-95 to 1996-97

<table>
<thead>
<tr>
<th>Proportion in the same low income group one year later</th>
<th>All children</th>
<th>Children under 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest decile group</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>Lowest quintile group</td>
<td>72</td>
<td>72</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proportion in the same high income group one year later</th>
<th>All children</th>
<th>Children under 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Richest decile group</td>
<td>67</td>
<td>64</td>
</tr>
<tr>
<td>Richest quintile group</td>
<td>72</td>
<td>75</td>
</tr>
</tbody>
</table>

Notes: a = Average of transitions between waves 1 and 2 and waves 2 and 3.

The figures in the third row of the table can be used to show that if 20% of children were in the lowest quintile group in one year, then 72% of these were still there in the following year (that is, 14.4% of the total sample — or \([20 + 20 \times 0.72]\)). Conversely, 28% were not still there the following year, so that 25.6% of the sample experienced one bout of low income in any two year period (that is, \([20 + 20 \times 0.28]\)).

This again suggests that there is much greater apparent mobility when looking at income deciles — but that most mobility only involves a move from one decile to an adjacent decile. Figure 1 shows the proportion of children remaining in the same decile over a one-year interval averaged over the period 1994-95 to 1996-97. For the majority of children — those in income deciles one to seven — between about 20 and 40 per cent remain in the same decile from one year to the next. For those in the top two income deciles, the picture is very different, with about half to three-quarters remaining in the same decile from one year to the following year.

Overall, 36 per cent of children remained in the same annual income decile over the period. Of the 64 per cent who did change deciles, most moved to an adjacent decile as shown by the top line in Figure 1. The proportion of children remaining in the same income decile or moving to an adjacent decile was around 80 per cent for the lower deciles, 75 per cent for the middle deciles, and 80 per cent for the top deciles. The foregoing confirms that there is more short-range stability at the top of the income distribution than at the bottom — but that right across the income spectrum, about four out of every five children remain in the same or an adjacent income decile from one year to the next.
Long-range Mobility

The above suggests that about one-fifth of children live in families whose circumstances change sufficiently to shift them two or more deciles across the income distribution. Are there many children shifting from the very bottom to the very top of the income spectrum — or from the very top to the very bottom? Table 3 summarises long-range mobility, defined here as the proportion of the richest group that move to the poorest group in the next period, and vice versa. The numbers indicate that very little long-range movement occurs, with only one in every hundred children in the highest income decile in one year falling to the lowest income decile in the following year. In our sample, no children were in families that catapulted from the lowest income decile in one year to the highest income decile in the following year. Even if the period of observation is extended from one year to two, the proportions are still quite small. This confirms that most of the children moving out of the poorest income groups or the richest groups do not move very far. While it must be emphasised that we are only looking here at a three-year period, these results nonetheless echo those for other countries, suggesting that for the majority of children their place in the income ladder is relatively fixed.
Table 3: Long-range Mobility: One- and Two-year Intervals, 1994-95 to 1996-97

<table>
<thead>
<tr>
<th></th>
<th>All children</th>
<th>Children under 15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 yr&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2 yr&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Proportion of richest income group moving to poorest income group</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Top decile to bottom</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Top quintile to bottom</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Proportion of poorest income group moving to richest income group</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Bottom decile to top</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bottom quintile to top</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Notes:  
<sup>a</sup> = Average of transitions between waves 1 and 2 and waves 2 and 3.  
<sup>b</sup> = Transition between waves 1 and 3.

Finally, Table 4 shows the full mobility matrix by income decile for children from one year to the next, key features of which were summarised in Tables 1 to 3. It sheds further light on the extent of mobility through the income spectrum. For example, it shows that 73 per cent of all children in the bottom income decile in one year remained in either the bottom or second income decile one year later — and that none moved up to the top decile. But Table 4 also shows us that about one-fifth of all those children in the bottom income decile in one year had moved up to deciles three and four a year later, suggesting a substantial improvement in their economic well-being. Conversely, however, of all those children who were in decile 3 in one year and thus above our financial disadvantage thresholds, about one in every eight lived in families that had slipped down to the bottom income decile one year later.
Table 4: Movement Between Annual Income Deciles: One-year Transitions, All Children, 1994-95 to 1996-97

<table>
<thead>
<tr>
<th>Decile in first period</th>
<th>% of all children</th>
<th>Decile in subsequent period</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: a = Average of transitions between waves 1 and 2 and waves 2 and 3.

Financial Disadvantage

We have used two measures of financial disadvantage based on the thresholds set previously: those children whose equivalent gross family income places them in the lowest quintile of children are defined to be in financial disadvantage, while those whose equivalent gross family income places them in the lowest decile are defined to be in severe financial disadvantage.

Persistent Financial Disadvantage

This section describes the extent of persistent financial disadvantage among children. It should be noted, however, that our definition of persistent financial disadvantage is constrained by the availability of data for only a three year period as we have no information on equivalent income before and after this period; other studies, such as Duncan, Coe and Hill (1984) define the persistently poor over a much longer time period (eight or more years out of ten). Further, this analysis requires that the respondents have provided data on their income for all three years, so sample sizes are even smaller.

We start here by identifying children who were in financial disadvantage in 1994-95 and then look at what happened to them over the following two years. Using the lowest income decile as our severe financial disadvantage cut-off, only
one in every 100 children were in persistent severe financial disadvantage over the three years — that is, they were in families with incomes below this financial disadvantage threshold in all three years (Table 5). About five in every 100 children were in severe financial disadvantage in 1994-95 and still in severe financial disadvantage in 1995-96.

Table 5: Proportion of Children in Financial Disadvantage by Number of Years, 1994-95 to 1996-97

<table>
<thead>
<tr>
<th>Financial disadvantage threshold</th>
<th>All dependent children</th>
<th>Children aged under 15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lowest decile</td>
<td>Lowest quintile</td>
</tr>
<tr>
<td>Proportion in financial advantage in:</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>All 3 years</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>2 years only</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>1 year only</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Proportion in financial advantage in:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All 3 years</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>At least 2 years</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>At least 1 year</td>
<td>19</td>
<td>28</td>
</tr>
</tbody>
</table>

If the threshold is raised to the lowest quintile income cut-off, the proportion persistently in financial disadvantage in all three years rises to twelve in every 100 children. About 13 in every 100 children experienced two years of financial disadvantage (Table 5). With respect to children under the age of 15 years, we find slightly higher proportions in persistent financial disadvantage (and correspondingly, slightly smaller proportions never in financial disadvantage).

While only a small proportion of children were in financial disadvantage in every year, many experienced financial disadvantage at some point during the three years. Starting with the figures based on the lowest decile cut-off, 1 per cent were in severe financial disadvantage in all three years, 5 per cent experienced severe financial disadvantage in any two years, and 13 per cent had a brush with severe financial disadvantage in only one year. These proportions imply that, during the three years, 1 per cent of the sample had three spells of severe financial disadvantage, 6 per cent had at least two spells of severe financial disadvantage, and 19 per cent had at least one spell of severe financial disadvantage (lower panel of table 5). In other words, the lives of one-fifth of all children in the sample had been touched by severe financial disadvantage over the three years.
If the financial disadvantage threshold is raised to the lowest quintile of current income, the proportion having at least one spell of financial disadvantage rises to nearly a third of the sample (28 per cent) — 40 per cent more than the proportion for a single year. Conversely, this means that over two-thirds of all children in our sample never experienced financial disadvantage within the three years, using the lowest quintile measure.

The relatively high proportion of children in the sample experiencing some financial disadvantage over the period 1995–97 is another manifestation of the earlier finding about mobility between income groups being common at all points along the income range.

The extent of persistent financial disadvantage can also be gauged from another perspective — the proportion of children in financial disadvantage at a point in time compared with the proportions in financial disadvantage throughout the entire period. Of the group of children defined to be in financial disadvantage based on the lowest decile cut-off in 1994-95, table 6 shows that 23 per cent were still in severe financial disadvantage one year later and 12 per cent were in severe financial disadvantage in all three years. When the cut-off is the lowest quintile of current income, the percentages are much higher. Of the children in financial disadvantage in 1994-95, 71 per cent were still in financial disadvantage a year later and two-thirds were in financial disadvantage in all three years.

Table 6: Children in Financial Disadvantage in 1994-95 and Still in Financial Disadvantage in 1995-96 and in or 1996-97

<table>
<thead>
<tr>
<th>Financial disadvantage threshold</th>
<th>All dependent children</th>
<th>Children aged under 15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lowest decile</td>
<td>Lowest quintile</td>
</tr>
<tr>
<td>Proportion in financial disadvantage in year 1</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Proportion still in financial disadvantage in year 2</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>71</td>
</tr>
<tr>
<td>Proportion still in financial disadvantage in years 2 and 3</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>67</td>
</tr>
</tbody>
</table>

Despite the sensitivity of the results to the financial disadvantage threshold used, the foregoing numbers show that a large proportion of those in financial disadvantage in the first year remained in financial disadvantage through all three years of the SEUP survey. These results also indicate that there is a greater likelihood of staying in financial disadvantage among those who have been in financial disadvantage at some point in time than among the population as a
whole. Jarvis and Jenkins (1996:16) indicate that the reason for this is straightforward — ‘those in the low income stock have disproportionately long low income spell durations compared to the population as a whole; those with relatively high exit rates and hence, shorter durations, leave first, leaving behind the longer duration people’.

Characteristics of the Persistently Financially Disadvantaged

Do children in persistently financially disadvantaged families have some particular set of characteristics, or are they just a random subset of those who are financially disadvantaged at a particular point in time? To address this issue, we briefly look at selected characteristics of children financially disadvantaged in all three waves, and compare the distribution by family type and economic activity for this group, with the corresponding distribution of children financially disadvantaged in wave 1, and all children in wave 1. In this section we focus the analysis on financial disadvantage defined using the lowest quintile of current income. The primary reason for choosing current income over annual income is that the variables important to our analyses (on family type and economic activity) are contemporaneously associated with current income.

Of the total sample of children in persistently financially disadvantaged families, 42 per cent were in sole parent families. The corresponding proportion for children who were financially disadvantaged in wave 1 was close to half that (22 per cent), while of the total sample of children in wave 1 only 12 per cent were in sole parent families.

Another major difference between the persistently financially disadvantaged and those financially disadvantaged in wave 1 is that the former are over-represented in families where one or both parents are not working. Forty one per cent of persistently financially disadvantaged children were in couple families where one or both parents were unemployed or not in the labour force, compared with 32 per cent for those financially disadvantaged in wave 1, and 8 per cent for the whole sample of children. In addition, 39 per cent of persistently financially disadvantaged children were in sole parent families where the parent was not in the labour force in wave 1, compared with 18 per cent for all those financially disadvantaged in wave 1, and 6 per cent for the whole sample.

Conclusion

The regular national income surveys conducted by the Australian Bureau of Statistics mean that we are accustomed to receiving up-to-date information about the extent of financial disadvantage and income inequality at a particular point in time. But for many policy makers and researchers, the degree of movement into and out of financial disadvantage and the accompanying mobility in income distribution are equally important social indicators. As Bradbury, Jenkins and Micklewright (2001:130) note, even when conventional cross-sectional child poverty rates are not increasing, there may still be a decline in mobility —
resulting in undetected increases in the proportion of children experiencing long spells of poverty. ‘In other words, changes in dynamic patterns can have as important an impact upon child well-being as changes in the cross-sectional poverty rate’.

Looking at income mobility first, while there was extensive movement between family income groups from one year to the next, the moves were generally not large. Some 64 per cent of children moved from one decile group to another between one year and the next. This suggests a reasonably comparable degree of income mobility in Australia to that found by Bradbury, Jenkins and Micklewright for Britain (59 per cent), the USA and West Germany (57 per cent), and Ireland (60 per cent) (2001:103). Our analysis showed that children who were rich in one year were unlikely to slip into financial disadvantage the next year, and children who were poor in one year were unlikely to be rich the next year. This lack of long-range movement across the income spectrum is also consistent with the findings for other countries (for example, Jervis and Jenkins, 1996:40; Gottschalk and Danziger, 2001).

The proportion of Australian children experiencing at least one brush with financial disadvantage during the three years was higher than the proportion in financial disadvantage in any particular year. Over three years, the proportion of children facing at least one bout of financial disadvantage was 28 per cent — about 40 per cent higher than the proportion estimated to be in financial disadvantage during any single year (using the bottom quintile cut-off as the definition of financial disadvantage). The proportion of children facing at least one year of financial disadvantage out of any two years in our sample can be calculated from Table 2 and stands at 25.6 per cent. This again appears to be highly comparable with those in other industrialised countries, with Bradbury, Jenkins and Micklewright reporting that the proportion of children who were ever in the bottom quintile over a two year period was 26.4 per cent for Germany and the USA, 25.3 per cent for Ireland and 27.1 per cent for Britain (2001:107).

While the SEUP data reveals relatively low estimates of persistent financial disadvantage across the three years, a look at the proportion staying in financial disadvantage after each wave shows that the majority of children in financial disadvantage in the first wave remained there in successive waves. Taken in conjunction with the information on high exit rates from poverty, the SEUP data shows that in Australia, while a significant proportion of children slip into and out of financial disadvantage, a small pool of children tend to remain in financial disadvantage, at least for the three years that we have data.

We can again use the figures in table 2 to calculate that 14.4 per cent of children stayed in the bottom quintile for any two consecutive years: this compares with 13.9 per cent for Britain and 14.2 per cent for the USA (Bradbury, Jenkins and Micklewright, 2001:107), providing support for the view that low income persistence in Australia does not appear to be lower than in other industrialised countries and perhaps contrary to the traditional view of Australia as the land of the ‘fair go’. Unfortunately, we do not have internationally comparable data for a three year span.
An important issue for policy makers is determining the causes of persistent financial disadvantage, and transitions into and out of financial disadvantage. As one might expect, children in persistent financial disadvantage were over-represented in couple families where one or both parents were not working and in sole parent families (particularly where the parent was not in the labour force). There was also a strong association between changes in a child’s family characteristics and changes in income or financial disadvantage status. Changes in the number of earners in the family had a clear-cut association with transitions out of or into financial disadvantage, while decreases in the number of children (via children reaching adulthood and leaving the parental home) were more important for transitions out of financial disadvantage (Abello and Harding, 2004:29). These findings are consistent with studies of income dynamics overseas, which suggest that households with two earners are less likely to be financially disadvantaged and more likely to subsequently rise again above the low income threshold if they do slip below it (OECD, 1998:184), while households headed by a sole parent are more likely to remain persistently poor (Jervis and Jenkins, 1996; Huff-Stevens, 1995; Oxley, Dang and Antolin, 2000; OECD, 1998; Desrosiers et al., 2002).

Our results, like those of the other income dynamics studies indicated above, thus suggest that government policies that focus on reducing unemployment and encouraging labour force participation — including employment and training programs and longer term measures to increase worker skills — should have a positive effect upon child financial disadvantage. The international evidence suggests that it is important to identify those who are persistently poor, with this group requiring greater policy attention than those who are poor for only a short period — and also that it is appropriate to complement income measures with other measures of deprivation to help policy makers identify and target this group (Nolan, Maitre and Watson, 2001).

In recent years, the policy landscape has changed greatly in Australia: the creation of new longitudinal datasets has challenged the perception of mobility among the sole parent population, with studies suggesting that the average period of receipt of income support may be as long as 12 years (Gregory and Klug, 2003).

In the 2005 Budget and subsequently, the Australian government announced a range of welfare to work measures, including requiring sole parents whose youngest child is aged over 7 years to seek at least 15 hours of paid work a week (Andrews, 2005). While this measure should help to ensure that Australian children do not remain financially disadvantaged over long periods of time, such sole parents are also to be placed on a form of income support which has a lower payment rate and a much harsher income test than the current pension (and thus higher effective marginal tax rates on earned income — Harding et al., 2005). As a result, the final impact upon child financial disadvantage will depend upon how these opposing factors play out in practice.

Given the importance of multiple income earners in the prevention of child financial disadvantage, the high effective tax rates faced by secondary income earners in Australia are also an important related issue. It is often the case that the
very people who stand to benefit from greater participation in the workforce are the ones that face the greatest financial disincentives to do so (Toohey and Beer, 2004). Child care subsidies can also play an important role in helping to boost the labour force participation of secondary earners and sole parents, so the persistence of child care affordability and availability problems is another area of potential policy attention (McNamara, Cassells and Lloyd, 2005).

References


Australian Bureau of Statistics (2004), Household Income and Income Distribution, 2002-03, Catalogue No 6523.0, Canberra


The authors wish to thank Bruce Bradbury and Anthony King for their extremely helpful comments on earlier drafts of this paper. Many thanks also to our anonymous referees for their comments.
Immigration Detention Reforms: A Small Gain in Human Rights

Savitri Taylor

Australia’s mandatory immigration detention regime came into force in 1994. From the outset, the regime has been vigorously opposed by many civil society organisations and some of its aspects have been criticised by Commonwealth Government agencies such as the Human Rights and Equal Opportunity Commission (HREOC) and the Commonwealth Ombudsman’s Office. Increasing media scrutiny since 2000 has also had an effect on public opinion (Taylor, 2001). Both the Labor Government, which introduced the regime, and the Liberal/National Coalition Government, which continued it, have sought to respond to political pressure by progressive mitigation of the harshness of the regime without significant change to the fundamental framework. The Prime Minister’s announcement of a number of changes to the mandatory detention regime on 17 June 2005 was the most recent bid to ease political pressure in this way (Howard, 2005).

Australia’s Immigration Detention Regime

The aims of the immigration detention regime are: to protect the community from individuals who pose a health, criminal or national security threat; to ensure that non-citizens who are in Australia in breach of immigration laws are available for removal from the country; and to deter prospective unauthorised arrivals (Liberal Party of Australia and Nationals, 2004). However, the design of Australia’s immigration detention regime makes it clear that the last stated aim is the primary one. The Migration Act 1958 (Cth) mandates the detention of ‘unlawful non-citizens’ (those entering Australia without a valid visa) until they leave or are granted a visa. Persons entering Australia on any type of valid visa and subsequently apply for a substantive visa are usually eligible for a so called ‘bridging visa’, which gives a person the status of ‘lawful non-citizen’ and allows that person to be at liberty with protection from removal from Australia pending determination of the visa application. In contrast, unauthorised arrivals are ineligible for a bridging visa except in very limited circumstances. If they apply for a substantive visa, they are usually kept in immigration detention, pending determination of the application (and pending removal if unsuccessful).

The deterrence aim is freely acknowledged in the rhetoric of Australia’s federal politicians (Vanstone, 2005a:35; Dodson, 2005). However, it is not officially acknowledged by Australia in any forum in which the international legality of its immigration detention regime is under consideration (see, for

Savitri Taylor is a Senior Lecturer in the School of Law, La Trobe University
example, HREOC, 2004:210). This is because the first two aims of immigration detention are legitimate under international law, but the last is not (Taylor, 1998). Detaining unauthorised arrivals regardless of the merits of their particular situation for the purpose of deterring others from following in their footsteps is considered illegitimate for one very simple reason. It involves treating human beings as mere means to an end and is, therefore, a repudiation of the fundamental premise of international human rights law, which is the equal worth of every human being.

**Human Rights Implications of Mandatory Detention**

Australia is a party to the International Covenant of Civil and Political Rights (ICCPR). Article 9(1) of the ICCPR provides that no one shall be subject to ‘arbitrary’ detention. Detention is ‘arbitrary’ if it is not permitted by domestic law or is not a necessary and proportionate means of achieving a legitimate end. Since deterrence is not considered to be a legitimate aim of immigration detention by the international community, the use of detention for that purpose would be 'arbitrary' and in breach of article 9(1). While community protection and availability for removal are considered to be legitimate purposes, detention for those purposes must still be a necessary and proportionate means of achieving the purposes. The only way to ensure that the requirements of necessity and proportionality are met is for detention decisions to be made on a case-by-case basis. The ‘necessary’ criterion is satisfied if it can be shown that there is no less restrictive means of achieving the end in question in the particular case. Satisfaction of the ‘proportionate’ criterion requires demonstration that the importance to society of the end to be achieved by detention outweighs the importance to the individual of physical liberty and the negative impact on the individual of deprivation of liberty. It is, of course, almost impossible to overstate the importance to the individual of physical liberty. Since the negative impact of detention on the individual tends to increase as the duration of detention increases, duration of detention is a relevant factor in assessing proportionality. The longer the period of detention, the more difficult it becomes to describe the detention as a proportionate means of achieving any legitimate end. Further, there are some especially vulnerable individuals for whom detention can be expected to be particularly harsh and this, too, is a factor in assessing proportionality.

Pursuant to the Optional Protocol to the ICCPR Australia has recognised the competence of the UN Human Rights Committee ‘to receive and consider communications from individuals’ who claim to be victims of ICCPR violations by Australia. In three recent cases involving complaints by immigration detainees (C v Australia; Baban v Australia; Bakthyari v Australia) the Human Rights Committee found that even if immigration detention was initially justified, the Australian Government had failed to justify continued detention ‘in light of the passage of time and intervening circumstances.’ In particular, the Human Rights Committee found that the Australian Government had not demonstrated ‘there were not less invasive means’ for ensuring the complainants were available for removal, ‘for example, the imposition of reporting obligations, sureties or other
Immigration Detention Reforms

conditions’. In all three cases, the Committee was accordingly of the view that detention had been arbitrary and constituted a violation of article 9(1) of the ICCPR. The Human Rights Committee also considered whether Australia was in violation of article 9(4) of the ICCPR which provides that:

Anyone who is deprived of his liberty by arrest or detention shall be entitled to take proceedings before a court, in order that court may decide without delay on the lawfulness of his detention and order his release if the detention is not lawful.

The Australian Government argued that adequate compliance with article 9(4) was achieved by the fact that an Australian court could order release if immigration detention was found to be unlawful under Australian law. However, in all three cases a majority of the Human Rights Committee took the view that the term ‘lawfulness’ in article 9(4) refers to ‘lawfulness’ under international law as well as domestic law. They found Australia to be in violation of article 9(4) since no Australian court had power to consider whether there was substantive justification for continued detention under international law, let alone order release in the absence of such justification.

In summary, the three main aspects of the mandatory detention regime that are of human rights concern are: the plight of especially vulnerable individuals; the problem of prolonged or indefinite detention; and the absence of independent review.

Situation of Especially Vulnerable Individuals

This section explores whether the recent immigration detention reforms have resulted in human rights gains for especially vulnerable individuals.

Alternative Places of Detention and Residence Determinations

Section 5(1) of the Migration Act provides that a person is in ‘immigration detention’ if held by or on behalf of a person who is authorised by the Minister in an immigration detention centre (IDC) or other place listed in the section or ‘another place approved by the Minister in writing’. Therefore, even before the June 2005 reforms, one option for dealing with vulnerable individuals ineligible for a bridging visa, but whose needs could not be met in an IDC, was to arrange a more suitable place for their detention. In practice, alternative detention was also subject to the detainee meeting health and character requirements and being assessed as not presenting a high risk of absconding or other management problems (Department of Immigration, 2002: paras 5.2.1, 5.2.6 and 6.1.2).

1 Formerly (before January 2006) Department of Immigration, Multicultural and Indigenous Affairs (DIMIA), now Department of Immigration and Multicultural Affairs (DIMA).
Moreover, the use of alternative places of detention was relatively rare prior to August 2001.

A significant step towards more extensive use of alternative places of detention was achieved with the establishment of a Residential Housing Project (RHP) near the Woomera IDC in August 2001 (HREOC, 2004:143) at a time of great political pressure (see Taylor, 2001). The Woomera RHP, and another RHP established in Port Hedland, were later decommissioned due to reduced numbers in detention, but can be re-commissioned if circumstances change (DIMA, 2005a). At present there is one operational RHP with a maximum capacity of 40 located at Port Augusta, with one adult male in residence (DIMA, 2006). An RHP of similar capacity will open in Sydney in mid-2006 and an RHP with a capacity of 12 residents in Perth in late 2006 (Vanstone, 2006). Examples of other alternative places of detention include foster care homes (primarily used for unaccompanied minors) and accommodation made available pursuant to community care arrangements between DIMA and non-government organisations (NGOs)(DIMA, 2002:paras 6.4.1-6.4.5 and 6.6.1-6.6.4).

Prior to the June 2005 reforms, officers of the Department of Immigration were instructed that unlawful non-citizen women and children, unable to be released from detention relatively quickly, ‘should as a matter of priority be given the option’ of placement in an RHP (DIMA, 2002:para 5.2.1). They could also consider giving ‘other detainees with special needs better suited to management in a housing environment, the option of residing in a RHP’ (para 5.2.6). Of course, placement in an RHP was subject to a place actually being available.

According to the same departmental instructions (para 5.1.2):

The purpose of a RHP is to permit detainees to live in a more domestic environment that is less structured than an [IDC], and to permit them a greater degree of autonomy over their own lives, while they remain formally in detention and available for processing and removal from Australia if necessary.

However, because adult males were not considered for RHP placement, families were placed in the difficult position of choosing between better conditions of detention for eligible family members and family unity (HREOC, 2004:145). Also within RHPs detainees had only restricted freedom of movement and were subjected to constant surveillance and a high degree of supervision. As one child who was detained at the Woomera RHP observed (quoted in HREOC, 2004:160):

The [detention] centre has its own problems and the housing project has also its own problems. Like I think both are equal. Just here is like ... the shape is different.

Prior to the June 2005 reforms, DIMA felt that more use of the community care alternative detention option could have been achieved with greater willingness by NGOs to enter into the necessary arrangements (Vanstone, 2005b). NGOs had, in fact, two good reasons to be wary of entering such arrangements.
One was inadequate resources. According to Melbourne’s Asylum Seeker Resource Centre (Karapanagiotidis, 2005):

The Community Detention Model places the burden for the care of such families onto what are usually non-government funded asylum seeker welfare agencies. [DIMA] rarely will fund the medical care, housing or material aid needed by asylum seekers despite them legally still being in detention and thus in their care. This has placed great pressure on the Community Detention Model places the burden for the care of such families onto what are usually non-government funded asylum seeker welfare agencies. [DIMA] rarely will fund the medical care, housing or material aid needed by asylum seekers despite them legally still being in detention and thus in their care. This has placed great pressure on agencies to meet the needs of asylum seekers who often have medical needs that require specialist care and management, income to support their family and stable housing to ensure their safety and well-being.

The other main reason for NGO wariness was the Migration Act’s requirement for nomination of not only a place of detention but also a person responsible for the detention. Serving as jailers was not a comfortable role for these organisations because of the potential conflict between serving the welfare needs of the client and discharging responsibilities under the Migration Act. To overcome this obstacle, a new subdivision entitled ‘Residence Determinations’ was inserted into the Migration Act as part of the June 2005 reforms (Vanstone, 2005b). The new provisions give the Minister for Immigration a personal and non-compellable power exercisable ‘in the public interest’ to make a determination that a specified person is to reside in a specified place and comply with certain conditions instead of being detained in the manner usually required by the Migration Act. According to the Government, the purpose of the power is to enable the detention of families with children, unaccompanied minors and ‘adults presenting with exceptional circumstances, to take place in the community where conditions are set to meet their individual circumstances’ (DIMA, 2005b; Commonwealth Ombudsman, 2005). Persons subject to residence determinations are ‘free to move about in the community without being accompanied or restrained by an officer under the act’ (McGuaran, 2005). Nevertheless, most provisions of Migration Act continue to apply to such a person ‘as if the person were being kept in immigration detention’ under section 189. Those subject to residence determinations are advised by DIMA, that in order to prove their status, if for example, they are picked up by police, they should carry with them at all times a letter provided to them by DIMA which sets out the residence determination arrangement as well as Departmental contact details (DIMA, 2005c:111).

The residence determination model of community detention is a considerable improvement on the alternative place of detention model. Not only are NGOs relieved of the burden of acting as jailers, they also appear to have been relieved of the burden of supporting detainees. DIMA, through the Australian Red Cross (ARC), is funding the provision of income support, accommodation and health care for persons subject to residence determinations (Clutterbuck, 2005:69; Greco, 2005:40). Moreover, ARC in partnership with leading welfare agencies such as the Brotherhood of St Laurence (BSL), St Vincent de Paul, the Salvation Army,
Anglicare and Uniting Care, is developing and implementing a comprehensive and coordinated approach to working with persons subject to residence determinations, including ongoing needs assessment and case management as well as service delivery (BSL, 2005:2; Greco, 2005:38).

While the residence determination model of community detention will certainly result in human rights gains for persons detained pursuant to such determinations, there is no actual compulsion on the Minister to use this form of community detention for a particular individual as opposed to declaring an alternative places of detention for that individual. However, the Government’s intention is that, in future, families (not only mothers and children but also fathers) will be detained in RHPs until primary assessment of visa applications is complete and will then be detained in the community under a residence determination unless community detention conditions are breached or removal is imminent (McGuaran, 2005). At 3 February 2006, there were 101 people, including 54 children, subject to residence determinations (DIMA, 2006).

Release from Detention on Bridging E Visa

As previously mentioned, a bridging visa is the device used in the Migration Act to avoid detention for authorised arrivals not holding a substantive visa. As a general rule, unauthorised arrivals are not eligible for a bridging visa. However, in an exception to the general rule, prescribed classes of unauthorised arrivals are theoretically eligible for the grant of a Bridging E Visa. The prescribed classes include protection visa applicants aged less than 18 or more than 75 or who cannot properly be cared for in a detention environment because of ‘a special need (based on health or previous experience of torture or trauma)’ (Migration Regulations). These categories of person are eligible for a Bridging E Visa, subject to there being adequate arrangements for their care or support in the community and subject to health, character and security requirements.

Although the relevant provisions have been in place since 1 September 1994, only a few unauthorised arrivals with the specified vulnerabilities have been granted bridging visas (HREOC, 2004:182-97). However, the fact that on 22 June 2005 there were 8 children and 24 special need individuals in the community who had been released on a bridging visa under the relevant provisions (DIMA, 2005d) may now indicate a greater readiness by DIMA to use these provisions.

Use of the provisions is not unproblematic. There are difficulties in meeting the requirement that there be adequate arrangements for their care or support in the community. For adults, the adequate care and support requirement arises from the fact that most persons released on a Bridging E Visa are denied both work rights and social assistance entitlements (Burn and Reich, 2005:162). Unfortunately, the NGOs that enter into care arrangements with DIMA do not have the resources to meet the basic needs of those released, especially over the longer term. Although DIMA recognises this and is willing to consider NGO requests to assist with living allowances, medical and health costs, administration costs and the like on a case-by-case basis (DIMA, 2005e), the fact remains that it does not underwrite
these costs fully or as a matter of course. For example, DIMA was contributing to the costs of maintaining only 7 of the 24 individuals in the community on 22 June 2005 who had been released on a bridging visa on special need grounds (DIMA, 2005f). Moreover, DIMA considers its duty of care to be at an end once release on a bridging visa has occurred. Thus bridging visa holders have no recourse if care arrangements prove inadequate or unsustainable as they often do (Domicelj, 2005:30). A 2003 study published by the Asylum Seeker Project Hotham Mission (2003:4) found that asylum seekers on Bridging E Visas ‘live in abject poverty’, experience ‘high levels’ of mental health problems, ‘a general reduction in overall health’, and ‘high levels of family breakdown’.

On a more positive note, DIMA is currently conducting a review for the purpose of making recommendations to the Minister on how greater flexibility in the availability of, and conditions attached to, bridging visas can be introduced so that DIMA officers can respond appropriately to the needs of particular individuals without compromising achievement of the Government’s immigration control objectives (McMahon, 2005).

Exercise of Section 195A Migration Act

Another serious failing of the immigration detention regime (prior to the June 2005) was the inability to grant a bridging visa to an unauthorised arrival falling outside the prescribed classes, whether or not such a grant was an appropriate response to that individual’s circumstances. Since 29 June 2005 the Minister for Immigration has had a personal and non-compellable power under section 195A of the Migration Act to grant a visa to a person who is in detention under section 189, if ‘the Minister thinks that it is in the public interest to do so’. In exercising this power, the Minister may grant any visa deemed appropriate, even if the person does not satisfy the specified criteria for granting that visa. This means, among other things, that the Minister can grant a bridging visa to an unauthorised arrival (outside the prescribed classes) for whom detention would have a harsher than normal impact. Moreover, since the Minister may grant any visa, the granted bridging visa could be of a class that carries work and social assistance entitlements.

Length of Detention

Only 25 per cent of all unlawful non-citizens detected by DIMA are taken into immigration detention — an even lower (7 per cent) proportion of all unlawful non-citizen women and children is taken into detention (Vanstone, 2005c). This is because most unlawful non-citizens enter Australia on a valid visa and are usually, therefore, eligible for a bridging visa after becoming unlawful. The problem is that a significant proportion of those who are taken into immigration detention remain in detention for long periods of time. At 23 September 2005, there were 748 immigration detainees of whom 116 had been in detention for between one and two years and 92 had been in detention for two years or more (DIMA, 2005g).
As a result of the June 2005 reforms, the number of long-term detainees is the lowest it has been in at least the last four years (DIMA, 2005g).

Processing Time Limits

To minimise the time unauthorised arrivals spend in detention pending the processing of protection visa applications, DIMA and the Refugee Review Tribunal (RRT) have long given priority to applications by immigration detainees and respectively have had processing time targets for these cases of 42 days and 70 days (DIMA, 2004:70; RRT, 2004:23-24). Since December 2005 the Migration Act also sets out legislated processing ‘time limits’ of 90 days each by DIMA and by the RRT. Non-compliance with the ‘time limits’ will trigger reporting requirements by the Department to the Minister and by the Minister to Parliament, but, as the Government emphasises, it ‘will not give rise to a right to the issuing of a visa or release from detention’ (McGuaran, 2005). It is also important to note that this reform doesn’t actually address the main factors which delay processing time of many applications. These factors are the inordinate amount of time taken by the Australian Security Intelligence Organisation to complete security checks (Law Institute of Victoria, 2005:21-2) and the inability of individuals with non-refugee protection claims to get those claims addressed until after they have received a negative decision from the RRT (Law Society of South Australia, 2005:13-14).

Release on Bridging Visa

Since 1994, the Minister for Immigration has had the power under section 72(1)(c) of the Migration Act to determine that an otherwise ineligible person is eligible to apply for a bridging visa if that person has been in immigration detention for more than six months since lodging a protection visa application without a primary decision having been made, and the Minister considers a determination to be in the public interest. The power is personal to the Minister and its exercise is non-compellable. At the time of writing, only four individuals had ever benefited from the exercise of this power.

The Australian Government often asserts that the long-term detainee population consists wholly of individuals who are not entitled to any substantive visa but are nevertheless trying to avoid removal by, among other things, misusing Australia’s legal processes. This is simply untrue. In the past two years there have been over 100 cases of long-term detainees eventually being granted protection visas (Rural Australians for Refugees, 2005; Georgiou, 2005).

Quite aside from the merits of their case, however, some unlawful non-citizens are for practical reasons unable to depart Australia. The Migration Regulations have been amended with the stated purpose of ameliorating the plight of some of these individuals. Since 16 June 2005, persons in immigration detention whose ‘removal from Australia is not reasonably practicable’ for the time being have been eligible for the grant of a Removal Pending Bridging Visa (RPBV), provided the Minister is satisfied that the person ‘will do everything
Immigration Detention Reforms

possible to facilitate’ their removal from Australia and any visa applications (with specified exception) have been finally determined. They also have to meet character and national security requirements. Moreover, only detainees invited to do so by the Minister for Immigration are able to apply for a RPBV. As at 14 July 2005 the Minister had invited 58 individuals to apply for a RPBV and 42 of them had taken up the invitation (Cobb, 2005). However, as at 13 February 2006 there had only been a total of 31 grants of RPBVs (Mann, 2006).

If granted, a RPBV enables the holder to remain at liberty in the community until removal from Australia becomes reasonably practicable. Importantly, the holders of RPBVs are given some social assistance entitlements and have the right to work.

Exercise of 195A Migration Act Power

As already mentioned, the new section 195A of the Migration Act gives the Minister a personal and non-compellable power to grant a visa to a person detained under section 189, if 'the Minister thinks that it is in the public interest to do so'. Thus the Minister is now able to bring any long detention situation to an end by granting the detainee a bridging visa or indeed a substantive visa regardless of whether the person in question meets the usual visa criteria.

Detention Review

The final aspect of the mandatory detention regime, which has been subject to extensive criticism, is the unavailability of an independent body to review an individual’s detention at regular intervals and to order release if detention is not, or ceases to be, ‘a necessary and proportionate means of achieving a legitimate aim’. The June 2005 reforms have not changed this.

The Secretary of the Department of Immigration is required by the new Part 8C of the Migration Act to report on the circumstances of a person’s detention to the Commonwealth Ombudsman once the detention period reaches two years and at six-monthly intervals thereafter. At 3 February 2006, DIMA had provided an initial report to the Ombudsman on the 152 detainees who had reached the two year mark before 29 June 2005 and provided second reports on the 51 out of those 152 detainees who still remained in detention (Commonwealth Ombudsman, 2006:1). DIMA had also reported on 59 detainees who had reached the two year mark after 29 June 2005 (Commonwealth Ombudsman, 2006:1). As soon as practicable after receiving a report, the Ombudsman is required to assess the appropriateness of the detention arrangements and to provide that assessment to the Minister. The assessment may include any recommendations the Ombudsman thinks fit to include — for example, a recommendation that the person should be held in another form of detention or that the person should be released into the community on a visa. However, the Minister ‘is not bound by any recommendations the Commonwealth Ombudsman makes’. The only thing that
the Minister must do is table a modified version of each assessment in Parliament within 15 sitting days of receiving it.

At 3 February 2006, the Ombudsman had provided a total of 43 assessments to the Minister (Commonwealth Ombudsman, 2006:1). At 9 February 2006 the Minister had tabled 14 assessments together with statements indicating that there had been one removal, four grants of substantive visa and one other release from detention in advance of the Minister receiving the Ombudsman’s recommendation. In two cases, release had occurred after receipt of the recommendations but it was not clear that release had occurred because of the recommendations. In the remaining cases the Minister was considering the action recommended by the Ombudsman. It therefore remains to be seen to what degree the Minister will feel politically constrained to give effect to the Ombudsman’s recommendations.

On the plus side the Government has provided the Commonwealth Ombudsman with additional funding of $12.8 million over four years in order to discharge the enhanced supervisory role in relation to immigration detention and other immigration matters (Vanstone, 2005d). In addition, legislation designed to facilitate the Ombudsman’s investigative activities commenced in December 2005. On the minus side, the Government has a track record of ignoring recommendations of ‘toothless’ watchdogs that have thus far attempted to hold the Government accountable for the administration of immigration detention, including recommendations made by the Commonwealth Ombudsman (Taylor, 2000). Another minus is the fact that the detention review mechanism kicks in only after a person has been in immigration detention for two years. As Laurie Ferguson (2005) has pointed out, Cornelia Rau was mistakenly held in immigration detention for ten months not two years, yet the Minister for Immigration (Vanstone, 2005a:23) is on record as saying: ‘Can I give an absolute guarantee that the detention of Cornelia Rau did not in any way further damage her condition? Of course I cannot give that guarantee’.

**Looking Forward**

Ever since the introduction of the mandatory immigration detention regime, its opponents have suggested alternatives to achieve the legitimate ends without imposing as high a cost on the individuals concerned. The suggested alternatives allow for unauthorised arrivals to be detained as a matter of course for a brief initial period in which identity, health, character and national security checks can be undertaken. Upon expiry of the initial detention period, the suggested alternatives require continued detention to be justified on the facts of the particular case as a necessary and proportionate means of ensuring community protection or availability for removal. Further they require that an individual’s detention be reviewed by a court at frequent intervals, with the court empowered to order release if it finds that detention has become unnecessary or disproportionate. The suggested alternative models require that all less restrictive means of achieving legitimate ends, such as release subject to reporting obligations or provision of
Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Immigration Detention Reforms

Surety or other conditions, be found unable to achieve the same ends in a particular case, before a finding can be made that detention is necessary. The more fully developed alternative models include a role for alternative places of detention to deal with those individuals who need not, or should not, be subjected to the full rigours of an IDC environment, but who for legitimate reasons cannot be released from detention either.

In addition to these features, the Reception and Transitional Processing (RTP) Model developed by the Justice for Asylum Seekers Alliance also contemplates the establishment of a case management system administered by an independent service provider. This includes the assignment of a case worker to each asylum seeker with the specific role of providing information, advice and assistance to the asylum seeker from the time of arrival until a substantive visa is granted or removal is affected. Among other things, the case worker would counsel the asylum seeker throughout the determination process to prepare that person for all possible outcomes, including removal from Australia in the event the claim is rejected. The premise is that the goal of ensuring that persons with no right to be in Australia are available for removal can be achieved just as well by helping individuals to accept a removal decision as by employing detention and other coercive means. The RTP Model is inspired by the model used in Sweden which has in fact been highly successful in procuring the voluntary return of rejected asylum seekers (Justice for Asylum Seekers Alliance, 2002:25).

In summary, human rights compliant alternatives to mandatory detention already exist on paper. The task which remains to be accomplished is that of turning the theoretical alternatives into actual policy. One positive outcome of the Cornelia Rau scandal is that this may conceivably happen. DIMA, in partnership with community sector agencies, is presently developing a community care model to be piloted in Sydney and Melbourne with 400 to 500 selected immigration detainees and others of ‘particular compliance interest’ (DIMA, 2005h: 6-7; DIMA, 2005c:66; Correll, 2006). At the time of writing the 12 month pilot was about to commence (Correll, 2006). According to DIMA, the community care model will be an augmentation of a wider case management framework being rolled out across its service delivery network to ‘holistically’ manage clients, ‘particularly those who are vulnerable or have complex circumstances’ (DIMA, 2005c:66). DIMA says ‘the expectation is that case management would include managing the transition to settlement or removal (DIMA, 2005c:66). Entrusting the case management function to DIMA officers rather than an independent service provider as called for by the RTP model is problematic. There is potential for officers to face serious conflict of interest if they are required to discharge compliance and welfare duties simultaneously. Leaving aside this reservation, however, the community care model pilot itself is welcome as it will additionally involve the provision of services ‘based on individual need which may include housing, assistance with living needs, counselling, medical and mental health intervention’ (DIMA, 2005c:66) while DIMA decisions are made regarding removal, or where appropriate, temporary or permanent settlement’ (DIMA, 2005h:6-7). The Government will consider wider implementation of the model.
once the pilot has been assessed (DIMA, 2005h:7). While the mandatory detention policy continues to apply, the changes that have been implemented do represent a small improvement from a human rights perspective and those being considered hold out the possibility of significant improvement in the future.

References


Baban v Australia (2003), Communication No. 1014/2001, 18 September.


Commonwealth Ombudsman (2005), ‘Briefing for Dr Savitri Taylor’, 5 December (Copy on file with author).


Department of Immigration (2005b), ‘Briefing for Dr Savitri Taylor’, 6 December (Copy on file with author).


Economic Framework for Melbourne Traffic Planning

Harry Clarke and Andrew Hawkins

Melbourne has the size, population density and car ownership levels of many North American cities but the levels of public transport infrastructure seen in European cities. Although possessing above-average public transport infrastructure it still faces problems of dealing with high car use and socially-costly traffic congestion that amount to losses of billions of dollars annually. The costs include wasted travel time, increased stress and fatigue, excessive fuel and maintenance costs, lower worker productivity, increased transport capital requirements, reduced business competitiveness and deterioration of an otherwise aesthetically pleasing environment. Although these costs can be reduced with demand-side policies, roads are the only public utility in Victoria not targeted for efficient pricing.

In this paper, economics-based policies are proposed for ameliorating Melbourne’s traffic congestion. A suite of demand-side strategies is suggested including a cordon pricing scheme for the central city, parking policies which address boundary problems on the cordon perimeter and pricing of major arterials and ring-roads. These policies should be accompanied by ‘traffic calming’ measures on smaller urban roads. Auxiliary public transport pricing reforms are advocated promoting competition. Current policies of concentrating urban expansion on city boundaries are criticised. The proposed policy mix is a stop-gap bringing congestion closer to efficient levels thereby facilitating transfer to comprehensive electronic pricing.

Comprehensive ‘first-best’ pricing of Melbourne’s roads by satellite monitoring is impractical due to high start-up costs relative to congestion. However, inexpensive demand-side policies can reduce congestion costs. This paper exposit such policies. Section 2 overviews Melbourne’s transport system. Section 3 analyses congestion policies emphasising demand-side tactics and cost-effective cordon pricing of the city centre. Supply-side strategies, information policies, indirect demand-side strategies and other policies are discussed. Section 4 outlines a plan to relieve radial and central traffic congestion by curb-side pricing, cordon pricing and ‘traffic calming’. Section 5 explains why direct demand-side strategies are generally unsuited to reduce cross-town congestion. Section 6 offers conclusions and final remarks.

Harry Clarke is Professor and Head of the Department of Economics and Finance, La Trobe University and A. J. Hawkins is an analyst and former La Trobe University student.
Melbourne’s Transportation

Melbourne’s sprawling metropolis wraps around Port Phillip Bay, stretching 50 km east to west and 70 km north to south. It is a large city covering 1,700 square kilometres and is home to 3.3 million residents.

Melbourne splits into two zones, a central-cum-inner-suburban region and an outer-suburban periphery. The CBD and inner suburbs were developed during the nineteenth century before the advent of motor transport. The compact CBD has high population density and few high capacity roads. As Melbourne’s commercial centre the CBD hosts one third of the city’s jobs. On an average weekday over half a million people travel within the City of Melbourne, with sixty-three per cent of these trips by car (City of Melbourne, 2000:4.3). The CBD is the largest retail centre with turnover three times that of the largest alternative free-standing shopping mall (Cervero, 1998:334).

Melbourne’s inner suburbs are compact, mostly affluent, residential areas with pockets of commerce and industry. Residents live in high population densities of 45-60 people per hectare, have easy access to local commerce, amenities and the CBD. Because they have convenient access to public transport, they are less car-dependent than those in the periphery (Cervero, 1998).

Melbourne’s periphery developed with the urban railway built in the late nineteenth century, which allowed residents to live in large freestanding houses distant from the workplace. This created low-density residential corridors along rail lines. The advent of the motorcar facilitated cross-town travel and land between rail corridors was eventually developed as urban sprawl.

The periphery consists mainly of detached single-family homes with gardens. Residential areas are often segregated from industrial and commercial areas with population densities of approximately 14.9 persons per hectare (pph). These are among the lowest urban densities outside the United States. Montreal and Toronto, cities with similar populations, have population densities of 33.8 pph and 41.5 pph respectively (Department of Infrastructure, 2002:58). The long distances that must be travelled by residents in Melbourne’s periphery rule out walking and bicycling while low population densities undermine economies required for mass transit. The periphery has high car dependence.

Since the 1960s Melbourne has experienced high growth in car dependence. From 1961-1991, while population grew 53 per cent and the total road network by 44 per cent, vehicles registered grew 237 per cent and ‘passenger boardings’ on buses, trains and trams fell 31 per cent (Kenworthy and Laube, 1999).

Of all trips in Melbourne 14 per cent are to work, 10 per cent are education-related, 20 per cent are for shopping and 26 per cent are social or recreational. Many journeys are cross-town (only one third of work-related journeys are towards the city centre and, of these, only one third by public transport) which poses specific problems. However, marginal congestion costs occur predominantly in Melbourne’s CBD (59 per cent of the total), followed by its inner arterial roads (20 per cent), freeways (14 per cent) and other arterials (7 per cent) (Litman, 2002).
Transport infrastructure

Public transport infrastructure has evolved as ad hoc adjustments to immediate needs at high cost. Responsibility for roads is shared, with the Federal Government responsible for national highways, and 32 Local Governments responsible for tributary roads. The State agency responsible for management of primary and secondary roads is VicRoads.

Melbourne’s road network follows a disjointed grid. Each major road forms a grid square composed of tributary roads. Many major cross-town roads do not run the width of the city so drivers need to change roads to make cross-town journeys. The arterial roads were later additions that radiate out from central Melbourne.

Most of Melbourne’s roads are not charged for directly and many are congested at peak times. Excluding charges, road users only pay for approximately 44 per cent of the social costs they impose with non-motorists bearing around 32 per cent (Litman, 1999). Motorists do not bear congestion, pollution, noise, road construction, maintenance and external accident costs. Drivers also benefit from fringe benefit tax concessions that favour salary packaging linked to private transport. Where road use is charged for at point-of-use, as on CityLink’s Tollway or where it will be practised as on the Mitcham-Frankston Project, it is cost-recovery not efficiency-based. Apart from minor concessions for heavy vehicles travelling late at night and lower rates for passenger vehicles on weekends, there is little differentiation between peak and off-peak tolls.

Melbourne has one of the world’s largest urban train networks and the fourth largest tram network. Yet only 9 per cent of motorised trips are on public transport (Department of Infrastructure, 2002:39). Indeed many Melburnians live beyond walking distance to public transport and only a small proportion of jobs can be reached by public transport. Hence, although heavily subsidised, demand for public transport is low. In the past, transport deficits have been over a billion dollars per year or $1,670 per passenger (Industry Commission, 1994:75-77). Subsidies will be $2.3 billion over the five years from 2004 (Heasley, 2004).

For many, public transport is an imperfect substitute for car travel being less convenient and more expensive. Public transport, moreover, caters primarily for radial trips, accounting for 36 per cent of such trips compared to 9 per cent of total trips (Department of Infrastructure, 2002:39).

Recently the Government privatised public transport as regulated and subsidised monopolies. The urban train network is operated by the private corporation, Connex, the tram network by Yarra Trams and the buses by several private companies. Trains provide convenient travel to the CBD, where 85 per cent of jobs are located within a five minutes walk of a city-loop station (Cervero, 1998). Melbourne has no cross-town railway lines. Connex currently receives a subsidy of $345 million per year (Yarra Trams, 2004). The tram network forms a grid covering the inner suburbs, with occasional offshoots into the periphery. The
high patronage of the trams for local journeys is due to the regular services provided and surrounding high population densities. Nevertheless, Yarra Trams receives a subsidy of $112 million per year (Yarra Trams, 2004).

Buses are concentrated in areas developed after the advent of cars but not serviced by rail or tram. Buses usually travel cross-town to train stations, schools or peripheral shopping centres where they often follow meandering routes through suburbs. Low patronage is due to irregular and slow services reflecting low population densities. There is a lack of financial incentives for operators to seek extra passengers since they are mainly paid on the basis of distance operated with franchises reviewed on a ‘grandfather’ basis.

**Causes of traffic congestion**

Melbourne’s congestion results from lack of effective urban planning and transport policy, inefficiently priced transport and urban development subsidies. Supply-oriented civil engineers at VicRoads have dominated transport management focusing on predicting demand at subsidised private transport prices and then seeking to provide infrastructure. Little attempt has been made to manage demand.

Road supplies in central Melbourne are limited because of the high costs of acquiring intensively developed land. Creating peripheral roads makes more sense but raises ‘triple convergence’ issues¹, discussed below in 3.6. More roads release latent demands for travel, which tends to restore pre-existing congestion. Thus, providing roads in the periphery reduces travel costs and increases the appeal of cheaper peripheral housing, resulting in urban sprawl, which offsets the benefits of the initial road expansions. Commerce and industry follow residents in locating far from public transport corridors, further increasing demands for cross-town travel.

With rising incomes and declining motoring costs public transport has come to be seen as a poor substitute for car travel. Public transport usage is known to be price-inelastic undermining the effectiveness of subsidy policies.

**Traffic congestion costs**

Congestion imposes costs. Increased travel time cuts into leisure and work time with costs at about 50 per cent of a commuter’s wage (Small, 1992). Other costs include stress, fatigue and increased costs of fuel and car maintenance. There are also specific costs to the transport services sector, which accounts for 10 per cent of gross product in Victoria. Indeed freight movements within Melbourne cost $2 billion annually (FDF Management, 2001). Congestion reduces the

¹ Downs’ (1992) *fundamental law of traffic congestion* states that peak hour traffic congestion rises to meet increases in capacity because initial reductions induces users of alternate routes to return (*spatial convergence*); fewer people tend to avoid peak-time travel (*time convergence*); and users of alternate transport modes return to cars (*modal convergence*).
productivity of transport capital by increasing its input requirements. This decreases economy-wide productivity through inefficient use of fuel, automobiles and creating slower, less reliable road-based public transport. Congestion increases emissions of dangerous pollutants because, with congestion, fuel consumption increases by up to 30 per cent (Litman, 2002). Moreover, pollution, vibration and noise caused by congestion decrease the environment’s aesthetic appeal.

There are no accurate, up-to-date estimates of Melbourne’s congestion costs. Earlier estimates suggest costs of $2 billion dollars per year (Industry Commission (1994:220) but these are gross figures not deadweight losses. They compare travel times with congestion times under free-flow (BTCE, 1996), exaggerating net costs in one sense but understating them by ignoring indirect costs such as pollution. The same gross figures are presented in Port Jackson Partners (2005), which forecasts congestion cost increases from $2.7 billion to $8 billion over 2001-2021.

Observed traffic delays suggest there are significant deadweight losses to Melbournians from congestion. BTCE (1996) estimates for 1995 suggest deadweight losses of $466 million. For 2005, assuming (i) 11 per cent population growth, (ii) 23 per cent inflation and (iii) 12 per cent growth in the per capita car fleet size, implies a rough estimated cost of $698 million. Supply improvements are ignored here as are nonlinearities in the relation between vehicle numbers and travel times. VicRoads estimated travel times on sections of the Westgate Freeway, the Western Ring Road, CityLink and the Eastern Freeway doubled from 1994-2003 (Silkstone, 2004). Further, this estimate ignores growth in vehicle usage intensity: experience suggests substantial increases in intensity reflecting increased travel distances to work (Kenworthy and Laube, 1999).

**Congestion Policies**

The existence of congestion suggests gains can be realised with active policy. The first-best intervention is to force drivers to internalise all social costs of travel by pricing road use to ensure efficient usage. If this is impractical, potential gains remain possible through policies that reduce road use to efficient levels.

Pricing policies must be socially acceptable and easy to use. Controls on road use promote inefficiency while market approaches, relying on price, select out users with maximum willingness-to-pay.

Pricing technologies must also be reliable and flexible enough to allow for fine-tuning and expansion. Tolls should account for the fact that vehicles impose different costs by measuring congestion impacts in terms of passenger car equivalents, so impacts are related to the size of vehicles. Policies also need to account for the difficulties low-income earners face because of lack of transport alternatives to cars.

Efficient pricing, whether exact or approximate, offers a double-dividend. It improves resource allocation and provides a superior source of revenue for States, such as Victoria, that currently depend on, for example, regressive gambling taxes.
Congestion charges will fall heaviest upon the time-poor affluent who often live in congested and polluted inner city areas but will allow the abolition of some socially-costly taxes.

**Comprehensive pricing**

The likely availability, within decades, of comprehensive first-best road pricing technology provides leeway for current transport policy. Rather than attempting to internalise all costs now, one can instead look for *stop-gap* strategies to contain congestion until comprehensive pricing is available. Comprehensive pricing remains a longer-term goal.

**Kerbside pricing**

Kerbside systems can only address congestion on major roads. These rely on interrogators located next to roads, or gantry points above roads, to pick up signals from transponders or tags in passing vehicles. A computer analyses usage and debits driver accounts to reflect tolls.

Unfortunately such partial pricing creates incentives to divert from priced to unpriced roads. The possibility of ‘rat running’ (touring around gantry points to avoid charges), implies that, with unpriced substitutes, priced roads should have greater capacity but lower tolls (Choe and Clarke, 2000). McDonald (1995) suggests there are big impacts of unpriced substitutes on gains from efficient pricing with toll revenues being only 25-50 per cent of those without priced substitutes and yielding only 8 per cent of the gains. Given administrative costs the case for efficient pricing can be destroyed with such substitutions. Further, if motorists change destination choices to uncharged locations, congestion itself relocates toward uncharged areas.

**Cordon pricing**

The priced cordon reduces congestion by charging an entry fee for access to an area creating economies by monitoring along a perimeter, rather than on individual roads. The entry fee can be varied to capture marginal social costs with payment made electronically by kerbside charges, tollbooths or pre-purchased windscreen stickers. London has recently successfully introduced cordon pricing. The cordon met congestion reduction targets and showed that large-scale road pricing projects can gain reasonable popularity.

Priced cordons can be an efficient way of capturing the social costs of road use because, unlike parking charges, they capture *through* and *terminating* traffic. To minimise adverse distributional impacts on low-income, high travel value citizens, alternative lower-cost public transport should be available for those unwilling to pay the cordon charge.
Information

Transport reform is politically sensitive. A prerequisite is public education on the costs of congestion, the case for transport subsidies and the logic behind demand management. Such information, as a public good, is underprovided. The case for efficient pricing needs to be explained and distinguished from cost-recovery-based construction tolls that fund supply-side initiatives such as CityLink and the Mitcham-Frankston Tollway. Construction tolls target average rather than marginal costs and do not target congestion. They are inefficient if only because they are uniform over the day. They should be higher at peak periods to reflect higher congestion costs. The popularity of efficient pricing can be improved by guaranteeing revenue neutrality by cutting other taxes once charges are introduced. Harrington et al (1998) show that when authorities guaranteed taxes collected would be offset by reductions elsewhere, support for road pricing in California rose by 7 per cent.

There remain political economy constraints on reform. Efficient pricing will not appeal to peripheral, car-dependent areas, in politically-important electorates. Huge public transport subsidies may likewise be difficult for citizens to accept. The target of education campaigns is to reduce such barriers to efficient pricing. While local residents may seek ‘free-rider’ benefits by leaving roads, such as the Mitcham-Frankston Tollway unpriced, the point should be made that appropriate charges would ensure efficient use of this road rather than cost-recovery, and that the same charges will be applied to other major roads. With appropriate redistributions of toll revenues all citizens are better off with such charging.

Finally, giving people choice improves the acceptability of pricing. Harrington et al. (1998) show that restricting pricing of a road to a single lane, leaving a more congested alternative to those wishing to avoid tolls, increases support for pricing by 9-17 per cent.

Non-congestion externalities

Motor transport produces by-product and congestion externalities including pollution, noise, vibration, accident risk and the aesthetic degradation of neighbourhoods. These costs are roughly proportional to the distance a vehicle travels and its weight, which is most conveniently related to fuel consumption. A fuel surcharge can internalise such externalities for urban if not non-urban commuters although, even for urban commuters, surcharges are blunt because of variations in engine efficiency. United States studies suggest policies that fully internalise these externalities can double motorist private variable costs (Litman, 1999:3). Such charges have a non-specific congestion impact.

Supply-side measures

Although popular with politicians and engineers, supply-side strategies cannot, in themselves, resolve congestion in the long-term. This is demonstrated by the
persistence of congestion in cities, such as Los Angeles, even given considerable supply-side investment. Increasing road supplies to the point where demand never exceeds supply is impractical in congested cities, given difficulties in scaling-up road capacity to keep pace with growth in travel demands. Supply measures are at best an expensive form of congestion relief. Their appeal stems from free-rider externalities and the absence, until recently, of cost effective technology to charge for road use. Individual motorists derive specific benefits from new roads but, as with public transport, costs are borne by the average taxpayer. Increases in road supply have weak effects on pre-existing congestion because of ‘triple convergence’ issues. The effective way of avoiding these adverse effects is to address congestion directly by charging.

**Indirect demand-side policies**

If it is technically and politically difficult to directly internalise congestion costs, one can set charges on goods complementary to private car use such as fuel, registration, parking, stamp duty, spare parts and, indeed, car prices themselves. Charging for complements increases motoring costs and lowers demands for private motorised trips. Such charges can focus on usage, and perhaps location, and should reduce fleet sizes. However, this approach does not specifically address congestion and creates inefficiencies by leading to under-utilisation of uncongested roads.

Indirect charges can target fuel (or even spare parts usage) which is roughly proportional to distance, pollution and risk. Since elasticities of usage with respect to fuel is low this implies low deadweight losses but, to significantly impact on external costs, fuel taxes must be high. Currently about half of gasoline costs accrue to government. Such charges proxy for maintenance, environmental and health externalities and encourage purchase of fuel-efficient vehicles. They only marginally target congestion. Such policies impose costs on all drivers regardless of the valuation or location of travel resulting in inefficiently few journeys. Increasing spare parts or new vehicle charges might undermine traffic safety as drivers postpone repairs or vehicle upgrades.

Fuel charges can be designed so heavily congested areas get charged most, but this requires large distances between congested and uncongested regions to prevent inefficiencies such as ‘fuel fetching’ (driving to distant locations to purchase fuel). It is impractical to levy such indirect charges in Melbourne.

Another approach is to levy charges on city parking that is related to the social costs of car travel. Parking charges can be peak-load priced to internalise social costs of peak hour travel but to not impact as much on travel at other times. However, parking charges increase congestion by encouraging motorists to roam in search of low parking fees and increase chauffeured and ‘through’ traffic.

**Alternative transport**

Alternatives to private road transport include walking, bicycling, use of jitney, taxi, bus, tram or train. These alternatives may not compete with cars because cars
offer space, time and route freedom and because many of the costs cars generate are external. Attempting to reduce congestion by subsidising alternative transport modes requires high cross-price elasticities of substitution which are, in fact, low. Subsidies themselves are ‘first-best’ policy prescriptions for industries dominated by large fixed costs, such as railways, where economies of scale are significant. However, if costs are primarily variable, subsidies imply inefficiencies because use then becomes inefficiently high.

While information on using congestion-inducing private cars to travel is a private good, information on congestion-efficient vehicles, non-motorised travel and public transport is a public good that may be under-provided. Marketing campaigns, such as the City of Perth’s ‘Travel Smart’ program, which informs travellers of the costs associated with travel decisions, can modify transport demands by moral suasion.

If congestion charging is introduced some travel would shift to subsidised public transport. This operates at close to capacity at peak times, so road pricing reform calls for increased provision and subsidies. Only when all urban transport is efficiently priced at social marginal cost will citizens bear the cost of their travel choices and travel and locate their residence, workplace and recreational activities efficiently.

The extent to which public transport should be subsidised depends on scale economies and externalities. Infrastructure (rails, roads and signals) are a natural monopoly subject to scale economies but rolling stock is not. Vehicles can be operated by relatively small, independent, unsubsidised, unregulated firms. Indeed firms can operate on the same routes, offering close substitute services. What is objectionable is not government enterprise *per se* but lack of competition. Privatised monopolies need be no more efficient than government monopolies and efficiency can fall with only limited competition as network effects are undermined by removing central coordination. This is particularly so if strong unions remain and since the Melbourne privatisations, union membership has risen after large falls in the 1990s (Masanauskas, 2002). Also with only a few suppliers, cartels form to inhibit competition.

With government regulation, regulated monopolies or oligopolies merely add to the list of special interest groups trying to rent-seek. This has been the liberalisation experience of the Chilean and English bus services. The Chilean bus service was deregulated in 1979. Bus companies quickly formed cartels to exploit the Government’s absence from the market, service levels deteriorated, prices increased and traffic congestion dramatically increased. The only beneficiaries were the cartels. The English bus service suffered a less severe fate than the Chilean, when it was deregulated. Service levels deteriorated and adverse network effects occurred as common ticketing disappeared so prices increased leading to increased traffic congestion (Mees, 2000). Some might argue that ‘Demsetz competition’ involving exclusive licenses that are periodically re-tendered may work better than open-slasher competition.

Note the equity implications of subsidies. While public transport is used mainly by the less affluent, some beneficiaries of subsidies are radial commuters.
heading toward well-paid CBD and inner city jobs. Others, less affluent tend to commute across-town to jobs in the periphery using private vehicles. The distributional impacts of subsidies must be calibrated. Providing such information is a research priority.

Finally, note that the benefit of public transport subsidies and regulation is absorbed by declining productivity as strong unions pursue excessive wage rises and bloated management structures continue to allow inflexible working conditions and over-investment in uneconomic services and infrastructure. In short, interest groups rent-seek imposing a burden on the majority of Melbournians who rarely use public transport. The best protection against such rent-seeking is competition.

Urban planning

Policy can seek to increase population densities in areas served by transport corridors. This reduces required private vehicle travel and increases demand for public transport. However, over the next thirty years, most of Melbourne’s new urban growth is planned for the fringe districts of Melton, Casey, Wyndham, Whittlesea and Hume (Department of Infrastructure, 2002) so population densities will hardly alter given current urban designs.

Clustering retail, employment, cultural and social institutions together in peripheral residential suburbs that are well served by public transport, can promote alternatives to the private car. However, this has not been the case with suburbs such as Box Hill, a large employment hub with 15,000 workers. While well-serviced by bus and train, only around 7 per cent of its workers use public transport (Cervero, 1998:328). Moreover, creating large commercial developments in residential suburbs, to reduce traffic pressures, can encounter local community opposition as public protests on developments in the Camberwell Station area highlight.

Traffic calming and other policies

Traffic calming involves slowing traffic to make streets safer and more useful for general residential life. This is achieved by altering road widths, reducing speed limits, building road obstacles such as speed bumps, and by encouraging slower driving. By adding street furniture and vegetation to make the environment visually attractive and less car-oriented, traffic calming reduces external costs of car use and accidents. It also limits local traffic flows and spill-overs onto local roads from the pricing of major roads alone. Even though calming reduces ‘rat-running’ and improves amenity values, it imposes costs by slowing down local traffic, increasing fuel costs and so on.

Other policies include ride-sharing which has the potential to reduce congestion by reducing the number of vehicle trips. This requires common origins, destinations and travel times. High occupancy vehicle lanes encourages ride-sharing by restricting supply which, however, has some effects in worsening congestion.
Introducing greater flexibility in working hours reduces travel demands during peak periods but also reduces the coordination benefits from synchronisation of work hours. Triple convergence will offset many flexi-time strategies and such strategies would need to be organised city-wide to significantly reduce congestion. With increasing use of mobile phones and the internet this may become more feasible, perhaps by allowing work one day a week from home. Another way of re-organising times to reduce congestion is to alter school hours to allow chauffeured trips to school to be carried out further from peak times.

*Park-and-ride* facilities allow travellers to drive to public transport, park and then transfer to public transport. Melbourne exploits park-and-ride, with over 23,000 car spaces provided near stations in addition to on-street parking (Mees, 2000:231).

Finally, *upgrading tow-truck procedures* by introducing roving recovery vehicles can reduce accident-induced congestion delays.

**CBD and Radial Traffic Reforms**

Melbourne’s most intense congestion occurs on radial and central roads which should be addressed by kerbside and cordon pricing, respectively. Such policies are simple and user-friendly with electronic monitoring allowing tolls to be levied at low cost. These policies do not interfere with traffic flows and pricing can be varied to reflect short-run social costs. While impacts are concentrated on high-income, high time valuation city-based workers, adverse distributional impacts, where they do occur, can be negated using compensations from toll revenues. Indirect demand-side strategies, such as traffic calming, can minimise boundary problems.

*Motorway pricing*

Melbourne’s radial motorways (Eastern Freeway, South-Eastern Freeway, Calder Freeway, West Gate Freeway, Princess Freeway, Western Freeway, Tullamarine Freeway, Hume Freeway and CityLink) spread from the city centre like spokes on a wheel. They are intended to provide high speed uninterrupted access to the city from the periphery but are congested during peak periods. These motorways are suited to kerbside pricing since their high traffic volumes justify the transaction costs involved and alternative roads can be managed as imperfect substitutes. The limited number of entry and exit points minimises the gantries required for monitoring. Indeed motorways would only require about ten gantry points, many of which could be strung from bridges.

The CityLink road network is the only free-flowing axis linking the South-Eastern Freeway, West Gate Freeway, Tullamarine Freeway and Melbourne’s ports. It is inefficiently priced with excessively high tolls leading to low patronage and excessive congestion on neighbouring substitute roads.

Melbourne already has experience with the technology, management processes and enforcement required to implement road pricing, via CityLink. This
technology has a high capture rate, needs little space, causes low environmental impact, has low fixed staff costs and low marginal costs. Gantry points, though expensive, can be minimised because of the limited entry/exit points: see Table 1 below. Melbourne residents are familiar with this technology. Many residents already have an E-tag and scope exists to create economies of scale in road pricing management by outsourcing management of new arterial roads to firms such as CityLink.

Table 1: Arterial and Cross-Town Roads that Can be Priced in Melbourne

<table>
<thead>
<tr>
<th>Existing roads with construction tolls</th>
<th>Approximate number of entry/exits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monash Freeway</td>
<td>15</td>
</tr>
<tr>
<td>Link Road</td>
<td>8</td>
</tr>
<tr>
<td>Bateman Avenue</td>
<td>2</td>
</tr>
<tr>
<td>Burnley Tunnel and Grant Street</td>
<td>2</td>
</tr>
<tr>
<td>Mitcham-Frankston</td>
<td>Under construction</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Roads not yet tolled</th>
<th>Approximate number of entry/exits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern Freeway</td>
<td>10</td>
</tr>
<tr>
<td>Western and Eastern Ring Roads</td>
<td>17</td>
</tr>
<tr>
<td>West Gate Freeway</td>
<td>7</td>
</tr>
<tr>
<td>Tullamarine</td>
<td>10</td>
</tr>
<tr>
<td>Calder</td>
<td>9</td>
</tr>
</tbody>
</table>

Cordon pricing

Cordon pricing operates with the same logic as kerbside systems but has the advantage of offering economies by monitoring road use only at the periphery of an area rather than monitoring each individual road within an area.

Melbourne’s cordon should cover Melbourne’s CBD up to, but not including, Victoria St, diverting down Clarendon St and Hotham St, running down Punt Rd, down Alexander Ave to Anderson St, Domain Rd, across St Kilda Rd to Park St, then down Kings Way, Sturt St, Power St and along Clarendon/Spencer St, up Dudley St around the inside of Victoria market and returning to Victoria St. Areas such as the Crown Casino car parks should lie inside the cordon to prevent motorists parking there and walking in thereby creating congestion on the cordon.
boundary. Boundary problems are further minimised by positioning the periphery within commercial and industrial areas that have significant parking restrictions and rigorous enforcement.

A plan of the proposed cordon is in Figure 1. This design captures most of the congested core and requires only about twenty gantry points so start-up costs are well below London’s £200 million cordon, with its over 200 entry and exit points.

**Figure 1: The Proposed City Cordon**

Through traffic can bypass the cordon by running around perimeter roads or using CityLink. Public hospitals, schools, and most emergency services remain on the periphery of the cordon, reducing claims for charging exemptions. The cordon excludes the entertainment and shopping precincts of Lygon St, Brunswick St, Victoria St, Toorak Rd, Swan St and Docklands and the shopping precinct of Victoria Market.

Because Melbourne’s centre also has tight parking restrictions it does not need to adopt an area cordon, like London, where cars are billed merely for being on city roads. Rigorous enforcement of parking restrictions allows Melbourne to operate a periphery cordon where drivers are only charged for crossing the cordon boundary. This internalises the social costs of trips out of and across, the cordon
better reflecting social costs of using road s. As no substitute roads exist, the
cordon charge accurately reflects marginal costs of using city roads. The charges
should vary to reflect the social cost of road use depending on travel time and
vehicle type allowing a more efficient allocation than London’s single price.

Unlike London, Melbourne should not exempt taxis, motorcycles, mopeds,
buses, recovery vehicles, repair vehicles and alternative fuel vehicles, as all
contribute to congestion. All vehicles should pay the external costs inflicted.
Motorcycles, mopeds, taxis and alternative fuel vehicles impose lower costs and
this should be reflected in charges. Taxis and buses impose congestion so tolls
levied should reflect this. The cordon charges internalise external costs associated
with using city roads leaving parking charges directed to capture the additional
specific costs of parking.

Cross-town Traffic Strategies

Improving the efficiency of cross-town travel is a challenge. The economies of
scale that allow direct demand-side strategies for radial travel and within the CBD
rarely exist in the sprawling plethora of cross-town roads. Yet serious congestion
occurs on cross-town roads such as Springvale Road. This is 20 kilometres from
the city centre and provides the main north-south road link in Melbourne’s Eastern
suburbs. A survey of 7,500 RACV members nominated the intersection of
Springvale and Whitehorse Roads as the worst congestion in Melbourne. The next
five most nominated sites were also in outer metropolitan areas

Rail infrastructure is primarily radial as population densities necessary for
mass transit are unavailable for cross-town journeys in the periphery. Land-
acquisition costs limit the potential for supply-side measures so policies are
limited to indirect demand-side tactics and occasionally, when traffic volumes
permit, to direct demand measures.

Roads

Melbourne’s cross-town roads generally do not run the length of the city, have
relatively low capacity, and are frequently interrupted by intersections. The only
arterial road that currently allows uninterrupted cross-town travel is the Western
Ring Road. A significant expansion of cross-town road supply is impractical
given land acquisition costs.

Cross-town roads are unsuited to direct demand-side strategies. Their price
elasticity of demand is high, as close substitutes exist. Also, because entry/exit
points frequently interrupt such roads, the number of gantry points required for
monitoring would be expensive. Road capacities are low so the scale required for
viable road pricing is not present.

The Western Ring road and the soon-to-be-constructed Mitcham-Frankston
tollway are the only cross-town roads with sufficiently high patronage and with
the limited number of entry/exit points required to enable kerbside pricing. Charges should be set below first-best levels since some substitute routes do exist. While cross-town travel can never be as efficiently managed as radial travel, it has the advantage of imposing lower congestion. Cross-town commuters on average have a lower value of time than inner city workers so the costs of congestion are lower.

With comprehensive road pricing, decisions about expanding existing roads and building new roads can be more informed. If road segments remain congested despite high pricing — in short, when they are earning good profits after accounting for all costs associated with their operation — there are grounds for expansion.

**Integrated urban and transportation planning**

An urban growth boundary can limit Melbourne’s transport problems in the city periphery. Restricting expansion halts encroachment into fringe wilderness, agricultural land and public amenity areas. Some of these resources — for example Greenbelt areas conserving biodiversity — are public goods and will be underprovided even if congestion and other externalities are internalised. Imposing a boundary encourages the development of viable satellite and regional cities as substitutes for sprawl. By increasing population densities, public transport and ride-sharing become increasingly viable in the periphery.

Unfortunately most population growth has occurred and is forecast to continue to occur on city fringes where it will be channelled into ‘growth corridors’ (Department of Infrastructure, 2002). Between the 1996 and 2001 censuses Melbourne’s population grew more than 7 per cent with more than half of this reflecting residential growth in outer suburbs such as Caroline Springs in the west, Roxburgh Park in the north and Narre Warren South in the south-east. Only 12 per cent has occurred in the local government areas of Melbourne and Port Phillip (ABS, 2003). The State Government sees intended corridors being serviced by radial transport via the Transit Cities component of its Melbourne 2030 plan (Department of Infrastructure, 2002). This encourages new development, housing, schools and shops near railway stations. However, since car travel in fringe suburbs is primarily cross-town this strategy is unlikely to be effective.

There are criticisms concerning the use of land use policies to address congestion issues. Some claim such policies have a ‘tail wagging the dog’ character. The objective is to make travelling efficient, not to limit it through restrictions on city size. However, given difficulties in applying efficient road pricing in low-density urban areas, land-use restrictions are a sound ‘second-best’ policy. Sometimes too, the imposition of urban boundaries is seen as impractical policy that will be undone by politicians in response to interest group pressures and, indeed, this seems to have already occurred in Melbourne’s west. The answer is to set boundaries that give room for current land-owners to realise
expected capital gains on long-term land holdings or provide financial compensation for lost property values as a result of greenbelt-type policies.

The inner suburbs in Melbourne have a mass transit-oriented public transport system. This allows for quality travel in all directions at virtually all times, as opposed to commuter public transport which only allows for efficient radial commuting at peak hour (Mees, 2000). Internal growth boundaries need to be developed to restrict high-density housing to transit-friendly locations within inner and surrounding suburbs. Reducing building and heritage regulations here would allow more people to live closer to areas of high employment and increase access to transit-oriented public transport.

Public transport

Government subsidies and regulation have not delivered a public transport system that efficiently services cross-town journeys in the periphery. Government regulated monopolies face information problems in attempting to cope with the dispersed times, points of departure and destination characteristic of cross-town travel. There is not enough ‘mass’ to make mass transit work so more specialised transport services are needed. Deregulated markets may offer superior outcomes and better coordination than centralised planning. Numerous unsubsidised and largely unregulated firms, with access to common infrastructure, may offer a superior solution to the dispersed travel requirements of cross-town travellers. Trams, buses, jitneys and taxis could provide substitute services. Competition would keep prices low while market forces would encourage wage discipline, reward good management, creativity, improved customer service and product differentiation. As an example, the Queens Van Plan, in poor areas not serviced by public transport in New York, receives no subsidies and charges patrons a flat $1 fee per ride (Winston and Shirley, 1998).

Government could focus on providing infrastructure, preventing anti-competitive behaviour and ensuring safety standards.

Final Remarks

Traffic congestion in Melbourne is a consequence of inefficiently priced transport services which lead to inefficiently high traffic demands. Little attempt has been made to even collect and collate appropriate data on the deadweight losses of congestion or the changes required to drive road use toward designed capacities. A component of any research strategy should be provision of this data. Research should also address the distributional implications of alternative transport pricing regimes and the issue of scale economies in public transport provision.

Melbourne, along with other cities, has consistently favoured supply-side solutions which impose costs on all taxpayers regardless of their travel decisions. These strategies have electoral appeal since consumers are largely unaware of the scale of the non-internalised travel costs and of the case for subsidising well-managed transport businesses with scale economies that generate positive
externalities. A second priority is to devise campaigns highlighting the benefits of demand-side strategies.

The first-best solution is to reduce traffic to socially efficient levels by comprehensive road pricing coupled with marginal cost pricing of public transport. This is currently infeasible but should remain a long-term objective.

Cost improvements will stem from adopting a suite of direct and indirect demand-side strategies. By-product externalities should be internalised through externality taxes on fuel, and variable registration and insurance charges. Arterial and CBD roads should be priced. Pre-setting urban boundaries would increase population density while internal growth boundaries would focus development, reduce travel, and increase numbers of people with convenient access to public transport. Public transport must be priced to encourage efficient modal substitutions.

A suite of demand-side strategies has the potential to be more efficient and fair in reducing traffic congestion than Melbourne’s current traffic management system.

References


*The authors are grateful to Lyn Martin, two anonymous referees and the Editor for helpful comments. Any opinions express are those of the authors alone.*
Olympic Economics


Reviewed by Ray Trewin

The author is an economist at Germany’s Johannes Gutenberg University, and a member of its Research Team Olympia. This book adds to his extensive list of publications on the Olympics. The dust cover states that ‘academics and researchers of sports economics, international economics, international business and competition will all find this fascinating book of great value’, and that ‘it will also appeal to those interested in the broader context of the Olympic Games and concerned by their commercialization and gigantism’.

Some of the big economic issues covered in this book include: costs versus benefits to the host city (usually jointly liable with the associated Organising Committee of the Olympic Games (OCOG), for debt incurred); the appropriate public and private sectors financing mix; and the increasing economic scale of what is basically a two-week event held every four years. It is disappointing that there is no similar consideration of the political economy of the Games, which might help explain the choice of a particular host city when purely technical considerations suggest that another might be more appropriate. By contrast, Simson and Jennings (1992) present a number of examples, such as excessive gift giving (now curtailed — presumably because of its obvious potential influence on the bidding process); Barcelona strategists splitting support for the Paris bid in 2005 by backing Albertville for the winter Olympics; and technical evaluations by an IOC committee being watered down before presentation).

Although the author makes a number of strong points, several other more debatable views are also advanced. In the introductory chapter, for example, a number of benefits attributed to the Olympics (such as better health through encouraging active participation in sporting activity, increased tourism, and increased trade) would seem to be of minor significance.

Chapter 2 describes the transformations of money values necessary for financial comparisons of various Olympics used in Chapter 3 to identify the shift in financing from the public sector to privately owned corporations. The mix ranges from mainly public (Montreal in 1976) to mainly private (Los Angeles in 1984). This change seems to be related to commercialisation of the Games, and a trend to more efficient private sector involvement in infrastructure development.

Chapter 4 deals with the expanding size of investment in the Games. An example of its consequences is that the debt incurred to run the 1976 Games is still
being repaid by Montreal citizens three decades later. More recent Olympiads have witnessed some constraint on this growth, such as concerns about the ecological impact of constructing large sports facilities. The author proposes a limited solution by way of relying on temporary facilities (for example, extra seating in the Sydney Olympic stadium; a plastic pool in an exhibition hall at the 2001 swimming world championships; and a velodrome in Atlanta).

Chapter 5 outlines methods for analysing the economic impact of the Games. Cost-benefit analysis (from the host city’s perspective) is proposed as the ideal approach. The similarity of revenues and expenditures suggests the importance of intangibles in the conclusions. The author argues that cost-benefit analysis is not possible, however, because of a lack of uniform financial data, so a more limited input-output analysis is attempted. This is restricted to just the tangible aspects; intangible effects and indirect impacts are not considered. Nevertheless, several rules that should be built into any attempt at cost-benefit analysis are listed:

- only additional net spending by tourists can provide a benefit, as distinct from redirected spending by local citizens;
- some Olympic expenditures can crowd out other activities, so the associated costs of transferred resources need to be accounted for;
- whether there is a boom or bust when planned Olympic investment is actually undertaken is important in determining the ultimate crowding out effect;
- inflation that results from the Olympics causing aggregate demand to exceed supply can reduce the overall benefits; and
- the use of a wide range of multipliers (from 1.5 to 3 in pre- and post evaluations of the Los Angeles Games, for example) to estimate secondary impacts is inappropriate.

A more appropriate approach than that set out in Chapter 5 might have been an institutional analysis that sought to explain seemingly irrational economic behaviour where intangibles override the basic economics. Examples include the continuation — despite low public interest and limited participation — of the modern pentathlon (which usually requires new facilities and heavy expenditures such as those associated with caring for horses); the foregoing of higher revenues from cable television in favour of more accessible free-to-air television; and the poorly attended cultural program run in conjunction with the sporting events. (The cultural program cost US$71 million in Barcelona, but only returned US$29 million in revenue.) There is some attempt to explain such aspects (with references to maintaining tradition, status seeking, and the attempt to build public support), but this is not presented in a consistent framework or in a focused manner.

In Chapter 6 the author looks, sometimes uncritically, at the alleged benefits to the host community in terms of increased tourism, exports and foreign investment. For example, he mentions post-Sydney Games research by PricewaterhouseCoopers on leveraging off the Olympics, which included a contribution of US$217 million in new sales and investment purportedly generated by an ‘Australian Technology Showcase’, whereas other research showed no
change in Australia’s attractiveness as a destination for investment as a result of the Games. He does include some reality checks, however. For example, he notes that the Australian Bureau of Statistics reported only 53,000 extra visitors to Australia — compared with the typical number of around 400,000 — over the month of the Sydney Olympics, which was much less than predicted. The chapter also contains a useful discussion (from the host city perspective) of negative impacts on tourism generated by the Games, such as city residents and potential visitors taking holidays elsewhere in order to avoid Olympic disruptions.

Chapter 7 discusses infrastructure investment, emphasising the importance of planning for the post-Olympics utilisation of the infrastructure, to avoid creating facilities that will become white elephants after the Games are over. Although this point should be obvious, it has been ignored by many host cities in their enthusiasm to make a strong impression and thus ensure the success of their bid.

Chapter 8 focuses on revenues. This is the longest chapter, perhaps reflecting the key concern of the IOC — also apparent in the transformation of the Committee for the Protection of the Olympic Symbol to become the New Sources of Financing Committee, under former IOC president Samaranch. The chapter outlines the expansion of IOC powers over host cities and OCOGs, and discusses some countervailing forces. For example, ‘over-commercialisation’ of the Games is held back by public disquiet (especially outside the US) about overly prominent advertising and merchandising, as in Atlanta: over 50 per cent and up to 76 per cent of spectators interviewed after recent Games thought commercialisation was a threat to their future. The IOC has reacted to this by banning advertising and merchandising at Olympic sites, as well as advertising on athletes’ clothing and equipment. Other ‘over-commercialisation’ aspects of concern include the disappearance of amateurism as a result of increasingly large rewards to athletes, which also encourages the use of drugs by some to gain an unfair advantage.

Since 1985 the IOC has limited the number of principal Olympics sponsors (those granted worldwide exclusive rights, within set categories, to incorporate Olympic symbols in advertising), thus increasing its revenues significantly. An interesting constraint to this power is also discussed: ‘ambush marketers’ who associate themselves with the Olympics — say, through sponsorship of well-known athletes — but pay no marketing fees to the Olympic movement.

The anticipated revenues for the 2008 Beijing Olympics presented in this chapter appear to be gross underestimates, reversing the increasing trend evident since Montreal in 1976, as do the anticipated expenditures presented in Chapter 9. It would be interesting to know if these numbers were deliberately underestimated in order to maximise the chances of achieving a successful bid, given a perceived desire for smaller Games. It is hard to imagine that spectators close to the Olympic village will really ride bicycles to the venues, as was stated in the bid.

Chapter 10 discusses the employment impact of the Games, making the important point that new employment generated is just a different way of looking at the boost to economic activity, not an additional benefit. Chapter 11, on the aftermath of the Olympics, deals mainly with concerns about price increases. In the case of the 2000 Olympics the author found no significant price differences
between Sydney and other Australian cities, but this seems a poor test of the Olympics’ impact on prices, as he states elsewhere that the crowding out effect would have a broad Australia-wide impact.

Chapter 12 draws some conclusions, one of which is a recommendation for a uniform accounting approach across all Olympics so that proper comparisons can be undertaken and the staging of the Games improved through a better matching of Games-related revenues and expenditures. The author presents a summary of the surpluses and deficits (both including and excluding the capital cost of investments) generated by the Olympics over 1972–2008, using 2003 estimates for Athens (2004) and 2001 estimates for Beijing (2008). Focusing on deficit and surplus outcomes, Montreal (1976), Munich (1972) and Sydney (2000) would be viewed as disasters by virtue of generating the largest deficits, and Los Angeles (1984) and Seoul (1988) as the greatest successes.

But success or failure needs to be interpreted much more broadly than from just the financial point of view suggested by some of the key calculations in the book. Despite the large deficit when investments are included in the calculation, the Sydney Olympics were considered the ‘best games ever’, presumably because of intangible subjective perceptions: for example, many respondents to post-Games surveys described them as ‘friendly’ and ‘fun’. One danger of giving top priority to generating a surplus (as could be implied from the book) is that this does not take into account the risk to sustainability of the Olympics resulting from excessive commercialisation, as was the perception following the Atlanta Games.

Generally, this chapter is disappointing. Threats to the Olympics, such as the IOC’s monopolistic behaviour causing cities not to bid for the 1984 Games except with a change of hosting arrangements, and causing sports consumers to shift their spending to other expanding international sporting events, are not covered in any detail. The IOC’s decision to limit the growth of the Games to avoid over-commercialisation has been mentioned, but there are a number of missing stories on aspects that challenge the perceived ideals and appeal of the Olympics (for example, the impact of corruption in the bidding process, the use of performance enhancing drugs, etc.). Another threat is terrorism, which raised its head at the Munich Olympics in 1972, and now imposes a large additional cost. The opportunity costs for top athletes to compete at the Olympics are also a threat.

Overall, this book fails fully to satisfy in relation to the wide range of interests it attempts to address. This is partly the result of becoming bogged down in accounting detail and the author’s use of a large number of complex charts. It is also rather difficult to read, and the indexing could have been improved.

Reference


Ray Trewin is Visiting Fellow, Asia Pacific School of Economics and Government, the Australian National University.
NON-AGENDA

With the view of causing an increase to take place in the mass of national wealth, or with a view to increase of the means either of subsistence or enjoyment, without some special reason, the general rule is, that nothing ought to be done or attempted by government. The motto, or watchword of government, on these occasions, ought to be — Be quiet...Whatever measures, therefore, cannot be justified as exceptions to that rule, may be considered as non-agenda on the part of government.

— Jeremy Bentham (c.1801)

Resource Management in New Zealand: Environment Court Decisions

Richard Hawke

The Resource Management Act (1991), the RMA, dramatically altered the environmental management framework in New Zealand. It defines the rights structure for protection of particular uses of environmental resources and provides a decentralised decision-making framework. These rights pose constraints, provide incentives, and create a framework for environmental conflict resolution. The Environment Court (EC) is crucial because its decisions impact upon property rights, which underlie individuals’ incentives and decisions on resource use. The use of an upper level court is a demonstration of society’s views on the importance of environmental considerations and offences (de Prez, 2000). By adjudicating between competing desires, the EC establishes principles for future decision makers. Given the importance of environmental costs, benefits and decisions to individuals and New Zealand generally this paper analyses a subset of EC decisions in an attempt to understand this adjudicator’s role. It is not possible to know what may have happened in the absence of the RMA. However, it is important to understand the EC’s adjudicating role because the EC’s decisions affect the future behaviour of participants in environmental management.

Decision-making Under the RMA

The passing of the RMA established a role for the EC and transformed the legislative framework for land, water and air following the long and arduous Resource Management Law Reform process (Young, 2001). The RMA relies on a
number of key components. First, rather than deal with natural resources in separate ways this Act enabled integrated management. For example, while previous water-related legislation dealt with the river but not riverbanks or land management the RMA ensures the environmental linkages are recognised in the management of the environment. Second, the overarching purpose of the RMA is sustainable management (Part II, section 5), which itself recognises there are competing desires for resources. Third, the Act defines the ‘environment’ and ‘effects’ very broadly. No longer are particular activities regulated (for example, forestry, mining, taking water for irrigation), but their effects are. The rationale for this is that potential third party effects of resource use should be managed rather than particular activities, for example, it should not matter how a particular water quality is arrived at, it is the quality that is important. Fourth, the primary resource management agencies are to be regional and territorial authorities.

The majority of decisions on environmental effects and conflicts are devolved to local councils and most do not involve public notification. Prior to recent amendments and the promulgation of national standards, local government received little central government guidance. Council decision-making has two crucial components. First, regional or territorial plans or policy statements — the framework documents that guide applicants, interested parties and the consent granting authority — set up the overarching framework within which decision-making takes place. Second, resource consents give effect to individual desires to undertake particular actions that affect resource allocation or use.

A recent survey of local authorities noted that in 2003-04 councils processed 54,658 resource consents, of which 4.8 per cent (2628) were publicly notified and 1.2 per cent (651) were appealed to the EC (Ministry for the Environment, 2005). Appeals on local authority plans and policy statements can also be lodged with the EC. Lodging an appeal does not necessarily mean a court hearing because cases may be withdrawn or agreements reached prior to a hearing. Decisions of the EC may be appealed to the High Court on points of law only.

**Methodology: Decision Analysis**

From 1992 to 31 May 2002 the EC and other courts (such as the High Court) made 4986 environmental law related decisions. EC cases are filed in Auckland, Wellington or Christchurch. These three centres keep separate registers that from 2001 to 2002 give an approximate indication of the time taken for a determination. In that period 53 per cent of cases took at least two years for determination with anecdotal, EC and Ministry for the Environment (MfE) information suggesting that the time costs associated with the EC were large (MfE, 2003). Since 2001 there have been a number of measures taken to address the delays, for example, the Court has received additional resources and improved services and in April 2004 the Court issued a Practice Note related to Case Management in an attempt to improve the prompt and efficient disposal of cases. Recent data suggests these measures have reduced delays (see Figure 1). Nonetheless, the ability to lodge an appeal and so delay a decision is recognised as a potential action to take.
The analysis of two important subsets of the decisions provides an insight into the EC’s adjudicating role. First, cases related to section 120 (section of the RMA on resource consents); and second, cases related to the First Schedule Clause 14 (section relating to territorial or regional authority plans). The written decision of each of these cases was analysed and the nature of the case classified.

**Resource consent cases**

In the period 1992 to 31 May 2002, there were 1397 resource consent cases. The cases were classified by type, plaintiff and respondent type, and outcome. The number and type of case decided each year has been variable (see Figure 2). The majority (57 per cent) of consent cases were classified as consent issues. However, year-to-year this percentage varied between 27 and 76. Interestingly the percentage of procedural cases (approximately 15 per cent) has been relatively constant. The major sub-classifications of the 206 procedural cases were: appeal procedures (43), waivers (36), strikeouts (34) and withdrawals (21). The appellants were individuals (50 per cent) or businesses (43 per cent) whereas in almost all cases (1367 out of 1397 cases), the respondent was a council.

**Figure 1: Environment Court Cases Awaiting Determination**

![Figure 1: Environment Court Cases Awaiting Determination](chart1)

Source: Ministry for the Environment (2004); Registrar of the Environment Court (2003); Ministry of Justice (2005).

**Figure 2: Type of Resource Consent Cases Decided**

![Figure 2: Type of Resource Consent Cases Decided](chart2)
The outcome of the resource consent cases was categorised into: win; loss; and, win with court imposed conditions. Only 5-13 per cent of cases (8 per cent average) in each year were actually ‘won’ by the appellant. An additional 14-38 per cent of cases (24 per cent average) were won with court imposed conditions. For procedural and costs cases, which do not really have a winner, the overall number (or percentage) won by appellants is difficult to quantify (see Figure 3).

**Figure 3: Outcome of Resource Consent Cases**

The distribution of possible outcomes was analysed for each of Auckland, Wellington and Christchurch. After excluding the cases that could not be categorised as win, loss, or win with conditions, there were 183 cases (40 per cent of cases) decided in Auckland, 268 (55 per cent of cases) in Wellington; and 190 (44 per cent of cases) in Christchurch. Won cases were 19 per cent in Auckland, 19 per cent in Wellington, but only 12 per cent in Christchurch. Similarly, the percentage of lost cases varied: 44 per cent, 41 per cent and 17 per cent for Auckland, Wellington and Christchurch respectively. In Christchurch the vast majority of cases (71 per cent) were won but with imposed conditions. Of the 1397, 97 per cent were heard by ten presiding judges and the variation in outcome by location is reflected in the variation in outcome by presiding judge (Figure 4).

**Figure 4: Resource Consent Case Outcomes by Presiding Judge**
There are many possible explanations for the variations by location and by presiding judge including, the type of case, the nature of the appellants and respondents — including their access to experts, the situation and the fact that many EC cases were heard by a panel of judges (or a judge and commissioners) and it is not known if the presiding judge had the same view as the majority. Previous research (Su-Wuen, 2000), which analysed the written decisions from 131 EC cases (and then followed this up with survey results from 36 participants), concluded that the number of experts was a strong influence on the outcome of the cases and that environmental groups often found it difficult to access such experts.

Plan (First Schedule Clause 14) cases

There were 696 plan cases in the period 1992 to May 31 2002. The majority of these (70 per cent) were plan changes, but the number of procedural cases was also significant. Unlike the resource consent cases, the distribution of cases across the three main centres appears non-uniform: the total number of cases decided in Auckland, Wellington and Christchurch were 190, 125 and 364 respectively; and the percentage of cases decided in each year in each centre varied.

The distribution of cases across the three main centres is reflected in the nature of the cases. It was possible to categorise 66 per cent of the cases (461 cases) by ‘resource’. Of these 59 per cent (270) were about ‘land’, including zoning and property development etc; 17 per cent (80) were coastal and 6 per cent (27) were about aesthetics. However, in Auckland 50 per cent (59) of the categorised cases (118; 62 per cent) were coastal and 36 per cent (43) were about land; in Christchurch 68 per cent (192) of the categorised cases (281; 77 per cent) were about land; and in Wellington 58 per cent (34) of the categorised cases (59; 47 per cent) were about land. Hence, the nature of the contested resource in the different regions affects the total number of cases. For example, 61 of the 192 ‘land’ cases in Christchurch were cases in which Queenstown Lakes District Council was the respondent. Ericksen et al (2003) note that nowhere in New Zealand are the conflicts between resource development and environment protection so sharply defined and the consequences so far reaching as in the Queenstown Lakes District. These differences, and the lack of policy guidance particularly in the area of assessing landscape quality, led to the considerable number of ‘land’ cases as a result of the district plan development process.

The outcome of plan cases was categorised into case won; case won with court-imposed conditions; win by consent between the parties; or, case lost. As a percentage, only 10 per cent of the cases in each year were actually ‘won’ by the appellant (similar to the percentage for resource consent cases). The majority of cases (49 per cent) were resolved by consent between the parties. Thus in most plan cases the result is a change in the plan and this is completed by either the case being won or the situation being resolved by the parties. This suggests mediation is an important component of the decision-making process.
Discussion

In general there are two broad categories of institutions: adversarial and cooperative. In an adversarial framework the two (or more) opposing parties in a conflict present their cases to a neutral party, who then makes the decision. In this situation the outcome of the conflict tends to be ‘all or nothing’. In contrast, resolution of conflict in a cooperative institution does not involve a decision by a neutral party. Instead, there are various negotiation groups and committees where the opposing parties tend to reach the decision together. Compromises tend to be a typical feature of these decisions (Bostedt and Mattsson, 1996). The recent growth of consensus-type dispute resolution in the area of natural resources has been driven by a desire to reduce costs (including those associated with time delays) and increase the likelihood of achieving solutions that are good for both the economy and the environment (Kahn, 1994; Marcus, Geffen and Sexton, 2002; Prior, 2003). It has also been argued that this type of dispute resolution favours more meaningful participation because it is a more open process and not just less adversarial (Pellow, 1999). While the use of the EC suggests environmental decision-making in New Zealand is more closely aligned to an adversarial process, the RMA contains a number of provisions to encourage the use of mediation. The results of the decision analysis can be interpreted in light of these two categories.

As with the adversarial approach there are problems with the consensus-type approach. First, it is difficult to define what consensus is. Is it unanimity or a compromise? Second, it may not be easy to obtain effective participation. Third, the differences between the parties in many environmental conflicts may be so great and the common ground so small that the process is very slow, and the potential for manipulation, abuse and co-optation is high (Pellow, 1999). Kahn (1994) noted that mediation is most successful when: the number of issues is small, all the parties believe that mediation best serves their interests, and broad issues or public policy are not being negotiated. However, it has been argued that in circumstances of extremely deep conflict the best approach is a collaborative process, not adjudication. An example is the process of ‘sustained dialogue’ where initial effort is concerned with exploring and learning about the scope of the relationship, probing the dynamics of the relationship, and experiencing the relationship together. Only after passing these stages is action considered. Such a process was developed for the post-cold war intra-state conflicts and continues to be employed for multi-level peace processes (Chufrin and Saunders, 1993; Saunders, 1999; Voorhees, 2002). Thus the relationship between the complexity of the case and the applicability of collaboration is not simple.

Decision-making under the RMA

Approximately 14 per cent of the decisions made by the EC since 1992 have been related to local authority plans and policies; however, when examined in detail 70 per cent of these cases involved plan changes and 24 per cent procedural issues. While 10 per cent of cases were actually won by the appellants, another 49 per
cent of cases were resolved by negotiation between the parties. Thus of the 487 plan changes 85 per cent were approved. While Court processes may appear adversarial in nature, the decisions made by the Court suggest that mediation is significant in many cases. The ongoing importance of procedural issues means that court time is being used to determine process issues rather than matters of substance and the ‘rules’ of the process are still being defined and interpreted.

The considerable time taken for determination of cases filed with the EC, and the high percentage of cases determined following court intervention or negotiation, demonstrate the significant role of mediated solutions. RMA cases have some particular characteristics, including: multiple parties (which, due to transaction costs, may make it difficult for the parties to negotiate a settlement, especially once a case becomes public because of the possible ‘loss of face’ that may affect subsequent applications); an open and young process (which admits those with limited understanding and makes it difficult to estimate ‘case quality’); an unequal balance in the gains to the parties following the judgement, especially because of unequal costs (for example, an application costly for the applicant but not for a party seeking to prevent a change in the status quo); and, an unequal balance in the parties’ trial costs. These characteristics suggest a role for mediation. However, a key advantage of an adversarial system is that parties are encouraged to reveal relevant information and a decision may be quickly reached.

Environmental matters are not usually private disputes or the wishes of the individual or group against the wishes of the collective, rather there is a triadic relationship of two individuals or groups and the state where the state must protect either one party or the other. A successful allocation mechanism relies on generating information about the costs and benefits of alternative resource uses and motivating people to use this information. Rights and costs are crucial in any allocation mechanism (Demsetz, 1964).

The EC plays a pivotal role in the framework for establishing principles of entitlement (such as appeals relating to plans) and impingements on those entitlements (such as appeals relating to individual resource consents). It explicitly adjudicates between competing uses of resources, using relevant qualitative and quantitative information to do so. The range of possible mechanisms for allocating resources among individuals and groups extends from markets to government and from private altruism to anarchy. There is no doubt that system ‘failure’ can occur in any system, and that no system can yield a perfect outcome (even if this were known). So there is no a priori reason to assume one system is unambiguously better than another. Rather, a system is best judged by the outcomes it produces and by how it affects transactions (including what incentives it creates).

Under the RMA, the government has devolved resource management and stewardship to a combination of legal, individual and group processes. The framework created by the RMA emphasizes decentralised decision-making. Decentralised decision-making allows local balancing of the preferences and information needs of decision makers (individuals, groups and courts) to achieve outcomes for the economy as a whole that are at least as desirable as centralised control. Although there will be variations across jurisdictions, this structure does
provide a systematic framework within which to manage the environment. The use of mediation seems highly applicable to cases where opponents are seeking specific outcomes that can be aligned with an applicant's broad objectives (for example, particular conditions). However, there are limits to the use of mediation. The use and consequences of mediation have been explored in the environmental decision-making literature, and some of its limitations have been noted. While it might be intuitively unappealing, given the complexity of most environmental decision-making situations, conflict is inevitable; indeed, deliberate and well intentioned collaborative approaches seldom resolve such conflict and given this reality, the discussion should be around the optimal framework for gaining the best outcome (Marcus, Geffen and Sexton 2002). Adjudication may be optimal for complex situations that are not as deep-rooted as international peace processes.

Given the inevitability of environmental conflict there have been attempts to assess the most appropriate form of conflict resolution. Bostedt and Mattsson (1996) in their study of conflict resolution in the Swedish and United States’ forestry industries noted that preference for the adversarial or consensus approach was partially conditioned by the party’s attitudes towards risk and the probability of recurring conflict. In particular, is there enough time to get the gains and losses from a court gamble (that is, do the parties expect to interact into the future), or is the situation a one shot game? The adversarial approach is usually associated with one-off game situations rather than long-term relationships. They also noted that it is not often recognised that the adversarial approach has the positive attribute that all interested groups may be involved even if consensus is seldom reached, whereas the consensus approach may be a very closed system.

Harrison’s (1995) assessment of environmental regulation compliance noted that the enforcement of the collaborative and cooperative Canadian system of environmental regulation in the pulp and paper industry did not yield the same environmental benefits as the United States adversarial approach. Khanna (2001) noted that the most successful voluntary regulation in the United States was when the regulation was backed up with stringent legislative threats; hence, even under a consensus-type approach there is a role for the adversarial system and vice versa.

From an environmental perspective it is unclear that mediation will yield long-run positive benefits: mediation often involves give on both sides and so the possibility of incremental decline in environmental quality may be higher than with a winner take all adversarial system (Kahn, 1994). Litigation has two principal functions: resolving an existing conflict between the parties; and the generation of rules to govern future behaviour. A particular benefit of litigation is that it internalises the costs and benefits of fact-finding to the litigants in a similar way to private property rights in contractual situations (Parker, 1995). In the United States there is considerable research attempting to understand which cases go to trial and the likely outcome of those trials (for example, Waldfogel, 1995)

In summary, while there is international literature on the role of adjudication, mediation and environment conflict resolution, RMA-specific literature is limited. Borrie, Skelton and Memon (2003) discuss the processes for conducting mediation and the legitimacy of mediation under the RMA while Montgomery and Kidd

The decision-making structure underlying the RMA attempts to balance the requirements for decentralised decision-making and planning. Past experience with planning, such as the environmental effects of the planned economies of Eastern Europe, demonstrate the perils of central planning. In addition, issue-specific legislation and an ad hoc approach to what is permitted is also an undesirable mechanism. However, enabling each resource management decision to be a one shot game does not encourage optimal outcomes nor fit well with the complex linkages evident in the environment. Hence, there has to be a balance between decentralised decision-making, obtaining the efficiency gains from being able to accurately measure demand and ensuring the value of options and alternatives are systematic, considered and acted upon. One of the results of the recent (2004/05) RMA review has been a movement towards more central government influence, although not control, to ensure that national interests are considered by local decision makers (Guerin, 2005; MfE, 2004a).

Environmental policy is, therefore, about establishing a rights structure that gives protection to certain uses of resources; while environmental management is about setting a decision-making process that results in optimal solutions regarding the manner, timing and allocation of resources within the economic, political, social and institutional framework provided. The RMA follows Bromley’s (1988) preferred solution of a process that yields standards of performance that have been collectively (politically) determined, and mechanisms for implementation that reward individual initiative, experimentation and efficiency. That is, planning for the social goals while relying on market-like processes to achieve the most efficient implementation of those goals. This was to avoid problems that occur when planners get too involved in operational issues and when the market has too much involvement over the broader policy issues.

The RMA sets down principles of entitlement (for example, affected parties and notification) and the process by which the general principles should be applied. This institutional arrangement that internalises externalities and limits the role of government to allocating initial property rights is aligned with the conclusions of Lai (1994). As such it is quite different from the system of zoning. The development of policies and plans are one mechanism for establishing property rights and the structure against which impingements on property rights are dealt with (Guerin, 2003). Over time the realisation that robust processes are crucial to ensure the proper allocation of property rights, has grown. For example, there has been a proliferation of material on planning and best practice that emphasises property rights (Quality Planning, 2002). Since 1994 (Wells, 1994a, b) it has been recognised that cost redistribution and investment certainty were significant RMA issues. This work also recognised that the RMA was designed to improve the quality of decision-making and that councils were required to monitor the state of the environment (this is related to their role of protecting the property rights of future generations and the intrinsic components of
the environment). Without a robust process, property rights would not be properly allocated (that is, they would be unclear) and poor resource allocation would likely result, due to the lack of defined responsibilities and incentives.

Not only is the RMA an attempt to identify rights it may also be seen as an attempt to leave decisions to the market, rather than use central planning, by properly allocating environmental property rights. ‘Effects’ in the RMA are the consequences of one person’s actions on another person’s property rights. Once they have been settled the applicant may make market judgements for any use consistent with the consent. That is, rather than specifying the appropriateness of an activity the RMA is about the nature of effects.

Conclusion

Resource management is fundamentally about allocating resources between competing uses; as such, many are complex and result in conflict. Given this reality, and the need for an open process, the EC is critical. It plays a pivotal role in the framework for establishing property rights (such as appeals relating to plans) and impingements on those rights (such as appeals relating to individual resource consents). Prior to, and even upon case filing, there is a place for mediation within the RMA. Section 99 allows consenting authorities to arrange meetings to encourage dispute resolution, a process that is common throughout New Zealand. Section 268 allows the EC to engage in mediation between parties awaiting a hearing. In fact, the EC offers a free mediation service. Recent evidence suggests that this service is being used frequently and promoted extensively (MfE, 2001 and 2003). The large time costs associated with the EC have been one of the drivers for encouraging mediation; however, the nature of previous case outcomes is an important constraint on the behaviour of all resource management stakeholders. While a court process is time consuming it does provide for external effects to be considered systematically in a manner that promotes decentralised decision-making by firms, individuals or local organisations. However, it is important that the process is not abused by using the court to create delays.

The trick will be for those involved in resource management decision-making processes to learn and for the processes to become less demanding of resources. In particular, it will be important for participants to understand the nature of their case and accurately predict the likely outcome of a court process because this would increase the probability of avoiding the time and expense of court processes. This should occur as participants become more familiar with court decisions (providing they are disseminated) and the results of procedural cases. A reduction in the time costs will also occur if the proportion of procedural cases drops and the court focuses on substantive issues. Furthermore, as the backlog of cases in the court declines there will be less incentive for parties to take a case to the court in the hope that this will buy them time for mediation because the delay will not occur. In this way the benefits of an adversarial system, mediation and individual decision-making can be gained.
References


This work was undertaken while the author was based at the New Zealand Institute for the Study of Competition and Regulation in 2003. The assistance and guidance of Professor Lewis Evans is gratefully acknowledged. The assistance and efforts of Melissa Storey and Emily Grace in analysing the Court decisions is also gratefully acknowledged. The paper was improved by the helpful comments from the reviewers and the editor.