Contents

Ian Gibbs and Paul Emery, 'Reforming Australia’s Commonwealth Business Programs' 3

Tony Makin, 'A New Rule for Monetary Policy' 17

Rhonda L. Smith and David K. Round, 'A Strategic Behaviour Approach to Evaluating Competitive Conduct' 25

David L. Ryan and Stuart Landon, 'The Political Costs of Tax Reform: A Canadian Perspective' 37

Jonathan Baldry, 'Abolishing Income Tax Deductibility for Work-Related Expenses' 49

Luke Gower, 'What Has Become of the Japanese Model?' 61

Ron Duncan, 'An Optimistic View of World Food Prospects' 73

Geoffrey Lindell, "Judge & Co.: Judicial Law-Making and the Mason Court' 83

George Winterton, 'Australia's Constitutional Convention 1998' 97

REVIEWS

David Preston reviews Reforming New Zealand Welfare: International Perspectives by Michael Jones 111

Chris Bell reviews Australian Local Government: Reform and Renewal edited by Brian Dollery and Neil Marshall 114

Stephen Rimmer reviews Contracting out Government Services: A Review of International Evidence by Graeme Hodge 117

Eric Jones reviews The Rise and Decline of the Asian Century: False Starts on the Path to the Global Millennium by Christopher Lingle 119
Dr Chris Jones has retired as Managing Editor of Agenda, a post he has held since 1995. The editors would like to record their appreciation of his substantial contribution to Agenda.

Editorial Committee

Jeff Bennett, University of New South Wales
James Cox, Independent Pricing and Regulatory Tribunal of NSW
John Freebairn, University of Melbourne
Tony Makin, University of Queensland
Gabriel Moens, University of Queensland
Tony Rutherford, Institute of Public Affairs
Grant Scobie, Centro Internacional de Agricultura Tropical
Judith Sloan, National Institute for Labour Studies
Tony Sorensen, University of New England
Michael Warby, Tasman Institute
Bryce Wilkinson, Capital Economics Limited
Alan Woodfield, University of Canterbury
Reforming Australia's Commonwealth Business Programs

Ian Gibbs and Paul Emery

INDUSTRY policy is once again being debated in Australia. Continuing high unemployment has prompted some commentators to argue that government should not only 'get the basic economic framework right', but must adopt a more interventionist approach to industry development.

This debate has already had an impact on public policy. For example, the Howard Government in 1997 decided to freeze import tariffs on motor vehicles and textiles, clothing and footwear from 2000 to the end of 2004, at levels several times higher than those applicable to most Australian industries.

As well, several recent reports on budgetary and other kinds of government support for industry display a more interventionist stance. Among these are the report of the Information Industries Taskforce (1997) (the Goldsworthy Report) and proposals by the Metal Trades Industry Association of Australia (MTIA/ElU, 1997). More rigorous and wide ranging is the Review of Business Programs (1997) (the Mortimer Review), which contains numerous recommendations designed to improve the overall framework for assessing and applying Commonwealth business programs. But parts of it also suggest a more interventionist role for government.

Apart from their direct cost to taxpayers, business programs influence resource allocation within the economy and alter the incentives facing industry. It is therefore important to ensure that business programs are appropriate and well designed.

Developed while the Howard Government was preparing its response to the recent reports, this article provides an overview of current Commonwealth business programs and of the proposals for change. Rather than offering a comprehensive assessment of existing programs or of the proposed changes, it concentrates on the underlying arguments for business programs. It also discusses the practical constraints on effective intervention and ways to improve program design and delivery. In doing so, it draws heavily on the Industry Commission's (IC) submission to the Mortimer Review (IC, 1997a).

Overview of Commonwealth Business Programs

The IC has estimated that Commonwealth support to industry, in all sectors of the economy, exceeded $8 billion in 1996/97. Of this, tariffs and agricultural marketing

---

Ian Gibbs is an Assistant Commissioner, and Paul Emery a Director, with the Industry Commission. The views expressed are those of the authors and do not necessarily reflect the views of the Industry Commission.
arrangements accounted for close to $5 billion. Government-funded programs and tax concessions assisting business activity accounted for the remainder.

Australian businesses also benefit from other Commonwealth measures which are not easily quantified, such as the anti-dumping system and government procurement policies. State and Territory governments provide a further $2 billion a year of direct support to business, and more than $3 billion a year in payroll tax exemptions.

The manufacturing sector is the major beneficiary of Commonwealth budget outlays on business programs. In 1996/97, it received around 40 per cent of the $1.8 billion spent on these programs (Figure 1). Agriculture accounted for around 30 per cent. However, as a proportion of sectoral value added, agriculture is the major beneficiary. It received budget outlays equivalent to more than 3 per cent of sectoral value added in 1996/97: three times the proportion for manufacturing and 15 times that for the services sector.

Figure 1: Outlays on Commonwealth-funded business programs, by sector, 1989/90-1996/97 ($m, 1995/96 prices)

Source: IC estimates.

Support for research and development (R&D) accounts for more than 35 per cent of total budget outlays on business programs. Export assistance programs are

---

1 The IC's estimate of the cost of these programs is less than the $4 billion cost in 1997/98 of the 70 programs covered by the Mortimer Review. Apart from timing discrepancies, the difference mainly reflects: the IC's exclusion of the $1.5 billion Diesel Fuel Rebate Scheme (which it views as the removal of an inefficient tax on business inputs rather than 'support'); the Mortimer Review's inclusion of various tariff concessions (which are counted in the IC's tariff data); and, partly offsetting these, the IC's wider coverage of programs.
the next most significant category of expenditure. (Figures 2 and 3 provide details on the break-up of program support in agriculture and manufacturing.)

**Figure 2: Commonwealth outlays on agriculture, by type, 1996/97**

- Rural adjustment: 16%
- Forest programs: 8%
- Landcare: 11%
- Other: 16%
- R&D: 49%

Source: IC estimates.

**Figure 3: Commonwealth outlays on manufacturing, by type, 1996/97**

- Export marketing: 11%
- Bounties: 12%
- Enterprise development: 3%
- Other: 14%
- R&D: 35%
- Factor f: 25%

Source: IC estimates.
The Commonwealth also provides a range of tax concessions to assist industry. The major concessions include the 125 per cent (previously 150 per cent) tax con­cession for R&D and income tax averaging arrangements for primary producers. Tax revenue forgone on such concessions totalled $1.6 billion in 1996/97.

Recent Industry Policy Proposals

The Mortimer Review. The Mortimer Review makes both general and specific recommendations to government. These include: adopting a per capita GDP growth target; boosting national savings; aggressively pursuing microeconomic reform; rationalising the current suite of business programs; applying consistent principles to the introduction, operation and evaluation of all programs; and establishing Invest Australia (with funding of $1 billion over five years) to attract additional do­mestic and foreign investment.

In some respects, the Mortimer Review's findings are consistent with the IC's submission to the review. For example, both support the need to apply a market failure rationale to government intervention; to rationalise the current plethora of programs; to apply consistent and transparent principles, including cost/benefit tests, to all business programs; and to fund generic activities, such as innovation, rather than particular sectors or industries.

However, there are some important differences. Notable among these are the review's call for the government to adopt a long-term economic growth target; its advocacy of a $1 billion fund to attract investment; its adoption of a much wider concept of 'market failure', including its extension to policy actions by foreign gov­ernments detrimental to Australian industries; and its acceptance of market failure as a strong rationale for government support for exports.

The MTIA Report. The MTIA Report has a narrower focus on manufacturing and an emphasis on investment attraction. It stresses the need for good macro­economic outcomes and for reform of taxation and industrial relations. But it also calls for the establishment of an investment agency, grants and tax concessions for major investments, and a 200 per cent R&D allowance.

The Goldsworthy Report. The Goldsworthy Report on the information industries also proposes measures to attract new investment. Its recommended National In­formation Industries Strategy includes: an investment attraction agency offering grants or tax holidays for new investments; 200 per cent tax deductions for R&D salaries; and the establishment of an Information Economy Development Fund to undertake feasibility studies on new opportunities. In addition, it argues for tax re­form to ensure that Australia's taxes are 'internationally competitive'; the removal of taxes on business inputs; and higher rates of depreciation for information technol­ogy to better match economic lives.
Rationales for Government Business Programs

In its submission to the Mortimer Review, the IC argued that soundly based and well implemented business programs can boost productivity and economic growth. However, the submission also stressed the need to focus programs in areas where governments can improve on market outcomes for the benefit of the community as a whole.

As all of the recent reports on industry policy emphasise, the most important form of government support for industry is a sound economic policy regime. Business needs a stable and sound macroeconomic environment, effective laws and institutions, and a business environment that rewards efficiency, flexibility and innovation. Recognition of the last of these has driven the microeconomic reform program over the past decade.

Yet clearly there are circumstances in which sound fundamentals and market competition alone do not result in the best outcome for the community and in which government business programs can help. The key issue, therefore, is how to determine when business programs benefit the community as a whole rather than merely the recipient firms and industries.

The following discussion indicates that the circumstances in which intervention by way of business programs is warranted are typically overstated. Moreover, in implementing programs, governments have often failed to recognise the practical difficulties and the costs of intervention. The upshot has been a range of business programs of doubtful benefit to the community as a whole.

The Mortimer Review recognises this and proposes rationalisation of existing programs into five main groups: investment, innovation, exports, business competitiveness, and sustainable resource management.

However, the review's finding that all of these areas have clear rationales for government support is debatable. The MTIA and Goldsworthy Reports' proposals for new or expanded government investment incentives are equally questionable.

**Investment.** Project-specific incentives to attract internationally mobile capital are a feature of recent industry policy proposals. One argument advanced for such subsidies is that they are needed to match similar support available in other countries. The Mortimer Review (p. 88) advances the further argument that foreign investment can bring positive externalities, such as new technology and skills transfer, which may warrant subsidies.

However, it does not automatically follow from the fact that some other countries provide investment subsidies that Australia should do likewise (IC, 1997d). In an economic sense, the provision of assistance by the governments of other countries is no more a 'market failure' than those countries enjoying lower costs because, say, labour is cheaper. Selective investment incentives must be paid for by taxpayers and other Australian industries. Hence, while investment is important to economic growth, such subsidies tend to shuffle investment among activities rather than boost investment overall. So, in the absence of external benefits, the case for incentives to attract investment is not compelling.
Moreover, even if some foreign investment brings external benefits, identifying and quantifying those benefits is difficult, especially in advance. It is hard to know whether a particular project would locate in Australia even without special inducements. Added to this is the difficulty of judging whether, once established, the activity would remain in Australia without continuing incentives. Where competition from other countries is intense, companies can play off one government against another, and so reduce or eliminate any benefit to the country that ultimately hosts the investment.

Indeed, domestic experience with investment incentives should warn against attempts to raise the stakes in this area. State and Territory governments continue to engage in competitive bidding for major investments and ‘events’. The IC (1996b) has found that this, at best, shuffles jobs between regions and, at worst, burdens taxpayers and ratepayers, and reduces the competitiveness of Australian industry and the income of Australians as a whole. The States’ experience in targeting individual firms to promote economic development sounds a similar warning. For example, the Victorian Economic Development Corporation, abolished in 1993, left taxpayers a debt of $110m.

A government that chose to implement investment incentives would therefore need to be extremely alert to the potential pitfalls. It would require stringent assessment and review procedures to determine whether the incentives were delivering national benefits, and a commitment to terminating such measures promptly if net national benefits were not forthcoming.

**Innovation.** R&D is the prime example of a business activity whose social benefits can often be significantly greater than the private benefits. Thus, in its report on R&D, the IC (1995) endorsed the need for government assistance to encourage firms to undertake additional, socially beneficial R&D activity. For the same reason, the Mortimer Review recommends continuing support for R&D.

Of course, determining the degree of externality, and therefore the required level of government support, is problematic. In its R&D inquiry, the IC effectively endorsed the then 150 per cent tax concession for business R&D. The Mortimer Review’s proposals (p. 107) would roughly restore that level of support. In contrast, the Goldsworthy (p. 88) and MTIA (p. 29) Reports call for much greater ‘average’ support; but both reports lack an analytical argument to support their proposals.

On some other R&D issues, there are significant differences between the IC’s report and the Mortimer Review’s approach. For example, the IC (1995:14) considered that CSIRO’s principal role is to undertake ‘public good’ research. It raised some concerns about the potential for the existing 30 per cent external revenue requirement to ‘privatise’ CSIRO’s publicly funded resources. In contrast, the Mortimer Review (p. 124) contains a recommendation for CSIRO to increase its level of external funding to 50 per cent.

The Mortimer Review (pp. 112-18) supports the facilitation of information on management practices and new technology, and their uptake by small- and medium-sized firms. The common argument for such support is that it can be a catalyst for
improving firm practice and culture. For example, enterprise development and business network programs are said to foster a 'best practice' culture.

Yet many firms already face strong commercial incentives to seek out better management practices, technologies and processes. This suggests that the absence of a best-practice culture must reflect more fundamental information problems or other impediments that retard the uptake of better business practice. Moreover, the presence of information deficiencies does not automatically establish a case for government intervention. Often, firms and markets develop ways to improve information: for example, there are many private consultants who can advise on appropriate business management.

It is easy to misinterpret problems in this area. For example, the high per unit cost for small firms of obtaining information is not necessarily a market failure. Small firms face relatively high costs for many inputs. The cost of obtaining information may be evidence of ideal business size rather than of malfunctioning in the marketplace.

Nevertheless, there are some information-related problems which markets may not address well. For example, a firm generating or purchasing information on best practice may have difficulty in preventing other firms benefiting from its efforts through demonstration effects. As with R&D, this may lead to the underprovision of such information. As well, firms may not fully appreciate the benefits of information on best practice, leading them to undervalue it. And firms may not properly use the available information. Where these problems are significant, program-style intervention may be appropriate.

It is essential to be clear about the source of the problem and about whether a business program would be the best way to address it. For example, where inflexible industrial arrangements prevent the adoption of better work practices, the problem should be addressed directly rather than through compensatory, 'best practice' assistance.

Exports. In seeking greater government support for exports, the Goldsworthy Report (pp. 48-9) highlights the growth of Australia's information and communication technology trade deficit. Similarly, the MTIA Report (p. 18) points to the deficit in Australia's trade in manufactured goods generally.

However, the rationale for trade rests on the presumption that the economic strengths of individual countries vary, creating the possibility of specialisation which benefits all. Viewed in these terms, deficits in some sectors and surpluses in others are inevitable and desirable rather than a cause for concern.

Export assistance can increase targeted exports, but raises costs for non-assisted exports and import-competing production. The broad-based nature of Australia's export growth in recent years suggests that factors operating to enhance economy-wide performance are more important to improving export performance. Indeed, in its microeconomic reform stocktake, the Commission (PC, 1996) found that reductions in the costs imposed on exporters by tariffs, and the speeding up of microeconomic reform, had met the preconditions for the removal of export market de-
development assistance set by the last comprehensive review (Committee for Review of Export Market Development Assistance, 1989).

Yet the Mortimer Review (p. 129) advances an externality argument for continuing government support to facilitate exports: ‘The main justification is the need to provide information and overcome private costs of entering or operating in foreign markets which exceed the net costs to society’.

The significance of such externalities is open to debate. Some pioneering firms may help to establish Australia’s reputation in overseas markets and provide market intelligence to other domestic firms. But where is the evidence that such spillovers arise from all export market development activity, or are generally sizeable? Moreover, in some industries, it may be possible to ‘internalise’ spillovers from export market development through an industry levy. That is, the levy, rather than government subsidies, would be used to fund generic export market development.

Business competitiveness. Under this heading, the Mortimer Review covers programs which seek to encourage adjustment or which offer compensation to business for the adverse impacts of other government actions.

Governments have an important role to play in facilitating adjustment to economic change, helping those adversely affected by that change and redistributing income more generally. However, the usefulness of business programs in promoting these objectives is often less clear.

To encourage adjustment, the Mortimer Review (p. 167) proposes short-term assistance for capital investment in plant and equipment, and one-off grants to firms facing unforeseen adjustment pressures.

The IC has readily acknowledged that there will be circumstances where specific adjustment measures or programs are warranted on efficiency or equity grounds. Thus, it has typically recommended phased rather than abrupt tariff reductions.

But, as the IC (1996c) found in relation to rural adjustment, it would be crucial that any specific measures were set up in a way that encouraged rather than delayed adjustment. For example, there is the risk that ‘transitional’ programs will become permanent. The retention of ‘temporary’ quotas on car imports from 1974 to 1988 illustrates this point. More generally, the provision of adjustment assistance to particular industries may encourage other industries to seek similar assistance as part of any reform proposals. Subsequent ‘horse trading’ over the level of support may act to delay necessary microeconomic reform.

For these sorts of reasons, the Commission (PC, 1996) has argued that the social welfare system and generally available retraining programs are usually better than specific measures as a way of assisting producers or employees adversely affected by change.

Several existing ‘business competitiveness’ programs aim to ‘compensate’ industries for the adverse effects of other government policies. For example:
• The Factor f scheme exists partially to offset the effects on the domestic pharmaceutical industry of drug price suppression under the Pharmaceutical Benefits Scheme.

• The Diesel Fuel Rebate Scheme offsets an inefficient impost on petroleum products used by farmers and miners. The excise on diesel fuel would otherwise distort production choices.

• Various tariff concession and import duty rebate schemes reduce the impact of tariffs on the cost of imported inputs.

As the Mortimer Review (p. 167) recognises, if 'problem' policies cannot be altered, compensating support may improve overall resource allocation.

However, designing compensatory policies is informationally demanding and this can lead to targeting difficulties. In this context, the IC’s recent report on the pharmaceutical industry (IC, 1996a) concluded that the Factor f scheme had overcompensated some firms and under-compensated others.

More broadly, compensating for the adverse effects of other policies can reduce the pressure to reform those policies. As the Mortimer Review (pp. 168-71) acknowledges, compensating assistance is very much a second-best option. General taxation reform could remove the need for the Diesel Fuel Rebate Scheme. And further tariff reform would render tariff concessions on imports redundant.

Sustainable resource management. The IC’s submission to the Mortimer Review did not address resource management programs in any detail. Nevertheless, it is clear that significant market failures can occur in this area. For example:

• Poor land management can lead to erosion and water run-off which reduce the quality of adjacent water supplies to downstream users.

• Community demands for conservation of biodiversity and other environmental attributes are often not reflected in decisions about resource use.

• Property rights to certain natural resources are inadequately specified, or there may be constraints on trade in these rights.

• There may be insufficient information on, or ignorance of, the effects of natural resource use.

Addressing these problems will often require better and more flexible property rights and regulatory regimes rather than business programs. For example, the IC’s recent draft report on land management (IC, 1997b) proposes recasting the regulatory regime to ensure that resource owners and managers take into account the environmental impacts of their decisions. It also proposes measures to create or im-
prove the markets for key natural resources and to encourage conservation on private land. In a similar vein, the Mortimer Review (p. 182) proposes the introduction of tradable quotas for all Commonwealth fisheries. However, the review also endorses business programs aimed at gathering and disseminating information on the consequences of poor land management and at encouraging energy efficiency. In the presence of such information market failures as those noted above, these programs may be useful. But the same caveats apply. For example, governments should be careful not to crowd out private markets: note the increased use of private consulting services which accompanied the partial withdrawal of State government extension services for agriculture (IC, 1996c).

Some Practical Considerations

Even when there is a good case for government business programs, there are invariably costs and constraints which can reduce the benefits or lead to negative outcomes.

Costs of intervention. Raising taxes to pay for business programs has efficiency costs. For instance, the efficiency cost of income tax could amount to more than 20 cents for each additional dollar raised (PC, 1996). And as well as the direct budgetary costs, there are the administrative costs for government and business alike.

Information problems. Proper assessment of proposed programs must take account of the nature and magnitude of the underlying problem and the impacts of alternative policies to deal with the problem (including doing nothing). Obtaining the information to make such assessments can be costly or even impossible. Moreover, information on the benefits of program support will typically be more comprehensive than that on the costs to the wider community. This increases the risk of programs being based on what is good for particular firms or industries rather than for the community as a whole.

Promotion of rent seeking. Government support for industry, even for good reasons, inevitably encourages lobbying for financial favours. From a national perspective, this is unproductive.

Incentives for program administrators. Administrators of business programs can face incentives that conflict with the goal of obtaining the best result for the community. For example, the performance of program administrators is likely to be judged on how well they have served the users of a program. This can lead to 'regulatory capture'. In addition, to avoid the appearance of failure, governments can be tempted to provide extra assistance to 'at risk' activities that have previously benefited from public support.

International obligations. Australia is a signatory to international trade agreements that limit the forms of government support to business. For example, World Trade
Organisation provisions prohibit any export subsidy outside agriculture. As the recent challenge to the import credits scheme for the textile industry highlights (IC, 1997c), Australian governments cannot ignore those commitments with impunity.

These problems do not condemn all business programs to failure. For example, as already noted, the IC (1995) found that tax concessions for R&D have benefited Australia, despite the difficulty of precisely estimating the extent of market failure in this area.

But the problems do highlight the need for proper assessment of the potential benefits and costs of intervention. Thus, the Mortimer Review (p. 72) recommends a threshold test for all business programs. This would require a formal assessment of whether a program would provide a net economic benefit, would address a market imperfection and would not breach Australia's international obligations. It would follow an assessment of whether the program was preferable to other approaches, such as changes to regulations or taxation measures. In many respects, this recommendation is in line with similar proposals in the IC's submission to the review.

However, as is apparent from the preceding discussion, the Mortimer Review's views on the circumstances in which net economic benefits are likely to arise go well beyond conventional views and those of the IC.

**Good Program Design**

Where business programs can improve community welfare, policy should focus on obtaining the maximum benefit for the funding involved. In its submission to the Mortimer Review, the IC (1997a:29-37) highlighted the key aspects of good program design.

**Effective targeting.** Programs should focus support on new activity rather than activity that would have occurred anyway. They should also target market failures generally rather than particular firms or industries. Thus, generally available support for activities such as R&D is preferable to firm-specific subsidies. And programs should be open to all firms rather than limited to (say) firms of a particular size.

**Transparent support.** Transparency facilitates monitoring and reporting on the outcome of programs. It requires clear eligibility criteria that limit administrative discretion in the allocation of funding to firms. An important finding emerging from the IC's inquiry (1996b) into State government assistance to industry is that secrecy in process and outcomes can lead to public suspicion that 'deals are being done', even when this is not the case.

**Maintaining market incentives.** Programs should retain the incentive for firms to operate efficiently and should not alter relative prices any more than is necessary to address the underlying problem.
Avoiding unnecessary compliance costs. Keeping programs simple, seeking industry input to program design, and relating the stringency of compliance requirements to the amount of support provided to individual firms can help to contain compliance costs. However, there will often be trade-offs between reducing compliance costs and meeting other objectives, such as targeting support. Hence, compliance costs must be assessed in conjunction with other benefits and costs.

Avoiding program duplication. Coordinating related programs and addressing duplication will reduce compliance costs. It will also diminish confusion for firms about what support is available. To help avoid duplication, program support at one level of government should take account of similar support at other levels.

Effective monitoring and public reporting. All business programs should be monitored regularly to assess their performance against underlying objectives. Public reporting of outlays and outcomes is important to allay suspicion that business programs are simply ‘business welfare’.

Periodic review by independent bodies. There is a need for periodic and independent review to consider whether programs remain relevant in the light of changes in the wider market and policy environment. This should be formalised through the imposition of sunset clauses in business programs.

Separation of policy from delivery. Responsibility for policy formulation should be separate from responsibility for program delivery.

Contestability in service provision. Where a program involves a service to industry — for example, extension advice — delivery should be open to the private sector and/or other government agencies.

User contributions to the cost of services. Market failures are rarely, if ever, significant enough to justify full public funding of business service programs. User contributions also help to avoid unnecessary use of programs and maintain pressure on providers to deliver services efficiently.

Of the recent industry policy reports, only the Mortimer Review looks at these issues in any detail. Its recommendations are, for the most part, identical to or consistent with the checklist in the IC’s submission. The main difference is that the Mortimer Review (p. 82) recommends annual performance reports and periodic evaluations against agreed objectives and performance indicators, with broader reviews only in exceptional circumstances. This would be less rigorous than the IC’s proposal for periodic independent reviews triggered by legislated sunset provisions.

Concluding Remarks

The over-riding objective of government business programs must be to improve community welfare. The focus of business programs cannot be simply on
'promoting industry' or particular sectors. The development and growth of particular industries or sectors is an outcome of industry policy, but should not be its objective.

The Mortimer Review explicitly recognises that business programs must address market imperfections and provide net economic benefits. Its proposals for assessing business programs are also consistent with most of the IC's; the other recent reports do not properly consider these issues.

Yet despite the Mortimer Review's greater rigour, there is a significant tension between its criticisms of the ad hoc nature and poorly specified objectives of many existing programs and its finding that the current level of expenditure on business programs is appropriate. More importantly, the review's proposals carry a significant risk because they are based on a very broad interpretation of market failure and the circumstances in which governments should intervene.

The adoption of investment incentives would represent a significant shift towards a more interventionist industry policy. The practical difficulties of identifying and quantifying any external benefits from new investments supported by incentives raise considerable doubts about the merits of such an approach. Certainly experience at the State level with investment attraction is far from encouraging.

More generally, with any proposals for business programs, it is imperative to ensure that the extent of market failures, and the capacity of governments to address them effectively, are not overstated. But where programs are implemented, the importance of good program design, including periodic independent reviews triggered by legislated sunset provisions, cannot be overstated.

References


(1996b), State, Territory and Local Government Assistance to Industry, AGPS, Canberra (Report No. 55).

(1996c), Submission to the Mid-Term Review of the Rural Adjustment Scheme, AGPS, Canberra.

(1997a), Submission to the Review of Business Programs, AGPS, Canberra.


A New Rule for Monetary Policy

Tony Makin

Disagreements about the conduct of monetary policy have traditionally centred on the relative importance of its goals and the best means of achieving them. Today, central bankers everywhere treat inflation control as the main objective of monetary policy, but retain considerable discretionary powers while operating with varying degrees of independence from governments. But do central banks know more about an economy's fundamentals than the financial markets, which act, directly or indirectly, on behalf of all agents in the economy?

Under present arrangements, the interest rate structure reflects the markets' best estimate of where the central bank will set cash rates. Interest rates vary considerably in response to the markets' continuous second guessing of discretionary changes. Discretionary interest-rate setting by the central bank thus contributes to longer-term interest rates being persistently higher than necessary, with adverse consequences for investment and long-term economic growth. If monetary policy were to become rule-based, however, market uncertainty about the future direction of the cash rate would be eliminated, and long-term interest rates would be lower than otherwise.

In this article, a new rule for monetary policy is outlined, involving the automatic setting of interest rates by the central bank. The proposed rule is based on a re-interpretation of the so-called expectations hypothesis of the term structure of interest rates. Its rationale is to minimise the level of discretion-induced economic uncertainty by effectively allowing financial markets to set all interest rates. Over time, the cash rate automatically set by the authorities in accordance with the rule would reflect, not the central bank's own discretionary view of its appropriate level, but the collective view of the financial markets based on their information set of the economy's fundamentals.

Current Operating Procedures for Monetary Policy

The modus operandi of monetary policy differs markedly across countries, both in terms of the intermediate variables targeted and with respect to the frequency and magnitude of central bank interventions to influence monetary conditions. Instead of attempting to control the money supply, as prescribed by monetarism, central banks now generally operate on one or both of the key monetary price variables: exchange rates and interest rates.

Tony Makin is Associate Professor in Economics at The University of Queensland.
Since Australia's exchange rate was floated, the Reserve Bank alters the stance of monetary policy by focusing on interest rates through the so-called cash rate. The cash rate is set for the purpose of influencing other interest rates, particularly longer-term rates. Thus, the linkage between short- and long-term interest rates plays a major role in the transmission of changes in the stance of monetary policy to the real economy.

Changes in interest rates are thought to have an impact on the spending patterns of households and firms through liquidity, wealth, exchange rate and expectations channels. For instance, higher interest rates can curb firms' investment expenditure (as witnessed in Australia during the last great policy induced recession of the early 1990s) because the cost of capital rises relative to its productivity. Higher interest rates are also likely to reduce equity values, as investors switch from shares to bonds, thus lowering the market value of firms' capital relative to its replacement cost. According to standard theory, this further weakens the incentive for firms to invest.

Moreover, to the extent that lower equity values depress aggregate household wealth, consumption spending may also be reduced. For monetary policy to operate effectively along these lines, it is important that interest rates at the short end of the spectrum influence rates at the long end, since investment spending in particular is likely to be more sensitive to longer-term interest rate movements, given business planning horizons. The interrelationship between monetary policy, interest rates and the economy is shown schematically in Figure 1.

Figure 1: Monetary policy, interest rates and the economy

---

1 Stevens (1996) elaborates on the Reserve Bank's view of the transmission mechanism of monetary policy.
The behaviour of the targeted cash market rate, and of indicative short-term and long-term interest rates over recent years, is depicted in Figure 2. What is evident is the close relationship between the cash rate and short-term (such as the 90-day bank bill) rates. There is also evidence of substantial co-movement between short-term and long-term rates, though long-term rates have been persistently higher.

International arbitrage ties long-term rates to foreign long-term rates, with any difference between domestic and foreign interest rates likely to arise from the markets' expectation of future exchange rate changes according to the standard uncovered interest parity condition. For instance, the gaps between similar long term rates in Australia and overseas rates can reflect market expectations of future exchange rate movements, as well as liquidity and risk premia which vary over time.

**Figure 2: Short-term and long-term interest rates in Australia, November 1993-April 1997**

![Graph showing short-term and long-term interest rates](source: Reserve Bank of Australia Bulletin (various issues, 1997).

**Monetary Policy and the Yield Curve**

At any given time, the so-called yield curve describes interest rates on securities which are ostensibly the same except for their terms to maturity. For instance, an ascending, or positive, yield curve shows that, for any given time interval, short-term rates on, say, five-week Treasury Notes are lower than on longer-term rates on, say, 26-week Treasury Notes. Descending or inverse yield curves convey the opposite (Figure 3). They can also be flat or hump-backed and the slope and position of the yield curve can vary from period to period.
The yield curve itself can be used as an indicator of expected macroeconomic conditions, particularly when comparing short-run interest rates and rates at the long end of the spectrum. For instance, long-term rates will reflect expectations of the inflationary effect of current monetary policy. There will therefore be some offset to the linkage between long-term rates and officially set short-term rates to the extent that a deflationary monetary policy which raises short-term rates also lowers the markets' expectation of future inflation. The longer the maturity structure, the more important this offset is likely to be.

**Figure 3: Yield curves for Treasury Notes, March 1997 and April 1997**

Various theories have been expounded over the years to explain why short-term and long-term market interest rates differ. The three main explanations of the term structure of interest rates that have been advanced are: the expectations theory; the risk (or liquidity preference) approach; and the segmented markets theory. Of these, the expectations theory is the most widely accepted. It presumes that investors regard default-free securities, like Treasury Notes, which are issued by the same entity but which differ with respect to their terms to maturity, as perfect substitutes. Accordingly, the interest rate paid on the longer-term security — say, a hypothetical 60 day Treasury Note — equals the average of the current rate on a hypothetical 30 day Treasury Note and the markets' expectations of the rate to be paid on 30 day Treasury Notes in a month's time.

In formula terms, the expectations hypothesis suggests that:

---

2 Cox et al. (1985), Culbertson (1957), Fisher (1896), Hicks (1939), Malkiel (1986), Melino (1988) and Roll (1970) provide expositions of theories of the term structure.
A New Rule for Monetary Policy

60 day rate = 0.5 (30 day rate + next month's expected 30 day rate)

Or, more generally, \( i_{2t} = \frac{i_t + i_{t+1}}{2} \)

where \( i_{2t} \) is the longer term rate, \( i_t \) is the current, or spot, short term rate and \( i_{t+1} \) is the short term rate expected in the next period.

For example: if the spot interest rate on a 30-day Treasury Note is 5 per cent and the market expects that the interest rate on a 30-day Treasury Note next month to be 7 per cent, then, according to this theory, the spot interest rate on a 60-day Note would be 6 per cent. It follows that an ascending yield curve depicting the short rate at 5 per cent and the long rate at 6 per cent contains information about the money market's expectation of the future short-term rate. In other words, this approach suggests that it is market opinion about future short-term rates which drives longer-term rates. The accepted corollary for monetary policy is that long-term interest rates can be changed by fixing the official short-term rate for sustained periods, since money and bond markets will base their expectations of future short-term rates on the current official rate.

The risk (or liquidity preference) theory of the term structure is effectively an augmented version of the expectations hypothesis. It suggests that investors consider short-term securities to be more attractive than longer-term bonds due to uncertainty about future short-term rates and that, accordingly, a premium is required on longer-term securities to make them as attractive to investors as short-term securities. The longer the maturity, it is argued, the higher this premium will be.

The segmented markets theory of the term structure is diametrically opposed to the expectations theory because it presumes that securities of different maturities issued by the same entity are completely non-substitutable. Accordingly, the slope of the yield curve results from the interaction of the demand and supply of securities in distinct, or separate, markets. In other words, no linkage exists between officially controlled short-term interest rates and longer-term rates.

The corollary for monetary policy here is that targeting short-term rates is largely ineffective in influencing long-rate-sensitive investment spending and macroeconomic activity. Instead, the theory implies that central banks should intervene in longer-term securities markets in order directly to influence long-term interest rates. However, this approach has few supporters and little empirical support.

Using Market Expectations to Set Interest Rates

Previous studies of the relationship between monetary policy and the yield curve have either focused on the yield curve itself as an indicator of future inflation and macroeconomic activity, or have empirically tested the validity of the expectations hypothesis.

---


4 An exception is Taylor (1992).
to establish whether discretionary monetary policy can in practice influence long-
term interest rates.

It is proposed here that the central bank regularly set the cash rate according to a
rule which is based on the expectations hypothesis and uses information contained in
the yield curve itself. The procedure whereby the cash rate is inferred from the yield
curve can be illustrated by reference to the hypothetical data in Table 1.

<table>
<thead>
<tr>
<th>Table 1: The rule illustrated</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Short-term rate (i*&lt;sub&gt;t&lt;/sub&gt;) (e.g. 30-day security)</td>
</tr>
<tr>
<td>First period</td>
</tr>
<tr>
<td>Second period</td>
</tr>
<tr>
<td>Third period</td>
</tr>
</tbody>
</table>

The cash rate which the central bank sets by intervening in the money market is
fixed at the same level as the short-term market rate shown in the first column. The
second column records a relatively longer rate which reflects the market’s expectation
of the future cash and short-term rates. For instance, if the cash rate is 5 per cent and
the longer rate is observed to be 6 per cent, it follows from the expectations hypothe-
sis of the term structure outlined above that the markets expect the future cash rate will
be 7 per cent. That is, from the earlier equation summarising the expectations hy-
pothesis, the expected future short-term rate is simply related to the current short-term
rate and the longer rate in the following way:

\[
i^*_t = (2 \times i^*_t) - i^*_t
\]

In this particular example, the cash rate can be automatically reset by the central
bank every 30 days. Yet it would also be possible for the central bank automatically to
peg the cash rate more frequently, for instance as regularly as every day, using a mov-
ing average of implicit expected rates, calculated using a designated security, such as
Treasury Notes. The money and bond markets could then express their collective
view about the appropriate stance of monetary policy by focusing on trade in that par-
ticular security, thereby determining the rate to be calibrated daily by officials. In ef-
fect, all domestic interest rates would then be market-determined.

The Exchange Rate

Questions may arise as to where the exchange rate fits into this proposal, since, in ad-
dition to interest rates, the exchange rate is taken into account by (for example) the
Canadian and New Zealand central banks when assessing monetary conditions.

Kaijakis and Moschos (1993), Lee (1995) and Lowe (1992) focus on the usefulness of the yield curve as
an indicator of economic conditions.
However, under the prevailing floating exchange rate system, the central bank, by
definition, should allow the exchange rate to find its own equilibrium level according
to the play of demand and supply in the foreign-exchange market itself. Hence there
is no explicit need to factor the exchange rate into the proposed interest-rate rule
(though rule-based changes in cash rates relative to foreign interest rates may still influ­
ence exchange-rate movements and hence real activity, as suggested earlier in Figure
1).

Inflation

Where would such a rule leave the low-inflation objective of monetary policy? It is a
fundamental axiom of monetary economics that central banks which ultimately have
monopoly control over the issue of the domestic currency cannot simultaneously set
interest rates and the money supply at pre-determined levels. Targeting the money
supply for given money demand implies that the interest rate becomes endogenous,
whereas targeting the interest rate for the same money demand implies that the money
supply becomes endogenous. Hence the money supply must always adjust in re­
sponse to the central bank’s interest targeting activity.

Accordingly, if the central bank strictly adhered to the proposed interest-rate rule,
then the money markets would also indirectly govern the rate of domestic money
growth and so ultimately decide the rate of inflation over the longer term. This may
suggest that the interest rate rule could lead to higher inflation. But such an outcome
seems unlikely in view of the strong aversion to inflation that money and bond mar­
kets have demonstrated in the past. Financial markets now react quickly to news of
higher inflation by immediately pushing up interest rates. Under the proposed rule,
the financial sector itself would thereby automatically induce higher cash rates, and
hence automatically check domestic money supply growth, which is the ultimate de­
terminant of inflation.

Conclusion

The current practice of discretionary interest-rate setting by central banks is clearly a
major cause of uncertainty in financial markets about the future course of short-term
interest rates. The Reserve Bank of Australia’s discretion in periodically changing
cash rates is likely to explain part of the persistent, though time-varying, premium on
Australia’s long-term interest rates.

Of the several rules that have been proposed to replace central banks’ discretion
over interest rates, probably the best known is the money growth rule derived from
the monetarist doctrine that growth of the money supply is the key determinant of
inflation. Yet this rule, though aimed at nullifying the potentially debilitating effects of
central bank discretion, presumed superior information on the part of the authorities
to the extent that the stipulated money growth rule had to be premised on the central
bank’s forecasts of future economy-wide activity.

The alternative rule proposed here recognises that policy-makers are likely to
know less about the behaviour of the economy at large than is known collectively by
the financial sector, operating on behalf of resident households and firms transacting in domestic and international markets. Since interest rates of varying yields are proximately determined by the money and bond markets, it follows that the information held by these markets about economic fundamentals should be used as a basis for setting official interest rates. Financial markets would then effectively determine interest rates, consistent with the way they set the value of the currency under a floating exchange rate regime.

References


The author gratefully acknowledges the comments of two anonymous referees and the editor.
A Strategic Behaviour Approach to Evaluating Competitive Conduct

Rhonda L. Smith and David K. Round

SECTION 46 of the Trade Practices Act prohibits Australian firms which have a substantial degree of market power from taking advantage of that power for the purpose of hindering competition in a market, preventing entry or damaging a rival. One of the earliest and most influential trade practices decisions in Australia is the so-called QCMA case of 1976, concerning a proposed merger of flour millers. Here, the Trade Practices Tribunal (now the Australian Competition Tribunal, ACT) determined a methodology for analysing competition issues which was subsequently adopted by the Trade Practices Commission (now the Australian Competition and Consumer Commission, ACCC) as well as by the Federal and High Courts.

The QCMA approach is to define the relevant market and then to assess whether the conduct at issue could result in an increase in market power and consequently a substantial lessening of competition. The approach is structural, based largely on the measurement of non-behavioural factors. The Tribunal concentrated on what can be called 'group-structural' features of the market because of its concern that, after a merger, the market could exhibit collusive tendencies as a result of increased concentration. The emphasis was on potential as well as actual substitute products and on the height of entry barriers. Little account was taken of the dynamic, rivalrous interaction between current sellers or between these incumbents and potential entrants. Much concern was expressed about the price effects of the merger; and price issues have dominated most of the subsequent applications of this methodology.

The QCMA methodology did refer to the broad dimensions of competition, including non-price elements; it was noted that 'there should be independent rivalry in all dimensions of the price-product-service packages offered to consumers and customers'. Nevertheless, dynamic non-price rivalry, which individual firms undertake...

Rhonda Smith is a Commissioner of the Australian Competition and Consumer Commission. David Round is Associate Professor of Economics at the University of Adelaide and an Associate Commissioner of the Australian Competition and Consumer Commission. The opinions expressed in this article are those of the authors alone.

1 Re Queensland Co-operative Milling Association Ltd., Defiance Holdings Ltd. (Proposed Mergers with Barnes Milling Ltd.) (1976) ATPR 40-012.
2 Ibid at 17,246.
Rhonda L. Smith and David K. Round

in order to derive a sustainable competitive advantage over their rivals, received little attention both in this decision and in subsequent cases.

In this article, it is argued that analysis of the competitive impact of a firm's behaviour should be primarily considered within a non-structural framework which recognises that conduct can differ markedly between firms in the same structural environment, or can be similar between firms operating in quite different structural market conditions, depending on the firms' goals, strengths, weaknesses and strategic directions.

Without doubt, the QCMA model provided a relatively simple but effective means of assessing the implications of conduct in relation to the Trade Practices Act at a time when neither the Federal Court nor the legal profession had much experience in dealing with a statute which is a mix of economics and law. In addition, in 1976 markets were less concentrated than they are today, and business initiatives and the environment were considerably less dynamic. However, time and theory have overtaken the QCMA approach, which now has less relevance to any assessment of the misuse of market power or of competition generally.

Strategic Conduct and Market Structure

Since 1976, markets in Australia have become more oligopolistic, particularly as a result of the mergers and acquisitions of the 1980s and the general rationalisation of the 1990s. The distinguishing feature of oligopolistic markets is not so much the number of market participants, or the extent of market concentration, or the precise height of barriers to entry, but rather a recognition by firms of their mutual interdependence and its associated uncertainty. However, the individual patterns of behaviour which follow from this recognition can differ greatly. Of two markets, one with only two participants and the other with considerably more but with otherwise similar structural conditions, the latter could display little competition while the former could be highly competitive. Indeed, within two markets with identical structural features, including firm numbers, behaviour could differ markedly. Structural analysis is insufficient. Something more is needed: an analysis of strategic behaviour.

Attention should be focused on rivalry between individual firms rather than on a structural concept of competition. An analysis of strategic behaviour by individual firms will assist in distinguishing competitive from anti-competitive conduct. In oligopolistic markets, strategic behaviour can alter market structure, for example by raising or creating entry barriers. Further, while one action or strategy on its own may appear to lessen (or not to lessen) competition, a firm's whole set of strategies may lead to the opposite outcome. This may be especially important in the context of s.46 of the Trade Practices Act, which covers misuse of market power, where purpose must be proved, even if by inference from conduct.

What Is Strategic Behaviour?

'Strategic behaviour' refers to actions which a firm takes to improve its competitive position relative to actual and potential rivals, in order to gain a permanent commer-
A Strategic Behaviour Approach to Evaluating Competitive Conduct

A firm's strategic behaviour can influence its long-run profits. Carlton and Perloff (1994:382) refer to actions 'to influence the market environment and so increase profits'; while Martin (1993:46) refers to 'investment of resources for the purpose of limiting rivals' choices'. Strategic behaviour thus refers to conduct which is not economically inevitable, but which is the outcome of a conscious attempt to shape the firm's market environment to its own lasting advantage and to the competitive disadvantage of rivals.

There are two categories of strategic behaviour. 'Non-cooperative behaviour' occurs when a firm tries to improve its position relative to its rivals by seeking to prevent them from entering a market, to drive them out of business or to reduce their profits. 'Cooperative behaviour' occurs when firms in a market seek to coordinate their actions and therefore limit their competitive responses (this does not necessarily imply explicit agreement). Here, only non-cooperative strategic behaviour is considered.

It is primarily under oligopolistic market conditions that a firm has an incentive to alter its relative position through strategic behaviour. The firm recognises its interdependence and the need to take into account other firms' reactions when making its own decisions; but it also recognises that it is free to make decisions to alter its commercial environment. These strategies are revealed over time through investment and through tactical moves and countermoves. Strategic behaviour can be manifested in:

- entry deterrence, for example through output expansion, the deliberate creation of excess capacity, pricing just below the level which would encourage entry, acquisition of essential inputs, predatory pricing, raising rivals’ costs, and making credible threats as to how the firm will react to entry;
- advertising and brand proliferation;
- R&D and technology choice;
- tying consumers in various ways where switching costs are significant; and
- various long-term contracting devices.

To engage in successful non-cooperative strategic behaviour, a firm must have some market power or advantage; it must be able to act before its rivals; and it must demonstrate credibly that it will follow its strategy regardless of the actions of its rivals (that is, it should be able to deter potential rivals by changing their beliefs about how aggressively it will behave in future).

Such conduct may not cause long-term damage to the competitive process if continual opportunities exist for all firms to initiate new bouts of strategic behaviour, and if they all have equal opportunity to initiate such actions. There is nothing wrong with

---

3 Textbooks, especially older ones, often suggest that these types of conduct are practised only by dominant firms; but modern theory accepts that such conduct can be undertaken by firms with relatively modest degrees of market power.
a firm seeking to get ahead of its rivals by developing a sustainable commercial superiority over them by, for example, developing better production techniques or introducing new and better products. This has been recognised by the High Court in Australia in the *Queensland Wire* case, where it was stated by Mason CJ and Wilson J that

Competition by its very nature is deliberate and ruthless. Competitors jockey for sales, the more effective competitors injuring the less effective by taking sales away. Competitors almost always try to 'injure' each other in this way. This competition has never been a tort ... and these injuries are the inevitable consequence of the competition s.46 is designed to foster.

**The Structuralist vs the Strategic Behaviour Approach**

The traditional structuralist approach and the strategic behaviour approach differ in several important ways. First, the structuralist approach is static, concentrating on two successive states of affairs but ignoring the process of the move from one to the other. It thus ignores the intertemporal issue usually associated with strategic behaviour. The strategic behaviour approach, in contrast, is dynamic and concentrates not only on conduct but also on responses to that conduct, and on the response to those responses, and so on, thus treating the firm as continually adapting to its environment.

Second, the structuralist approach tends to play down informational and demand asymmetries between firms, and the consequent differential expectations and the strategic responses which their existence may attract.

Third, while the adoption of new technology is accounted for in the structuralist model by downward movements of cost curves, in the strategic behaviour model it may also affect the conduct of one or more of the firms in the market.

Fourth, the strategic behaviour model tends to take a neutral view of firm size and market shares. While the structuralist approach may simply use market share as an indicator of the need to examine other structural factors, often size as such is taken to indicate the potential for collusion or misuse of market power; a line of reasoning that may be quite misleading in relation to oligopolistic markets.

Fifth, the structuralist model treats entry barriers as the most significant factor in assessing whether particular conduct is likely substantially to lessen competition in a market. In *QCMA*, it was said that

Of all these elements of market structure, no doubt the most important is ... the condition of entry. For it is the ease with which firms can enter which establishes the possibilities of market concentration over time; and it is the

---


5 As suggested in the *Merger Guidelines* prepared by the ACCC (1996).
threat of entry of a new firm or a new plant into a market which operates as the ultimate regulator of competitive conduct.

The strategic behaviour model, meanwhile, places less significance on current barriers to entry (but by no means gives them an insignificant role) because, even in the absence of entry barriers, a firm may still seek to achieve a strategic advantage over its rivals by exclusionary or other behaviour which will create barriers to entry or expansion in the future. The absence of barriers to entry is therefore insufficient for establishing that conduct will not result in a substantial lessening of competition or in misuse of market power in the long run.

**Why a Framework Based on Strategic Behaviour?**

Whether strategic behaviour is anti-competitive in its effects depends on its nature and purpose. It may be socially acceptable under some market conditions but not under others. But the static, structural, group-centred approach based on QCMA is not really appropriate for fully analysing the anti-competitive effects of such conduct. Under that approach, if the structural features of the market do not appear likely to facilitate the exercise of market power, the conduct is generally not analysed further. However, the conduct in question may ultimately alter the market structure. Recent antitrust literature (see, for example, Baker, 1997) indicates that even firms with quite a small market share may exercise unilateral market power through strategic behaviour.

Not only has market structure become more oligopolistic, but business structures have become more complex. Most firms now operate in multiple markets that are linked in some way (for example, through common manufacturing or distribution processes, inputs or information) which provides synergies for the business as a whole. Thus, to understand conduct in any one market it may not be sufficient to look simply at the structure of that market in isolation.

Kay (1993:227-8), for example, suggests that firms should segment their markets (price discriminate) in order to appropriate added value most effectively. The firm should operate in as many markets as possible, given its competitive advantage. Pricing and marketing strategies should be designed to suit the characteristics of each market: for example, a firm might lower its prices in order to increase its market share so that it can earlier enjoy the size needed to undertake technological change. Firms can also engage in what might be called ‘strategic packages’, that is, a coordinated set of strategies jointly determined to further the multiple goals of the firm in one or all of its commercial environments. An example of this multiple-pronged strategic conduct is provided by the well-known du Pont case, where in order to derive a lasting commercial advantage in the market for titanium dioxide, the firm not only developed and patented a new, less polluting process, but at the same time greatly expanded production capacity such that it was capable of satisfying 95 per cent of forecast increases in demand. These actions significantly lowered du Pont's costs and it became the dominant firm. This illustrates what a firm can do to achieve lower costs and thereby derive

---

6 QCMA, op. cit.: 17,246. This again emphasises the group-specific nature of the QCMA approach.
a sustainable competitive advantage; but the package had the effect (and possibly the purpose) of market foreclosure.

The role of strategic behaviour within firms raises concerns not only in relation to identifying and pursuing breaches of Part IV of the Trade Practices Act, but also in relation to remedies for breach. Porter (1981:453) argues:

That an overall strategy guides strategic interaction also implies that a remedy aimed at one aspect of a firm’s behaviour must be probed to see how it will affect the ability of the firm to carry out its previous strategy, and whether the firm is likely to adjust other elements of its strategy to compensate or redefine its strategy completely. The firm will strive to maintain an internally consistent approach to competing, and one to which it is uniquely suited.

An attempt by the ACCC or the ACT to correct behaviour in one market could thus have socially undesirable consequences in another, if firms pursue coordinated strategic behaviour across several markets. This problem is not new; but it needs to be addressed with a strategic behaviour approach to examining firm conduct.

Some Examples of Strategic Behaviour

One of the most difficult tasks for antitrust investigators is to determine, in the absence of ‘the smoking gun’, whether identical prices between rivals reflect successful collusion between them or are the outcome of genuine competition. Equally difficult is dealing with conduct such as alleged misuse of market power, where investigators confront the problem of distinguishing between action which is anti-competitive and (as it is usually characterised by the perpetrator) action which is pro-competitive. Some examples will illustrate this point.

In a Federal Court case involving a price war between an established suburban newspaper and a new entrant, the issue was whether the lowering of prices by the incumbent, the Wentworth Courier, was justifiable as a necessary competitive response to a new entrant, Eastern Express, with superior technology, or whether it was designed to drive out of the market an otherwise efficient rival. Alternatively, with just two participants in the market, did the new entrant have a strategy for driving the incumbent from the market? With existing technology, was the market large enough to sustain only one participant? Was the observed conduct by both parties that which could be expected to occur in a competitive market, or did it go beyond the boundaries of socially acceptable good, hard competition?

In a matter before the Trade Practices Tribunal in 1977, Ford sought approval for franchising restrictions which it wanted to impose on its dealers, arguing that it led to

---

greater promotion and sales of Ford vehicles and therefore enhanced competition in the market. The Tribunal stated:

We reject the argument that an enhancement of competitive strength of a major participant in a market necessarily increases competition. To assess the effect on competition, it is essential to examine whether that enhancement has resulted from the imposition of a restraint on potential or existing competitors. Ford's argument would involve accepting that a dominant company which increased its competitive strength by weakening or eliminating its rivals was thereby increasing competition. Such action would increase that company's competitive position but instead of increasing competition in the industry concerned it would have the effect of decreasing competition.  

Clearly all firms would like to secure from their rivals as many of the most suitable retailing outlets as possible, as this raises the costs of existing or new rivals and forces them to invest heavily in their own networks or to use less suitable alternative distribution outlets. But does this damage long-term prospects for competitive behaviour in the market? The Chicago school of economic thought would argue that such exclusive dealing is pro-competitive, as it enhances economic efficiency and promotes inter-firm market rivalry. 10 There is no clear-cut answer.

'Creeping acquisitions' have become increasingly common in Australia in recent years. Here a major player in a market gradually acquires smaller rivals, leading to an increase in the costs of competing for the remaining (smaller) rivals. A good example of this is the so-called 'cheque-book competition' by the major supermarket chains. The purchase of independent retailers does not appear substantially to lessen competition in the relevant market as numerous other retailers remain: that is, there has been no significant structural change in the market. However, if such purchases are part of a strategy to deprive independent wholesalers of access to key outlets or sales volume so that their average unit costs rise and their ability to compete with the chains is diminished, consideration of this activity under a strategic framework might well result in a different conclusion.

Providing an Analytical Framework

The following analysis provides a simplified example of how strategic behaviour might be assessed in practice.

In 1994 Rank Commercial Ltd, a New Zealand company, announced a takeover bid for the assets of Foodland Associated Limited (Foodland). An agreement be-

---

10 For a summary of the arguments for and against vertical restrictions, see Carlton and Perloff (1994:522-44).
11 We wish to thank Stephen King for his helpful suggestions on the presentation of this example.
tween Rank and Coles Myer Limited (Coles) would have seen Foodland’s Australian assets pass to Coles. Foodland at that time was the only wholesaler of groceries to independent retailers in Western Australia. It also owned a number of retail grocery stores operating under the Action banner. It was estimated that with the acquisition, Coles’s share of grocery sales would have increased to 75 per cent (prior to the merger Coles’s share was 23 per cent).

In order to analyse the proposed acquisition, the following assumptions are made:

- Significant economies of scale/scope exist in grocery wholesaling and retailing.
- Following the acquisition, Coles would obtain improved terms from its suppliers, if only because of the increased volume of purchases. Economies of scale, as well as pecuniary benefits, exist in obtaining grocery products from manufacturers.
- Barriers to entry into grocery wholesaling are high. A new independent wholesaler would be unlikely to set up in Western Australia if Coles were to exercise its market power following such an acquisition. Nor is there any real likelihood that the existing independent wholesaler in South Australia would sell into Western Australia.
- At the time of the proposed acquisition, Coles, Woolworths, Action Stores and the other independent retailers were competing strongly with each other for the retail supply of groceries to consumers.

After the acquisition, Coles would face the following options:

- Retain the cost savings from the acquisition (no price changes) and increase its profits.
- Use the reduction in supply costs to lower the prices to independent retailers and to Coles’s own retail outlets (full pass through).
- Raise prices to the independent retailers in order to shift customers to its own retail outlets (this would also give some benefit to Woolworths). Coles’s own prices might increase, decrease or remain constant. It is assumed that they remain constant, and that as a result Woolworths’s prices are also unchanged.

Now consider whether the third option, raising rivals’ costs, is rational. If Coles makes no changes, it benefits from the acquisition to the extent of the cost savings. To undertake an alternative strategy in relation to its role as supplier to the independent retailers requires that this results in a more profitable outcome, at least in the long run. Whether or not this is the case depends on, first, the extent of the flow of retail customers from the independents to Woolworths (such movement results in a loss of sales to Coles), and second, the extent to which the sales made from wholesaling to
independent retailers are less profitable than supplying Coles’s own vertically integrated retail outlets.

The possible outcomes from this option, relative to simply taking the extra profit associated with the cost saving, can be represented as follows:

<table>
<thead>
<tr>
<th>Wholesale profits from supply to independents relative to supply to Coles’s own outlets</th>
<th>Flow of customers from independents to Woolworths</th>
</tr>
</thead>
<tbody>
<tr>
<td>High profits</td>
<td>High flow</td>
</tr>
<tr>
<td>Low profits</td>
<td>Low flow</td>
</tr>
</tbody>
</table>

Unfavourable Favourable

Assume that Coles derives at least the same profit from supplying independent retailers as from supplying its own retailers (denoted as high profits in Table 1). If the proportion of former customers of the independent retailers shifting to Woolworths is high, it is not rational for Coles to raise supply costs and hence the prices of the independents. It will be worthwhile to do so only if the profits from supplying independent retailers are at least as high as from supplying Coles’s own retail outlets and if the proportion of former customers of independents shifting to Woolworths is low.

Now assume that the profits from supplying independents are lower than from supplying its own vertically integrated outlets (denoted as low profits in Table 1). Then it will be in Coles’s interests to raise rivals’ costs to force them out of business irrespective of whether a high or a low proportion of customers shifts from independents to Woolworths. The difference in average store size suggests that supply costs to independent retailers are higher, so higher profits are likely to be derived from supplying Coles’s own outlets. In addition, profits are derived only from wholesaling in the case of supply to independents, but are derived from both wholesaling and retailing in relation to Coles’s own outlets.

Under a structuralist approach, after defining the market attention will focus on the consequences for competition and hence consumers of forcing the independent retailers out of the market. The high (and possibly heightened) barriers to entry, the vertical integration of the retail chains and the high level of market concentration suggest that, absent the independent retailers, Coles may be able to raise prices with a reasonable expectation that Woolworths will follow. Further, it is unlikely that much consideration would be given to Woolworths’s responses to Coles’s initiatives.

Now assume that for several years before the acquisition Coles spent heavily to establish a reputation for ‘service and low everyday prices’. Consumers have to incur significant search costs if they want to be well-informed about grocery prices: the number of grocery items is very large, prices tend to change fairly frequently, and consumers do not make their purchasing decision on the basis of individual grocery items.
but on a package of groceries and associated services. Generally, consumers undertake only periodic searches because of the opportunity cost of search time, and rely instead on information conveyed by the reputation of the retailer.

However, if the retailer ‘cheats’ on its reputation, customers will go elsewhere and will be difficult to retrieve. The retailer would need to offer more than was originally promised in order to win them back. This could be a lengthy and costly process. Under these circumstances, it is much less likely that, following the exit of the independent retailers, Coles would be able to achieve a sustainable increase in prices. Maintaining its reputation would certainly be a constraint on Coles’s pricing decisions. However, the constraint is not a structural one and is separate in time from the acquisition. It is not clear that the structuralist approach to analysing the acquisition would consider such a constraint.

**Consequences for Australian Competition Policy**

Incorporation of a strategic behaviour framework would not have necessarily significantly changed any previous decisions made by the ACT or the courts in Australia based on the QCMA approach. But this structural approach, with its focus on whether there is a causal sequence from market structure to conduct to performance, generates a significant potential risk of a wrong assessment of market conduct with respect to its effect on competition and hence as to whether there is a breach of the Trade Practices Act. Greater flexibility of analysis is needed.

While market behaviour should not be assessed without at least a careful consideration of the group or collective effects of market structure, the conduct of individual firms in the relevant institutional context — past, present and future, as firms jockey to make the most of their competitive environment — should become the prime focus of analysis. At the same time, the focus needs to be broadened away from the traditional area of the market to the firm’s entire competitive environment, as the intricately intertwined operations of modern firms transcend the narrow market context.

The strategic behaviour approach adds several dimensions to the QCMA approach. It focuses more on individual behaviour than on group conduct; it looks to the future rather than to the past for its solutions; it recognises the need to examine the entire integrated pattern of a firm’s behaviour across all the markets in which it operates; it recognises that there is no uniquely predictable outcome for any given situation; and it is more applicable to the analysis of misuse of market power scenarios under s.46 of the Trade Practices Act than is the QCMA approach. The central

---

Footnotes:

12 Feedback effects are allowed in the structural model, but usually are recognised only for their impact on structure.

13 We are not implying here that analysis of issues brought under s.46 involves a structural approach. Rather, we simply make the point that analysis of competition matters in Australia based generally on a strategic behaviour framework would make for a more unified and consistent approach than is currently the case.
task for those sitting in judgment on firms is to distinguish pro-competitive from anti-competitive strategic behaviour.

A recent unreported decision by the New Zealand High Court has laid a foundation for a departure from a pure market-power standard for testing whether conduct might substantially lessen competition. The Court acknowledged that 'even a monopolist has rights to compete', and cited with approval decisions in the US where it was concluded that a firm with lawful market power had no general duty to exercise competitive restraint and should be expected to compete aggressively. Of course, while it is encouraging to see this endorsement of the social and judicial acceptability of strategic behaviour, the Court did not provide (and, realistically, could not hope to provide) a definitive or universal test for distinguishing between pro-competitive and anti-competitive strategic behaviour. It all depends on the facts of each case.

As the noted British academic and consultant John Kay pointed out in his comments on an earlier version of this article presented at the Trade Practices Workshop, the key question to ask when a firm is charged with acting anti-competitively is whether its actions are essentially profitable in their own right (that is, independent of what rivals do), or whether profitability depends on the firm's expectations about the actions and reactions of its rivals. In other words, could the action be taken anyway, on a stand-alone basis, regardless of the responses of rivals, or alternatively, could such behaviour be observed in a market normally judged to be competitive? If the latter, the strategic behaviour is not likely to damage the competitive process.

This test cannot be taken as the exclusive test of whether strategic behaviour is anti-competitive. If this were the case, all technically efficient strategic behaviour could be found to be competitive, regardless of its outcome on structure. Also excluded from the reach of the law would be behaviour which is privately profitable to the firm because of a combination of both efficiency and market power factors. Clearly a balancing of the two factors is necessary. Rather, a test based on strategic behaviour adds a further dimension to the determination of whether conduct substantially lessens competition.

Conclusion

Through strategic behaviour firms aim to achieve a sustainable, commercial advantage over their rivals. A firm engaged in strategic behaviour does not simply react to rivals' moves. The very essence of oligopoly is a conscious striving by firms to surpass their rivals, in the process securing as many sales as possible and a strong brand image, immune from rivals' strategies. Other firms, in order to survive, will react and seek to counter these actions. This is part of the vigorous process of competition which competition policy seeks to encourage and preserve.

In a market where firms behave strategically and non-cooperatively, the structural features of the market may be altered or there may be non-structural con-

---

14 Clear Communications Limited v Sky Network Television Limited & Ors, New Zealand High Court, unreported judgment, 1 August 1997, at pp. 69-70. We are grateful to Frances Hanks for bringing this decision to our attention.
straints on the conduct of firms. Under certain circumstances, strategic behaviour may result in the creation of market power which does not offer a stimulus to other firms to strive to match or surpass that firm's strategies, but which has the effect of foreclosing the market to new entrants: as in the case where strategic behaviour can seek deliberately to erect entry barriers even though currently they may not be present. When this occurs, it leads to a substantial lessening of competition.

On its own, a structuralist approach to analysing competition may focus on only one aspect of the firm's strategies; failure to analyse the entire strategy may suggest a substantial lessening of competition when none is likely, and vice versa. The danger of making such errors will be substantially reduced if firm conduct is assessed under an approach that takes into account all of the strategic activities engaged in by firms.

The recognition of strategic behaviour has far-reaching implications for competition policy and enforcement. Adoption of a strategic behaviour framework would undoubtedly extend the scope of conduct to be considered. Unilateral conduct by firms with reasonably small market shares might be judged to result in a substantial lessening of competition; more firms would be at risk of having action taken against their conduct. At the other extreme, properly addressing the strategic aspects of a market might eliminate concerns in some cases even where market concentration is high. But most important, incorporation of strategic behaviour into the analytical framework ensures that in competition analysis it is recognised that oligopoly and interdependence are not the exception but rather are the norm.

References

This article is a revised version of a paper presented to the 1997 Trade Practices Workshop conducted by the Business Law Section of the Law Council of Australia in Adelaide on 19 July 1997. We are especially grateful to John Kay for his support for the ideas presented here, and to Frances Hanks for her constructive comments on a previous draft.

---

15 All firms would like the quiet life which market power brings with it. But as market power cannot be exploited without the protection of entry barriers, seekers of market power need to engage in conduct designed to create such entry barriers. A firm can be lucky enough to be in the right place at the right time, as could happen, for example, with an advertising campaign for a small new firm which propels it into the limelight. But this is an example of conscious strategic behaviour.
The Political Costs of Tax Reform: A Canadian Perspective

David L. Ryan and Stuart Landon

THE Commonwealth government's forthcoming proposals to reform Australia's taxation system could well include the introduction of a goods and services tax (GST). This would be a broad-based consumption tax that may either be hidden in the price of goods or, as in Canada, added on explicitly at the point of sale.

Political considerations are likely to play a major role in the choice of modifications to specific taxes. For example, voters may be concerned only with the total tax paid, along with the total amount of government services provided (and, possibly, associated changes in government debt if revenues and expenditures do not balance). In this case, tax reform that increases the efficiency of the tax system and is revenue-neutral, as the government of Canada claimed of its GST when it was introduced in 1991, should not affect the government's electoral chances. However, if voters do not treat all types of taxes as being indistinguishable, then there are different political costs associated with alternative types of taxes and these may differ substantially from the economic costs of these taxes. To the extent that different forms of taxes affect voters differently, governments may choose those tax instruments which have lower political costs rather than those that have lower economic costs. As noted by Hettich and Winer (1984), the politically optimal tax structure is one in which the marginal political cost, rather than the marginal economic cost, of raising an additional dollar of tax revenue is the same for all types of taxes.

Different types of taxes might be expected to have different effects on voting behaviour, and hence on the political success of the incumbent party, for several reasons. For example, some taxes may be difficult to perceive because they are hidden in prices (excise taxes) or deducted at source (corporate taxes, resource rent taxes), so that the voters are unaware of the true size of the tax burden. Other taxes may be relatively small and paid only infrequently (licence fees), while others may

David Ryan and Stuart Landon are Associate Professors of Economics at The University of Alberta.

1 'At the proposed rate of 7 per cent (and allowing for other coincident fiscal measures) the GST is intended to be revenue neutral - yielding ... roughly the same amount that would have been raised in 1991 under the federal sales tax system' (Mitchener, 1990:3). Estimates of the tax revenues under the then existing tax system and the proposed 7 per cent GST, which indicate the revenue neutrality of the proposed tax, can be found in Canada, Library of Parliament (1990).
continually reinforce the voter’s perception of the tax (sales taxes, especially those added explicitly at the point of sale). Furthermore, the costs of taxes may be distributed unevenly across the electorate, and this distribution may differ for different types of taxes. Since the political costs of taxes may differ, and these differences may affect the types of taxes chosen, political costs may have a significant impact on the tax structure and, in particular, on whether a tax such as a GST is adopted in Australia.

In this article, the relative political costs of different types of taxes are evaluated empirically, with particular attention being paid to a visible GST-type tax. Unfortunately, this type of analysis is not feasible using Australian data because there does not exist in Australia, at either the State or federal level, a tax which is similar to a visible GST. Further, even an evaluation of the political costs of the hidden federal sales tax is not possible given the small number of post-war federal elections (and, thus, observations). While there have been a larger number of State elections during this period, no equivalent State-level tax exists.

Given the absence of relevant Australian data, the political costs of different taxes are evaluated here using data for Canada. Canada’s GST is a federal tax that was introduced at the beginning of 1991. As there have been only two Canadian federal elections since that time, and no variation in the GST tax rate, sufficient Canadian federal data do not exist to evaluate the political costs of this tax. However, all Canadian provinces except one have for some time levied visible sales taxes. Like the Canadian GST, these sales taxes are explicitly added to the purchase price at the point of sale and, therefore, might be expected to have similar political implications as the GST. In addition, these taxes have varied in size across provinces and through time, and data are available for a relatively large number of Canadian provincial elections. As a result, the evaluation of the political costs of a visible GST-type sales tax presented here is based on results obtained using data on Canadian provincial-level taxes and elections. However, it should be noted that the results obtained will reflect the impact of a change in existing tax rates rather than the introduction of a new sales tax.

Since a sales tax is an alternative to other types of taxes, in order to determine the relative political cost of sales taxes it is also necessary to evaluate the costs of other forms of taxes. Hence, in addition to sales taxes, in the analysis that follows we consider the (potentially) distinct effects on political success of direct taxes on persons, direct taxes on businesses or corporations, property taxes, taxes on fuel or specific natural resources, and licence or permit fees. Further, since the political success of an incumbent party will depend on the services the government provides as well as the taxes it levies, it is also necessary to evaluate the political costs and benefits of different types of expenditures. The alternative government expenditure instruments to be considered include government purchases of goods and services as well as transfers from government to persons, transfers to businesses (subsidies and capital assistance are considered separately), transfers to other levels of government, and transfers to hospitals.
Modelling Voting Behaviour

Since the political cost of a particular fiscal policy depends on its impact on voting behaviour, an empirical analysis of political costs must begin with the specification of a model of voting behaviour. Voting can be viewed as a process in which each voter votes for the political party which, if elected, is expected to make him or her better off than any other political party, in the light of the voter's forecasts of what changes will occur if that party is elected and of how, and to what extent, the voter will be affected by these changes. Following Deacon and Shapiro (1975), our analysis is based on an economic model of voting behaviour in which voters focus on expected changes in their gross income as well as on expected changes in various types of taxes and government expenditures and in the level of government debt.

According to this model, voters consider what policies they expect, and hence how well off they are likely to be, if the incumbent party is re-elected, and then decide whether or not to vote for that party. They use information on how the incumbent party has performed since the last election as a signal of the incumbent's future performance. In particular, they base their decisions on the changes since the previous election in various types of taxes and government spending, as well as in the level of government debt and gross income.

Voters have access to additional information which may signal an incumbent's future behaviour, and which may, therefore, have a systematic effect on the probability that a voter will vote for the incumbent party. This information may include variables reflecting the incumbent's overall economic performance, such as changes in the unemployment rate, as well as other variables that reflect particular characteristics of the incumbent party, such as years in office.

Modelling Political Costs

The political cost (or benefit) of the actions taken by a political party can be measured in terms of the impact of these actions on the objectives of the party. According to Downs (1957) and Riker and Ordeshook (1973), the goal of candidates and of parties is to win elections. While Downs equates winning elections with maximising votes, Riker and Ordeshook provide four different interpretations of winning, implying four corresponding possible objectives for political parties and candidates: maximising their plurality; maximising the number of votes received; maximising the proportion of votes received; and maximising the probability of winning. With a first-past-the-post electoral system, as used in Canada, a party's proportion of seats and its proportion of votes may differ significantly.

Nevertheless, empirical models of voter behaviour have tended to concentrate on explaining the percentage of the vote received (Deacon & Shapiro, 1975; Peltz-

---

2 A similar type of analysis has been used recently by Peltzman (1992), although he concentrated on the political impact of different types of expenditures rather than different types of taxes, and by Schram and Van Winden (1989) and Schram (1990), although they do not examine the impact of taxes and consider only a limited disaggregation of government expenditure.
man, 1990, 1992), perhaps in part because this measure is thought to utilise more information than whether a party simply won or lost. However, except in single-jurisdiction two-alternative contests, a party's vote percentage (like its number of votes or percentage of seats) does not generally imply anything about victory or defeat. A more important reason for concentrating on the vote percentage is that a political party which is interested in maximising the probability that it will obtain the vote of a randomly chosen voter can be viewed as if it is maximising its percentage of the vote (Deacon & Shapiro, 1975). In this context, the marginal political costs of particular fiscal policies can be defined as the effect on this vote percentage (that is, on the probability that a randomly chosen voter will vote for the party) of changes in various taxes and expenditures.

In terms of evaluating political costs, however, it may be more useful to view a political party's objective as being maximisation of its probability of victory rather than maximisation of its percentage of the vote. There are several reasons for this, but the most important for present purposes is that, when the objective is defined as the maximisation of the number (or percentage) of votes, the preferences of all voters, including decisive voters as well as voters in marginal non-decisive groups, are given equal weight. Assuming that the objective of a party is political victory, the appropriate measure of the political cost (or benefit) of a policy is its effect on the probability of victory, in which case the preferences of decisive voters are weighted more heavily than the preferences of marginal voters. Indeed, a political party that is interested in maximising the probability of victory can be viewed as if it is maximising the probability that it will receive the vote of the decisive voter.

In view of these considerations, in the empirical analysis reported here the political costs of particular types of taxes and government expenditures are evaluated in terms of their effect on both the percentage of the vote (where the objective is to obtain the vote of a random voter) and the probability of winning (where the focus is on the vote of the decisive voter). As a result, two different equations are estimated: one describing the percentage of the vote won by the opposition and the other describing the election outcome (where 0 indicates incumbent victory and 1 indicates incumbent defeat). In order to examine the impact of different types of taxes and government expenditures on each of these measures of political success, in both cases the explanatory variables include changes in various types of taxes, transfer payments and government spending, as well as changes in income and in the government debt. Additional variables which may signal the incumbent's future behaviour (which are described more fully below) are also included as explanatory variables in both equations.

While policies that voters believe to be beneficial are likely to increase both the vote percentage and the probability of winning, the different weights attached to the preferences of decisive and non-decisive voters by the two objectives may cause

---

3 Complete details of the models estimated and the method of empirical implementation (grouped probit for the percentage vote equation and probit for the probability of incumbent defeat equation) are contained in Landon and Ryan (1997).
significant differences in the quantitative measures of the political costs of these policies. A comparison of the empirical results obtained under each of these two alternative political objectives may indicate the importance of correctly identifying the government's objective when attempting to evaluate the political costs of changes in its fiscal program.

Application to Canadian Provincial Data

The economic data used to generate the explanatory variables required to estimate the equations of percentage vote and probability of incumbent defeat are taken from Canada's annual provincial national accounts for the years 1961-90 inclusive. During this period there were 82 provincial elections in Canada's ten provinces, although one election in the province of Newfoundland (in March 1972) occurred less than five months after the preceding election and was omitted since it could not be separately associated with annual economic data. These provincial elections, which follow no particular timetable, were matched with annual economic data by associating the previous year's data with an election which took place on or before 30 June, and data from the current year with elections that occurred on or after 1 July. In view of the differences in the size of government across provinces and the sample period, and since elections do not occur at specified intervals, variables were normalised by dividing by the level of real provincial per-capita GDP for the election year, and were converted into average annual changes since the last election. Since the data refer to changes between elections, the 81 included elections yielded 71 useable observations.

Relatively minor aggregation of expenditure and tax categories left data available for the real annual per-capita change (since the previous election) in the revenue obtained from eight different types of taxes: direct taxes on persons, corporate taxes, gasoline taxes, natural resource taxes, sales taxes, miscellaneous indirect taxes, licences, permits and other fees, and the provincial property tax; and in six types of government spending and transfers: expenditure on goods and services, transfers to persons, transfers to business (subsidies), transfers to business (capital assistance), transfers to local government, and transfers to hospitals.

As noted above, additional variables which may signal the future behaviour of the incumbent are also included in the empirical specification. The change between elections in the provincial unemployment rate is used as a possible signal of the incumbent's overall economic policy success. Since voters might believe that their well-being depends on whether the incumbent party can cooperate with or counterbalance the federal government, a dummy variable is included to reflect whether the incumbent provincial party was the same as the governing national party. As well, since a new leader might not be held responsible by voters for past actions of the government, we include a dummy variable which indicates if the incumbent party

4 Tax variables are calculated from total revenues rather than from tax rates since the appropriate data are unavailable and these revenues represent the cost to individuals, in terms of lost consumption, of the different taxes.
changed leaders within the twelve months prior to the election. Finally, two additional variables — the number of years that the incumbent party has held office, and this value squared — are included since voters may perceive the incumbent party to be more alienated from the public the longer it is in power, and this perception may not grow at a constant rate.

Note that since monetary policy is a federal responsibility, and is taken as given by voters in provincial elections, inflation is not included as an explanatory variable. While inflation may have an effect on taxes paid if the tax system is not inflation neutral, these effects should be accounted for by the use here of tax revenues rather than tax rates.

Empirical Findings

As explained in the preceding sections, the political costs of different taxes are determined by estimating two equations, one describing the percentage of the vote received by the opposition, and the other describing whether the opposition won or lost the election. The opposition's percentage of the vote ranges between 37.35 per cent and 80.35 per cent over the 71 different elections, with an average value of 54.49 per cent. Thus, on average, the probability that a randomly chosen voter would vote for the opposition was almost 0.55. However, the opposition won only 24 of the 71 elections, so that the probability that the decisive voter would vote for the opposition (in which case the incumbent would lose) was only 0.34. Hence, differences might be expected in the results obtained with the two estimating equations.

As indicated by the results presented in Table 1, the only tax variables that have a statistically significant impact on political success in both estimating equations are sales taxes and licences, permits and other fees. Increases in sales taxes reduce both the incumbent's vote percentage and probability of winning, while increases in licences, permits and other fees improve both these measures of the incumbent's political success. In addition, based on the estimates for the equation describing the probability of incumbent defeat, both increases in direct taxes on persons and increases in gasoline taxes significantly improve the opposition's electoral chances. A test of whether all the tax variables have the same marginal impact on the incumbent's political success (that is, whether they all have the same marginal political costs) indicates that the effects of the different taxes are statistically significantly different.

The only type of government spending that significantly improves the incumbent's electoral position is spending on goods and services. Other types of government expenditure either have no significant effect or actually significantly assist the opposition's electoral position, particularly transfers to persons and to local government (in both models) as well as business subsidies (only in the probability of incumbent defeat equation).

As for the other variables included in the estimating equations, increases in per-capita provincial debt decrease the incumbent's vote percentage, but have an insignificant effect on the probability that the incumbent will lose. This suggests that
decisive voters misjudge the future budget consequences of current spending and taxation policies, or they have sufficiently high discount rates or, in the context of provincial elections, they believe that they can avoid future taxes by moving to another province. Increases in real per-capita income tend to raise the incumbent's vote percentage and probability of victory, but not significantly. Although not reported in Table 1, the empirical results also indicate that, as its term in office lengthens, the percentage vote for the incumbent as well as its probability of victory decrease. In addition, there is a significant increase in the incumbent's vote percentage, but not probability of winning, if both the provincial and the federal governments are from the same party. However, neither of the other variables (the change in the unemployment rate and a variable reflecting a change in leaders for the incumbent party prior to the election) is found to have a significant effect on either measure of electoral success.

Table 1: Impact of a one-dollar increase in specific taxes and expenditures

<table>
<thead>
<tr>
<th>Tax variables:</th>
<th>Effect on the opposition's vote percentage</th>
<th>Effect on the probability of incumbent defeat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct taxes on persons</td>
<td>.0008</td>
<td>.0040*</td>
</tr>
<tr>
<td>Corporate taxes</td>
<td>.0002</td>
<td>.0058</td>
</tr>
<tr>
<td>Gasoline tax</td>
<td>.0009</td>
<td>.0127*</td>
</tr>
<tr>
<td>Natural resource taxes</td>
<td>.0005</td>
<td>-.0025</td>
</tr>
<tr>
<td>Sales taxes</td>
<td>.0017**</td>
<td>.0052*</td>
</tr>
<tr>
<td>Miscellaneous indirect taxes</td>
<td>.0007</td>
<td>.0004</td>
</tr>
<tr>
<td>Licences, permits and other fees</td>
<td>-.0031***</td>
<td>-.0195**</td>
</tr>
<tr>
<td>Provincial property tax</td>
<td>.0007</td>
<td>.0065</td>
</tr>
</tbody>
</table>

| Expenditure variables:                             |                                           |                                             |
| Expenditure on goods and services                  | -.0006*                                  | -.0041***                                   |
| Transfers to persons                               | .0010*                                   | .0051*                                      |
| Transfers to business: subsidies                   | .0011                                    | .0093**                                     |
| Transfers to business: capital assistance          | -.0037                                   | -.0002                                      |
| Transfers to local government                      | .0012*                                   | .0062**                                     |
| Transfers to hospitals                             | .0011                                    | -.0059                                      |

| Other economic variables:                          |                                           |                                             |
| Debt                                               | .0001*                                   | .0004                                       |
| Income                                             | -.0001                                   | -.0003                                      |

Note: ***, **, and * indicate significance of the estimated coefficient on the corresponding variable at the 1 per cent, 5 per cent and 10 per cent levels, respectively.

The main difference between the results for the two objective functions is that the effects of the different types of taxes and expenditures on the probability of victory tend to be larger and are estimated more precisely (that is, more coefficients are statistically significant and they tend to be significant at a higher level of confi-
The Magnitude of Marginal Political Costs

Numerical estimates of the marginal political costs of each tax and type of government spending are presented in Table 1 for both political cost models. These estimates reflect the impact of a one-dollar increase in each of these variables on either the vote percentage of the opposition or the probability that the opposition will win the election. Since the magnitudes of these effects differ from election to election, the values presented in Table 1 are averages over the 71 observations (elections). These results indicate that a one-dollar increase in per-capita revenue from the visible sales tax would reduce the incumbent's vote percentage by 0.0017 and reduce the incumbent's probability of winning by 0.0052. This latter effect is smaller than the impact of an increase in the gasoline tax on the probability of defeat (the gasoline tax has an insignificant effect on the percentage of the vote), but larger than the impact of an increase in direct taxes on persons (which also has an insignificant effect on the vote percentage). Note, in addition, that increased debt financing has an insignificant effect on the probability of defeat, although it has a significantly negative effect on the incumbent's vote percentage. Thus, debt financing (at least until taxes must be raised to meet interest and principal payments) may be preferred over sales taxes by the incumbent party.

While the magnitudes of the estimated effects associated with corporate taxes, natural resource taxes, miscellaneous indirect taxes, and the provincial property tax are large in some instances, these effects are not statistically significant. Thus, while these taxes may have had a large impact on some elections, they have not had a systematic effect across elections.

In contrast to visible sales taxes which have a relatively large political cost, increases in licences, permits and other fees provide a large marginal political benefit to the incumbent, with the magnitude of this effect dwarfing that of any other tax (or expenditure). Unlike other taxes, these fees tend to be service specific, and voters may see them as resulting in better service (such as shorter queues). In addition, the average voter may not use many of the services funded by these fees and so may prefer that they be financed on a user-pay basis rather than out of general revenues.

In terms of the effects of the various expenditure variables, expenditure on goods and services would increase the incumbent's percentage of the vote by 0.0006 and the incumbent's probability of winning by 0.004. In contrast, transfers to local government would decrease their vote percentage by 0.0012 and their probability of winning by 0.0062. While no other expenditure variables have significant effects in the vote percentage equation, in the probability of winning equation the largest sig-
significant effect on the incumbent party's success is associated with transfers to business, which decrease the probability of victory by 0.0093.

**Political Expediency vs the Economic Efficiency of Different Taxes**

It has often been noted that political decisions do not necessarily correspond to the types of decisions that could be justified on purely economic grounds. Although the analysis here has concentrated on the political costs of different types of taxes (and government expenditures), there is a considerable literature concerning the economic cost or efficiency of alternative forms of taxation. A comparison of the political cost estimates obtained here with information about the economic costs of these different types of taxes may provide some indication of the extent to which economic and political tax-choice decisions are likely to coincide. It may also provide an indication of whether political and economic arguments are likely to correspond or diverge in the debate on tax reform in Australia.

In a study of the marginal efficiency costs of various US taxes, Jorgenson and Yun (1991) found that corporate taxes were the least efficient, followed in order of increasing efficiency by individual income taxes, sales taxes, and property taxes. In the empirical work reported here, corporate taxes and property taxes, the two taxes at the opposite ends of this list, are found to have statistically insignificant effects, indicating that the marginal political costs of these two tax instruments are statistically indistinguishable. However, sales taxes are found here to have a higher marginal political cost than direct taxes, so that the estimated political cost ranking of these two taxation instruments is the reverse of their efficiency cost ranking. Thus, to the extent that the efficiency ranking of tax instruments in the US is applicable more generally, the marginal political cost estimates provided here indicate that governments that wish to reduce the political costs of generating revenue may not choose taxes that are the most appropriate from the point of view of economic efficiency.

**Predictions of Vote Percentages and Election Outcomes**

In common with most economic models, the results reported here rely on empirical specifications that embody a number of assumptions. For example, in the specified models it is assumed that voters make electoral choices based on economic factors which include lagged changes in income, disaggregated tax revenues, disaggregated government spending, and government debt. While the assumptions embedded in the estimated models are not directly testable, one way of assessing the overall usefulness of the modelling approach employed here (as opposed to, say, an approach in which voting depends more on sociological or psychological factors) is to compare the predictions of the models with observed election outcomes. To provide a more stringent assessment of the predictive ability of the models, predictions are made for elections that are not included in the sample used to estimate the two models. To this end, the models are re-estimated with data from the last seven provincial elections (each for a different province) omitted.
Based on these revised estimates, predictions of the incumbent's vote percentage, and of the election outcome (based on predictions of the probability that the incumbent will lose) are made for each of the seven omitted elections. Following conventional practice (Greene, 1993:651-2), values of the probability of losing that exceed 0.5 are viewed as indicating that the incumbent will lose.

Table 2: Actual and predicted election outcomes

<table>
<thead>
<tr>
<th>Province</th>
<th>Date of election</th>
<th>Actual incumbent vote percentage</th>
<th>Predicted incumbent vote percentage</th>
<th>Actual change in govt.?</th>
<th>Predicted probability of loss by incumbent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nova Scotia</td>
<td>Sept. 1988</td>
<td>.435</td>
<td>.476</td>
<td>No</td>
<td>.006†</td>
</tr>
<tr>
<td>Alberta</td>
<td>March 1989</td>
<td>.443</td>
<td>.366</td>
<td>No</td>
<td>.431†</td>
</tr>
<tr>
<td>Newfoundland</td>
<td>April 1989</td>
<td>.476</td>
<td>.413</td>
<td>Yes</td>
<td>.947†</td>
</tr>
<tr>
<td>Prince Edward Island</td>
<td>May 1989</td>
<td>.607</td>
<td>.420</td>
<td>No</td>
<td>.393†</td>
</tr>
<tr>
<td>Quebec</td>
<td>Sept. 1989</td>
<td>.500</td>
<td>.454</td>
<td>No</td>
<td>.013†</td>
</tr>
<tr>
<td>Manitoba</td>
<td>Aug. 1990</td>
<td>.420</td>
<td>.519</td>
<td>No</td>
<td>.282†</td>
</tr>
<tr>
<td>Ontario</td>
<td>Sept. 1990</td>
<td>.324</td>
<td>.361</td>
<td>Yes</td>
<td>.984†</td>
</tr>
</tbody>
</table>

Note: Based on the prediction that the incumbent will lose if its probability of defeat exceeds 0.5, a † indicates that the model correctly predicted the election result.

Table 2 lists the province and month of the seven elections for which predictions are made, along with the actual incumbent vote percentage and the result of the election. The predictions of the probability of defeat for the incumbent are always less than 0.5 in the five cases where the incumbent won (and in two of these cases the predicted probabilities are almost zero), and above 0.5 (in fact above 0.9) in both cases in which the incumbent lost. Thus, the actual election result is correctly predicted in all seven cases. Estimates of the incumbent's vote percentage provide no direct prediction of the election result, and thus cannot be evaluated on the basis of actual election outcomes. However, based on the relatively low predicted incumbent vote percentage in the provinces of Alberta (36.6 per cent) and Prince Edward Island (42.0 per cent), one might have expected these governments to be defeated when they were actually re-elected. Along with the relatively imprecise coefficient estimates associated with the percentage vote model, these prediction results suggest that the probability of defeat model might be preferred to the vote percentage formulation.

Conclusion

The introduction of a GST-type visible sales tax is being considered as part of the current discussion on tax reform in Australia. Using Canadian provincial-level data, in conjunction with an economic model of voting behaviour, we obtain estimates of the impact on the success of the incumbent party of changes in a visible sales tax as well as the impact of changes in other forms of taxes and government expenditures.
The model performs well, in terms of both the significance of coefficients and its predictive power, especially when the objective of political parties is taken to be the maximisation of the probability of victory. This suggests that the framework employed here is a useful way of evaluating the political costs and benefits of different fiscal policies.

Whether the objective of the incumbent party is to maximise its percentage of the vote or its probability of victory, increases in visible sales taxes are found to have a significantly negative effect on the incumbent party's success. The effect of a sales tax increase on the probability of defeat is smaller in magnitude than an increase in the gasoline tax, but larger than the effect of an increase in direct taxes on persons (both of which are, at least to some extent, visible taxes). Less visible taxes, such as corporate taxes, natural resource taxes, and other indirect taxes, are shown to have a statistically insignificant impact on the incumbent party's success. The significantly positive impact of licence and other fees on the incumbent's political success suggests a distinct voter preference for user-pay methods of financing publicly provided goods.

Government spending on goods and services is the only type of spending that significantly reduces the probability of incumbent defeat and the opposition's vote percentage, while increased transfers to individuals, businesses or local governments have either a neutral or a detrimental impact (from the point of view of the incumbent). These results suggest that governments could maximise their political success by increasing spending on goods and services and reducing transfers to individuals, local governments and businesses. While an increase in the level of government debt appears to cause a very small but significant reduction in the incumbent's percentage of the vote, it does not have a significant effect on the probability of incumbent defeat. As a result, governments may prefer issuing debt rather than raising taxes in order to cover short-term revenue shortfalls.

Based on the results presented here, it would appear that governments wishing to improve their electoral chances are likely to reduce their reliance on broad-based visible taxes such as sales taxes, gasoline taxes and income taxes, and concentrate instead on raising revenue from less visible revenue sources such as natural resource royalties, corporate taxes, and user fees. Since the largest political costs are associated with the most visible taxes and, in particular, with a type of sales tax which is very similar to a visible GST, these results may have important implications for tax reform in Australia. A federal sales tax of this type was introduced in Canada in 1991 and became a major election issue during the 1993 Canadian federal election. In fact, the opposition party's promise to replace this tax, a major plank of its campaign, was probably one of the reasons it won a majority (although this promise was not kept during its first term in office). In contrast, the party which introduced the GST went from holding a majority of the seats in the Canadian parliament to holding just two of 295 federal seats.

Even though a visible GST-type tax can be supported on efficiency grounds, the associated political costs, as evidenced by the results reported here and the 1993 Canadian federal election, may be sufficiently large to make its introduction politi-
cally unappealing. Indeed, the results presented above suggest that governments considering tax reform may tend to concentrate on increasing taxes which are less visible or increasing user fees, since these are the least politically costly policies. Alternatively, if a GST is introduced, the results here suggest that this tax may be politically less costly if it is included in the price rather than added on explicitly at the point of sale. Despite the political appeal of taxes which are less visible or a hidden GST, tax reform that concentrates on these types of taxes is unlikely to be desirable either on efficiency grounds or in terms of promoting government accountability.

References


The authors gratefully acknowledge the helpful suggestions of three referees and the editor.
Abolishing Income Tax Deductions for Work-Related Expenses

Jonathan Baldry

UNDER the Australian personal income tax, wage and salary earners are allowed to claim tax deductions for work-related expenses (WREs). In income year 1993/94, 67 per cent of wage and salary earners (59 per cent of all personal income taxpayers) claimed deductions for WREs amounting to $4.86 billion, or 2.6 per cent of total declared wages and salaries. The most important component of the claims was for 'other work-related expenses', which includes tools, professional and trade journals, subscriptions to professional bodies, and conference fees; the total claimed amounted to $2.05 billion. The second most important category was motor vehicle claims, which totalled $1.54 billion. The most common claims were again for 'other work-related expenses' (4.10m taxpayers) and 'uniforms and laundry expenses' (2.86m taxpayers).

For most of these taxpayers, calculating claims for WRE deductions represents the major and often the only complication in completing the tax return. WRE deductibility is a major factor in the high compliance costs associated with the Australian personal income tax. It also generates a range of distortions and consumption caused by the implicit subsidies it gives to certain activities. Contrary to its presumed purpose, it seems to undermine the equity of the tax system: the differential access to deductible WREs available to the various occupational groups introduces an arbitrary element into the distribution of effective post-tax income, while in general conferring significant tax advantages on the better-off. These undesirable consequences together make a compelling case for the total or partial abolition of WRE deductibility in Australia, a move which would also put the Australian tax system in line with practice in the majority of developed economies. New Zealand, for example, abolished WRE deductibility in 1987.

The Law and Practice of WRE Deductions in Australia

Personal income taxpayers have been able to claim deductions for WREs since the first Income Tax Assessment Act (ITAA) of 1915. The relevant part of the 1936
ITAA, which (with amendments) defines current income tax law, is Section 51(1). This section defines an allowable ‘general deduction’ in these terms:

All losses and outgoings to the extent to which they are incurred in gaining or producing assessable income, or are necessarily incurred in carrying on a business for the purpose of gaining or producing income, shall be allowable deductions except to the extent to which they are losses or outgoings of capital, or of a capital, private or domestic nature, or are incurred in relation to the gaining or production of exempt income.

To be a legitimate deduction, a particular expense must be related to (but not necessarily required for) the generation of income. It must not be an expense being used to build up the wealth of the taxpayer. As well, the expense must be one which would not normally be incurred in a non-employment context — for example, as a cost associated with consumption or with non-market, non-taxable production. An interpretation of this qualifying clause is that an expense which generates a direct gain (‘utility’), or an indirect gain through the production of non-assessable income, shall not be allowable, because the direct or indirect gain is itself not taxed.

The main sources of interpretation of s.51(1) are legal judgments on appeals against assessments and the Australian Taxation Office’s (ATO’s) system of Rulings and Determinations. Determinations are a more recent innovation, which generally relate to the circumstances of particular occupational groups or to particular types of expenses. The large body of Determinations and Rulings has resulted in some general principles, many arbitrary and often contradictory distinctions, and a number of rules of thumb concerning ‘reasonable limits’ for particular types of WREs. One rule of thumb concerns the substantiation requirements for WREs in aggregate. In the ATO’s 1996 TaxPack, it is stated that ‘you must be able to substantiate … your claims if the total claimed is greater than $300’. However, for smaller total claims, ‘you need to be able to show how you worked out your claims … you do not need written evidence’ (p. 48). A recent crackdown on tax agents who automatically claim close to $300 in WRE deductions for all of their clients has shown that the $300 guide is not to be interpreted as a standard deduction.

The notion of ‘reasonable limits’ is a recognition that WRE deductions may be (improperly) used to subsidise consumption, or to promote other non-work purposes, if no restrictions are defined. The various interpretations of s.51(1) as it applies in particular groups of cases aim to establish both the relevance of an expense to an occupation and the extent to which the expense would be likely to be incurred in another context. This latter condition, which reflects the second, limiting clause of s.51(1), lies at the heart of most Rulings and Determinations related to WREs, and produces most of the arbitrariness and contradictions. For example, a

---

2 The status and nature of Determinations are detailed in the first of a new series of Determinations, reference TD 92/100. A previous series of Determinations was issued on capital gains matters, and numbered CGT 1 - 60.
shearer may claim deductions for jeans used as working clothes (TD 94/48), even though many people buy and wear jeans, and many shearsers would buy them as non-working clothes. By contrast, the clothes worn by plain-clothes police officers are not allowable deductions (TR 95/13). Flight attendants, required to be well-groomed on the job, cannot claim personal grooming expenses (TR 95/19) (as distinct from moisturisers, hair conditioners and the like, which are allowable deductions in recognition of the harsh working environment), while physical training instructors are able to claim the off-the-job costs of keeping fit, which is a necessary requirement for their occupation (TD 93/110).

Such arbitrary and inequitable outcomes arise because whereas it is easy to establish that a particular expense is associated with a particular occupation, it is difficult to establish that a group of people would not have incurred certain expenses if they had not had the particular occupation. For this reason, most of the Rulings and Determinations relate to the qualifying clause of s.51(1). The notable exception, which has generated a major anomaly, concerns deductions for self-education. To be a legitimate deduction, this expense must be related to a taxpayer’s current occupation and required for continuation or advancement in this occupation. But self-education expenses unrelated to the current occupation cannot be claimed as deductions, even if they will (or can) lead to a higher-paying occupation. As well, self-education expenses only in excess of $250 a year are deductible, even though no such restriction applies to other WREs.

Who Claims WRE Deductions?

High-income taxpayers and those in a small number of (generally highly-paid) occupations make much more use of the deductibility provision than others. High-income earners are big claimers of all types of WRE deductions, except for uniforms and laundry expenses. With this exception, both the proportions of wage and salary earners claiming, and the size of the average claim, increase with income for all categories of WREs. Regression analysis using 1993-94 ATO occupational data shows that the size of the average claim in each category, and the size of the average claim for all WREs, are more sensitive to income or salary than is the proportion claiming. An increase in annual salary of $1,000 leads on average to an increase of 0.4 per cent in the proportion claiming any WRE deductions. For individual WRE categories, the changes in proportions claiming range from minus 0.3 per cent per $1,000 increase in salary for ‘uniforms and laundry’ (the only category for which the proportion claiming decreases with income) to 0.9 per cent for the ‘other’ category. The size of the average WRE claim increases by $49 per $1,000 increase in salary, with the changes in the different categories ranging from zero for laundry and clothing expenses to $35 for ‘other’ expenses. Overall, the most income-sensitive claims are those in the ‘other’ category, which includes conferences,


\[4\] See FCT v Wilkinson (1983) ATC 4295.
books and journals, tools, and professional and union subscriptions. ‘Other travel’ and motor vehicle claims are the next most income-sensitive categories.

Table 1: Characteristics of WRE claims by five highest-claiming occupations

<table>
<thead>
<tr>
<th>Occ. code</th>
<th>Ave. salary ($)</th>
<th>Ave. income ($)</th>
<th>Motor (%)</th>
<th>Travel (%)</th>
<th>Clothing (%)</th>
<th>Education (%)</th>
<th>Other (%)</th>
<th>Total (%)</th>
<th>Ave. claim as % of salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>45,808</td>
<td>62,910</td>
<td>46</td>
<td>27</td>
<td>14</td>
<td>5</td>
<td>92</td>
<td>102</td>
<td>20.3</td>
</tr>
<tr>
<td></td>
<td>(4,367)</td>
<td>(4,822)</td>
<td>(192)</td>
<td>(1,461)</td>
<td>(8,179)</td>
<td>(10,776)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>32,440</td>
<td>40,193</td>
<td>70</td>
<td>12</td>
<td>16</td>
<td>7</td>
<td>89</td>
<td>118</td>
<td>14.4</td>
</tr>
<tr>
<td></td>
<td>(5,410)</td>
<td>(1,641)</td>
<td>(201)</td>
<td>(768)</td>
<td>(1,692)</td>
<td>(5,634)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>48,258</td>
<td>61,380</td>
<td>44</td>
<td>13</td>
<td>32</td>
<td>12</td>
<td>92</td>
<td>122</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>(2,455)</td>
<td>(1,412)</td>
<td>(179)</td>
<td>(1,437)</td>
<td>(2,121)</td>
<td>(3,448)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>53,092</td>
<td>56,063</td>
<td>22</td>
<td>28</td>
<td>69</td>
<td>7</td>
<td>95</td>
<td>115</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>(1,185)</td>
<td>(3,062)</td>
<td>(210)</td>
<td>(2,076)</td>
<td>(1,356)</td>
<td>(2,699)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>32,192</td>
<td>35,835</td>
<td>48</td>
<td>20</td>
<td>22</td>
<td>14</td>
<td>95</td>
<td>113</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>(922)</td>
<td>(2,077)</td>
<td>(207)</td>
<td>(1,179)</td>
<td>(1,269)</td>
<td>(2,246)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Occupational codes: 11: Legislators and government-appointed officials (including judges). 61: Investment salespersons. 23: Health diagnosis and treatment practitioners (medical practitioners, etc.). 33: Air and sea transport employees. 25: Other teachers and instructors (including lecturers). Since the average ($) claims in each category refer only to the amounts claimed by those claiming the relevant deduction, the average total amount claimed by those claiming any deductions is not the sum of the averages for the individual categories.

Source: Unpublished ATO data, 1993/94.

Table 1 shows that WRE tax deductions are extremely important for a small group of occupations, and are spread non-randomly around the employed population in terms of proportions claiming, average claim, and the ratio of average claim to income. If these WREs contain a large consumption component, then the groups listed in Table 1 are effectively receiving an income subsidy from the tax system. If, by contrast, the WREs are ‘purely productive’ and contain a zero consumption component, then the groups listed are being only partially protected by the tax system from necessary costs of employment. For example, a person on a 30 per cent tax rate required to purchase $100 of work-related items with zero consumption benefits receives only a 30 cents in the dollar cushion from the income tax against such expenses.

The degree to which WRE deductibility may affect the after-tax distribution of ‘true’ income is illustrated in Table 2. Deductibility is of the greatest benefit to high-income taxpayers in proportional terms. Although the ratio of claims to average wages and salaries decreases slightly as income increases, the effects of increasing marginal tax rates in the upper-income ranges, which make WREs more valuable to taxpayers, tends to offset this. However, if all WREs claimed are ‘phantom’ WREs (those that they are not required to perform a job and do not increase productivity) with a 100 per cent consumption component, the impact of deductibility is broadly neutral, though low- and medium-income taxpayers gain least in proportional terms,
Abolishing Income Tax Deductions for Work-Related Expenses

and high-income taxpayers the most. If WRE claims relate to ‘purely productive’
items of expenditure, then, in the absence of labour market adjustments which
would compensate employees for particular expenses of work, deductibility clearly
protects those on low incomes. To some degree, though, it provides even more
‘protection’ to high-income taxpayers. To remove deductibility would in the short
run make the income tax slightly more progressive.

Table 2: WRE tax subsidies and income

<table>
<thead>
<tr>
<th>Taxable income ($)</th>
<th>Ave. salary ($)</th>
<th>Ave. WRE claim ($)</th>
<th>Ave. WRE as % of salary</th>
<th>Ave. tax benefit ($)</th>
<th>Tax benefit as % of salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 10,000</td>
<td>8,063</td>
<td>344</td>
<td>4.3</td>
<td>69</td>
<td>0.9</td>
</tr>
<tr>
<td>10,000-14,999</td>
<td>11,824</td>
<td>386</td>
<td>3.3</td>
<td>77</td>
<td>0.7</td>
</tr>
<tr>
<td>15,000-19,999</td>
<td>16,946</td>
<td>492</td>
<td>2.9</td>
<td>98</td>
<td>0.6</td>
</tr>
<tr>
<td>20,000-24,999</td>
<td>21,957</td>
<td>552</td>
<td>2.5</td>
<td>196</td>
<td>0.9</td>
</tr>
<tr>
<td>25,000-34,999</td>
<td>29,282</td>
<td>741</td>
<td>2.5</td>
<td>263</td>
<td>0.9</td>
</tr>
<tr>
<td>35,000-49,999</td>
<td>40,354</td>
<td>1,033</td>
<td>2.6</td>
<td>456</td>
<td>1.1</td>
</tr>
<tr>
<td>50,000+</td>
<td>66,861</td>
<td>1,486</td>
<td>2.2</td>
<td>698</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Notes: Average Wages and Salaries are averages for each taxable income group. Average WRE claim is the average for those who claim. Tax benefit is computed by applying the relevant marginal tax rate to the average claim for each group.

However, given that (as shown below) labour markets adjust in the long run to
compensate employees for particular expenses of employment, deductibility pro-
vides short-run protection only; in the long run, the impact would largely be neutral
when WREs are purely productive. The shares of the subsidy and protection com-
ponents of WRE deductions are unknown, but this does not affect the equity argu-
ment for abolition. If deductions subsidise consumption, then, regardless of
whether the subsidy is progressive or regressive, they should be abolished and the
progressivity of the tax system should be made transparent through the rate struc-
ture itself. If WREs are purely productive, then the protection afforded by de-
ductibility is not needed in the long run.

The Equity Impact

In the apparent absence of any explicit justification for tax deductibility of WREs in
any of the parliamentary debates leading up to the first ITAA in 1915 or in any sub-
sequent debates or official publications, it is implicitly justified by reference to eq-
uity considerations.

Suppose occupations A and B each pay a gross wage of $500 a week, but B re-
quires the employee to purchase items costing $100 for use on the job as a condi-
tion of employment. Initially, suppose too that the WRE items are purely produc-
tive and have zero value to the employee outside the workplace. They are essen-
tially just productive inputs which, for whatever reason, are supplied by the em-
ployee rather than the employer. Finally, assume that there is a fixed income tax
rate of 30 per cent.
Without deductibility, employees in each occupation have $350 a week after tax, but the B-employee is left with only $250 after paying for the required WREs. If WREs are tax-deductible, the position of the A-employee is unchanged. However, the B-employee’s taxable income is now only $400, on which tax of $120 is payable. Take-home pay is thus $380, of which $100 is needed for WREs. This leaves the B-employee with $280, which is still less than the A-employee’s $350. Nevertheless, the broad principle of ‘vertical equity’ is satisfied. At its most general, this principle states that tax contributions should be related to ability to pay, as measured by gross income. The true gross incomes of the two employees are $500 for A and $400 for B, after subtracting necessary WREs. Since A pays more tax than B, vertical equity is satisfied. Without deductibility, the principle is violated, since each pays the same amount of tax even though A’s true gross income is higher after allowing for B’s expenditure on purely productive WRE items.

Deductibility of WREs is also consistent with the principle of horizontal equity, which holds that people with the same capacity to pay as measured by gross income should pay the same taxes. Consider a person in occupation C, with a pre-tax income of $600 a week, out of which $100 is required for necessary WREs which generate no consumption benefits. True gross income of a C-employee is $500. With deductibility, this employee pays tax of $150, the same as an A-employee with pre-tax income of $500 and no WREs. Without deductibility, the C-employee pays $180, which violates the principle of horizontal equity since A and C have the same true gross incomes.

The equity rationale of WRE deductibility is superficially plausible. But it assumes that there are no consumption benefits derived from WRE items and that wages paid do not adjust to the relative costs or benefits associated with tax-deductible WREs. Consider the question of consumption benefits. Most WRE items can, in practice, generate (often significant) private benefits for those incurring the expenditure; and if the WREs are tax deductible, private consumption is effectively subsidised. A building worker’s tools can be used for non-market and non-taxable activities, such as home renovations. Work-related travel costs can be used to finance leisure travel. Books purchased for professional purposes can, depending on the nature of the profession, generate significant consumption benefits: a teacher of English literature could have a ‘professional library’ consisting of books which most of us would buy purely for leisure reading. True, the extent to which deductibility can be used to subsidise private consumption is limited by legal judgments and Rulings and Determinations, as well as the ATO’s procedures and rules of thumb and the apportionment provision in s.51(1), which requires that particular expenses are apportioned between work-related expenses and those which are private in nature. But varying levels of subsidy are an inherent feature of such open-ended opportunities, especially if procedures designed to prevent abuse and to clarify the employee’s rights and duties are imperfect in their effects.

Such subsidies have some adverse efficiency implications (considered below). To the extent that their impact is not moderated by market forces, they also have clear and effectively arbitrary equity implications. As well, the big-ticket items, like
motor vehicles, 'other travel', and 'other' WREs, are concentrated among the higher-income groups, and so make the income tax effectively less progressive.

These observations assume that labour markets do not adjust for the differential consumption benefits, or indeed, where these are zero or are outweighed by the non-consumption characteristics of WRE items, for the differential costs associated with the WRE requirements of particular occupations. However, depending on the flexibility of labour markets and the ease of entry into particular occupations, some adjustment could reduce the arbitrary equity impact of deductibility. If the WRE items required by a particular occupation do not generate any consumption benefits, workers entering that occupation will require compensation for the additional net-of-tax costs associated with that occupation. If, for example, $100 of WREs with no consumption benefits are required, the supply price of labour to that occupation will increase by $100 by comparison with similar occupations with no WRE requirements: with a tax rate of 30 per cent and deductibility of WREs, a wage increase of $100 will give a worker a net wage increase of $100 if this is matched by deductible WREs of $100; and after this amount is spent on WRE items, there is no change in the amount available for private consumption. If the supply of labour to the occupation were perfectly elastic, the gross wage would rise by $100 in this case. But in general, the outcome would be a rise of less than $100 and a decrease in employment in that occupation: the workers who remain in the occupation would be less well off than they would have been without the WREs-plus-deductibility, but they would not be $70 worse off, which would be the case if there were no labour market adjustment.

At the other extreme, where all WREs are 'phantom', WRE deductibility is one of the non-wage benefits of the job. Depending on how much they value these benefits, workers will be willing to work in this occupation for a lower wage than for comparable occupations without access to the WRE consumption subsidy. The supply price of labour to this occupation will be lowered by comparison with similar occupations if the labour market is at all flexible, and to this extent the consumption subsidy is partly offset by a lower wage. Employer-provided on-the-job benefits have a similar impact.

To summarise: while the arbitrary equity impact of WRE deductibility may be to some extent moderated by market forces, neutralisation is not perfect. Deductibility to some extent protects workers from incurring 'true' WREs from their costs, but not perfectly. At the same time, where there is a large consumption component, market reactions will to a degree moderate the benefits of the de facto consumption subsidy, but not perfectly.

Of themselves, these arguments do not constitute an overwhelming case for the abolition of WRE deductibility on equity grounds, because the position of workers required to incur 'pure' WREs which generate no consumption benefits would arguably be worse without deductibility, even if abolition of deductibility were to eliminate the consumption subsidy going to certain occupations. This would be true if workers incurring pure WREs were to continue shouldering this expense after abolition (though if this were the case, market forces could again be expected
to moderate the impact. However, it is likely that abolition of WRE deductibility would shift the cost of these items from employee to employer.

**Shifting the Cost of WREs to Employers**

'Pure' WRE items are simply productive inputs which happen to be supplied by employees rather than employers. For such inputs to be supplied by employees, the relevant items must be comparatively cheap, because an employee would otherwise be unlikely to be able to afford them. With WRE deductibility for employees, it is a matter of indifference in the long run (allowing for labour market adjustments) to both employer and employees who shoulders the cost.

Suppose the assumptions of the earlier example hold (WREs of $100, with a 30 per cent employee tax rate) and, in addition, the tax rate on the employer's profit is 40 per cent. If the cost of the WRE items is shifted from employee to employer, then the employer's net profit is reduced by $60 per employee, which is the net cost of the employer's increased expenditure on WRE items. However, holding employment constant, the gross wage will fall by $100, reflecting the elimination of the employee's WRE cost, and this will increase net profit by $60. The overall impact is that the net positions of both employer and employee are unchanged.

Why then do employees bear the cost — as WREs — of some productive inputs? One possible explanation is that use of the relevant inputs is difficult to control, and that the employer saves on monitoring costs by letting the employee bear the cost of the WRE items. A more likely explanation is that deductibility of WREs generates some consumption benefits, and that letting the employees bear the burden of deductible WREs is a way of capturing a wage subsidy. Suppose that while $100 of WREs are required for 'purely productive' purposes, access to deductibility also allows the employee to receive a tax subsidy by purchasing in excess of this, because the operation of the tax law in practice does not allow for determination of what is 'really' required for the job. If these consumption benefits are worth $300 (as measured by the increased consumer surplus accruing to the employee), then the net impact of deductibility is a benefit to the employee of $230, which is the consumption benefit minus the net cost of the purely productive WREs. If the cost of the WRE items were shifted to the employer, the employer would (presumably) pay only for the 'purely productive' items ($100 gross), which reduces net profit by $60 per worker. If employees retain their access to deductions for the alleged WREs they purchase for consumption benefits, their wage would fall by $100; they would be no worse off (because they do not now bear the $100 of purely productive WREs), and this wage reduction would offset the increased expenditure by the employer on WRE items. However, if employer responsibility for these items means that the deductions formerly available to the employees (the effective tax subsidy for consumption) are withdrawn, employees will require a wage increase of $428, which is the gross value of the consumption subsidy forgone. Hence, overall, employees would receive a wage increase of $328 to get them back to their original position, while the employer would incur increased costs of $328 per employee, which is a reduction of $197 in net profit.
While market adjustments would generally ensure that the final outcome would be a loss shared between employer and employees, this example shows that deductibility of WREs for employees is attractive to both. So long as there are any consumption benefits associated with tax deductibility of WREs for employees, this provides an effective wage subsidy, and any wage subsidy, whether implicit or explicit, is attractive to both sides. But what is important in this case is that this type of subsidy (WRE deductibility) provides an incentive for the employer to leave the responsibility for supplying particular types of productive inputs — those which are deductible for employees and provide consumption benefits — in the hands of employees. If deductibility were abolished, the incentive would largely disappear.

Implications for Allocative Efficiency

A slight modification to the previous example shows that deductibility may have adverse efficiency implications. The simple modification required is to recognise that the amounts of productive WRE-type inputs used by an employer are affected by their net cost.

In a world with no tax-deductibility of WRE items for employees, though such items are de facto tax-deductible for employers, it is a matter of indifference (for employers and employees) as to who bears the cost. This was demonstrated in an earlier example. It was also demonstrated there that if employee finance of WRE items is accompanied by employee tax deductibility of WREs, and if access to deductibility allows employees to benefit from an effective consumption subsidy, then there are net joint gains to be made by employers and employees from shifting the costs of WRE items to employees. The division of the net gains would depend on the nature of the labour market (and in particular on the elasticities of supply and demand for labour), but in general both employees and employer would expect to share in the spoils.

So long as employers share in these gains from WRE deductibility, the effective net price to the employer of the WRE items required for purely productive purposes will be lowered, as will the effective price of hiring an employee. In total, the price of a 'package' consisting of an employee plus the productive items required by that employee will be lower than otherwise, and the employer will hire more such packages than otherwise. (Strictly speaking, this argument assumes that there is no substitutability between workers and WRE items in production. If this is not true, the composition of the chosen 'package' will vary depending on the productivity of the two component of the package. However, the essentials of the argument are not affected.)

It is a standard proposition in economics that if a profit-maximising enterprise pays the full price for an input and sells its output in a competitive market, that input will be used at the efficient level as defined above. An input which is subsidised will generally be over-used, and a taxed input under-used. In the absence of the subsidy provided by tax deductibility of WREs, items which generate consumption benefits for employees, labour and the WRE inputs will be used at their efficient levels. With the subsidy, they will be over-used. This is the basis for the allocative
inefficiency argument. The cost of these inefficiencies is a matter for conjecture, but is probably small. But in the present context they provide a further reason for abolishing WRE deductions for employees.

However, a further efficiency cost should be taken into account. Insofar as employees purchase some WRE items purely for the consumption benefits they generate, and since tax deductibility for these items lowers their effective cost to employees, these items will be consumed more than otherwise. This is a distortion in that the consumption value of these items to the relevant occupational groups is, at the margin, lower than the value to other groups, and lower than the marginal cost of production. Again, the magnitude of the costs involved is conjectural, but the argument for removing the tax subsidy on these grounds is similar to the argument for moving towards a general consumption tax as a means of correcting a distorted consumption pattern.

**Compliance Costs**

The compliance costs of tax deductibility of WREs alone provide a compelling case for abolishing them. Such costs comprise the time, money and effort incurred by taxpayers in particular, and by the private sector in general, in attempting to satisfy their tax obligations. The only comprehensive study of income tax compliance costs in Australia was that carried out in income year 1986/87 by Pope, Fayle and Duncanson (PFD) (1990), who used the sample survey method employed by a number of international researchers. This technique has been criticised on a number of grounds (Rimmer & Wilson, 1996). Most important, it may improperly identify costs incurred for other reasons as compliance costs, and the often low response rate (only 16.3 per cent in the PFD study) generates doubts about the randomness of the effective sample, since self-selection by those taxpayers who find compliance particularly irksome and costly is quite likely. These factors could exaggerate their estimates, perhaps by as much as 50 per cent. This possibility is allowed for in what follows.

On the other side of the ledger, since the PFD study the personal income tax system in Australia has undergone many radical changes, the introduction of self-assessment and *TaxPack* (in 1987) being the most important. These changes are likely to have had a considerable and positive impact on compliance costs, especially for wage and salary earners. Some indication of this is given by the increase in tax agent use associated with self-assessment and *TaxPack* (Baldry & McKinstry, 1995). To an unknown extent, these changes are likely to have partially offset any over-estimates in the PFD study.

The PFD study estimated personal income taxpayer compliance costs at 10.4 per cent of income tax revenue. A rough estimate of the amount which would be saved by abolition of WRE deductibility can be obtained by using those PFD estimates which relate to the type of tax return form completed.

In 1986/87, salary and wage-earners were required to complete and lodge Form S, so long as dividend and interest income was moderate, and so long as income was not earned from other major sources such as trusts, self-employment and rental
property. In short, Form S was the form which most wage and salary earners without other significant income would lodge. Form S-lodging taxpayers would presumably find their affairs considerably simplified if they could not claim WREs as deductions. The only major exceptions would be those taxpayers whose income was largely derived from pensions and unemployment, sickness or other benefits. An extreme assumption is that compliance costs for all Form S taxpayers would be reduced to zero by abolition of WRE deductions. A more moderate assumption is that these costs would be reduced by 50 per cent. Using the PFD data, rough calculations following from these assumptions suggest a reduction in compliance costs from 10.4 per cent to either 7.98 or 9.2 per cent of revenue, depending on the assumption made.

Using the data given in PFD (Table 5.5), the savings for low-income taxpayers would be between 2.3 and 1.15 per cent of taxable income, and, for high-income taxpayers, between 0.45 and 0.9 per cent of taxable income. Since about half of the total compliance costs for the average taxpayer comprise time costs, abolition of WRE deductions would not generally lead to massive visible savings for wage and salary earners, but rather a saving in time, frustration and worry, heavily concentrated around the end of the tax year. However, there would be visible monetary savings as well, especially in tax agents' fees, for many taxpayers. Possibly 30 per cent or more of taxpayers would cease to use the services of tax agents. Currently about 72 per cent of taxpayers use agents, though the proportion for wage and salary earners is probably much lower. Since data on agents' fees paid are not available, the tax agent savings cannot be estimated. However, the crude PFD data suggest that a 50 per cent reduction in compliance costs for salary and wage earners, where 50 per cent of the compliance costs saved are monetary, would lead to an estimated total monetary saving of 1.9 per cent of income tax revenue. If this figure were to be applied to current (1993/94) statistics, it would imply a saving of $905m, or $119 for every taxable person.

These figures indicate that the gains available from abolition are not massive, but neither are they negligible. Moreover, other savings would result. Most important, the considerable effort which the ATO currently puts into compliance enforcement in respect of WRE claims could be diverted to other, more lucrative areas of enforcement.

Table 5.4 of PFD presents the average estimated compliance costs of Form S taxpayers for each of four taxable income groups, and Table 5.5 lists the numbers of taxpayers in each income group. 64.8 per cent of taxpayers completed form S (Table 3.4). The assumptions made here are, first, that each of the form S taxpayers in each group would save either all or 50 per cent of the average compliance costs listed in Table 5.4; and second, that 64.8 per cent of the taxpayers in each income group completed form S (the numbers of form S taxpayers in each group are estimated as 64.8 per cent of the total number of taxpayers in that group). Total savings are estimated as the average, or 50 per cent of the average, compliance costs for each group, multiplied by the estimated number of form S taxpayers in that group. Total savings are estimated at $859m or $430m. These figures are subtracted from the total compliance costs as estimated by PFD, and the result expressed as a percentage of the Net Tax, as reported in Table 2 of the 1986/87 Taxation Statistics.
On the other hand, abolition of deductibility would provide more incentives for employees to become self-employed. The aim is to access the deductible business expenses which can provide (subsidised) consumption benefits. Employers have a similar incentive to do this, being able to economise on wages and avoid many of the other costs and regulations associated with employment (including the requirement to deduct tax at source under PAYE). The extent to which this would happen is unknown, given that self-employment involves a loss of benefits for employees and a less secure contract for employers. However, it is worth noting that the ATO has identified the self-employed as a high-risk area in terms of tax revenue, and is in the process of implementing procedures to limit the access of the self-employed to business expenses. In short, the ATO has the means of limiting the extent of this leakage of revenue, and shedding the enforcement costs associated with WRE deductions would release resources to tackle this area of revenue risk.

The Real Challenge of Tax Simplification

Abolition of WRE tax deductions would be a comparatively simple yet real structural reform to the income tax, and would generate net gains for the economy. It will cause losses to some groups in the economy, though less so in the long term (as labour markets adjust) than in the short term. Some employers are likely to oppose it as strongly as employees, because they would foresee the resultant pressures for wage increases and the need to bear some of the costs previously borne by employees. Other groups, especially those producing the over-used, tax subsidised items, would also object. However, unless the arguments presented here are refuted and a positive case made for the retention of deductibility, this provision of the tax law must be seen merely as historical baggage which remains a feature of the tax system simply because of inertia.

References


I am grateful to two anonymous referees for many helpful comments on an earlier version of this article. I am grateful also to John Gudgeon of the ATO for assistance in obtaining information on the occupational distribution of WRE claims which was kindly supplied by the ATO from its 1994 tax-return database while I was in residence during February 1996. Ludmilla Ruvinsky assisted with the statistical processing of the data.
What Has Become of the Japanese Model?

Luke Gower

No nation projects images of a distinctive system of political economy in quite the way that Japan does. During the 1970s and 1980s, many observers became convinced that Japan's decades of consistently high economic growth were due to the unusual institutions of its capitalism. These idiosyncrasies were thought to constitute a 'Japanese model', from which policy-makers in other countries could learn lessons in economic management.

A review of that idea is now in order. Japan's GDP grew by only 2.1 per cent between 1992 and 1995, compared with an average of 6.1 per cent in the other members of the Group of Seven. Unemployment at 3.4 per cent is near a record high, and the financial system is in disarray. These problems are likely to persist. For although growth in 1996 was firm, the plunge in GDP in the second quarter of 1997 was the largest in 23 years. Moreover, the damage that this has done to economic recovery has been seriously accentuated by a crisis in the financial system. The effects of persistently weak productivity improvements will soon be compounded by the retirement of a large cohort of workers. Labour inputs and aggregate savings ratios are therefore likely to dwindle, taking long-run growth down with them.

These vicissitudes of economic maturity and the measures that will be required to cope with them have disillusioned many erstwhile believers in the Japanese model. Japanese policy-makers in particular are now less confident in their interventionist approach, and show correspondingly greater interest in market-oriented reform. This article explains how the change in thinking has come about.

Defining the Japanese Model

Two broad perspectives have dominated debate about Japan's high growth rate in the post-war era: the growth-accounting approach and the Japanese model. Proponents of the former perspective, which has been elevated to prominence in an East Asian context by Paul Krugman (1994), decompose economic growth into flows of resources into productive sectors of the economy and the improvement of technology. Although growth-accounting theorists disagree about the relative importance of these two factors, most would agree that they jointly explain Japan's economic growth in the 20th century. Most would also agree that, because Japan has almost exhausted the potential for any further exceptional expansion from a rapidly in-

Luke Gower is Kiyoshi Kojima Fellow and Associate Director of the Australia-Japan Research Centre at The Australian National University.
creasing resource base or from reallocating existing resources, the difference between Japanese growth rates and those in the rest of the developed world is likely to diminish. Some would add that the contraction of this growth differential is all the more certain given that Japan is approaching technological parity with the countries that it once strove to emulate (Saxonhouse, 1994). In short, growth-accounting theorists think that tightening resource and technology constraints will overshadow whatever positive influence domestic economic institutions may have on Japan’s potential growth rate.

Proponents of the Japanese model find this reasoning too pessimistic and mechanical. They argue that Japan is a special case, with indigenous political and economic institutions that contribute to economic growth in ways not recognised by growth accounting (Johnson, 1982). In particular, they infer a causal link between the rapid rates of Japanese economic growth after World War II (see Figure 1) and the distinctive institutions of Japanese capitalism which were emerging at the same time. Since many of those institutions are now fixtures of the Japanese economy, proponents of the model often doubt that Japanese per capita growth must necessarily decline over the long run.

**Figure 1: Real per capita income percentage growth rates in Japan and the United States, 1901-94**

![Graph showing real per capita income percentage growth rates in Japan and the United States, 1901-94.](image)

Note: The observations for 1945 and 1946 were outliers and have been suppressed. Source: Maddison (1995).
Among proponents of the model there has been only a limited consensus on the question of which institutions have been most responsible for economic growth. Many mainstream economists have drawn attention to the more unusual properties of Japanese corporate governance and industrial organisation, which they think are well suited to mobilising resources. But another influential group has contested that emphasis, arguing that the Japanese state has used industrial policy with extraordinary dexterity, and that it has successfully identified and supported those industries which were essential to growth. Although the Ministry of International Trade and Industry (MITI) is the most conspicuous agent of these policies, other government agencies, such as the Bank of Japan and the Ministry of Finance, have also played key roles in the planning of economic activity.

For at least two reasons, it is those versions of the Japanese model which emphasise the importance of state planning and regulation that warrant the closest scrutiny. First, they are better known and more influential than most of the industrial organisation theories. The state's planning role in post-war Japan is widely appreciated and admired, particularly among policy-makers from other (commonly East Asian) countries who seek to replicate Japanese economic history. Second, the state's ability to deliver economic growth by means of economic planning and regulation has become increasingly contentious. The government's failure to stimulate recovery in the course of the current slowdown seriously undermines many of the claims of its supporters, and the threat of slower growth over the long term raises strong arguments against a regulated and managed economic system.

The Recent Slowdown

The Japanese model suffered its first serious setback in the early 1990s, when the rate of economic growth suddenly fell. International competitiveness dropped as the real exchange rate appreciated; the financial system was destabilised by the collapse of asset prices; and conditions in the labour market deteriorated sharply.

Although the causes of the downturn are still debated, monetary factors and speculative behaviour feature prominently in most explanations. During the late 1980s, stable consumer prices and an appreciating exchange rate produced an accommodating monetary policy. In contrast to stable consumer price inflation, and in response to the loose monetary policy, asset prices rose substantially. Between January 1987 and December 1989, stock prices on the First Section of the Tokyo Stock Exchange doubled. Over a similar period, the price of urban land in the six major cities did the same. The eventual correction in asset prices was a substantial one: stock prices fell to 40 per cent of their peak level and, after several years of slow recovery, they have not risen far above that point. At the same time, urban land values have collapsed by 22 per cent since 1990 and they are yet to reach their nadir.

1 See Aoki (1990) for a survey.
2 Johnson (1982) is the classic contribution.
The problems of unstable asset prices were exacerbated by financial institutions, which mistook the temporary hike in asset values for a real and permanent shock. On the assumption that asset prices would remain high, financial institutions extended credit on a massive scale to borrowers who offered land as collateral. When the correction in land prices occurred and borrowers defaulted, lenders were left with low-value collateral and bad loans which, even now, are conservatively estimated to be of the order of ¥27 trillion (A$314 billion).

The recent crisis in financial markets underscores the severity of these problematic bad loans and it is an embarrassment to admirers of the Japanese model. Previous financial crises — most notably that of 1965 — were dealt with decisively by government and business. The inability of those same parties to deliver similar outcomes in the 1990s is a poor reflection on contemporary Japanese economic management.

The financial crisis has had severe consequences for the personal and non-financial corporate sectors. And its recent deepening does not augur well for the future. Those who had borrowed against the value of land have been driven into bankruptcy in large numbers (Figure 2), and banks have become much more cautious in their lending policies, particularly where individuals and small businesses are concerned (APEG, 1997:50). This is highly problematic for a country like Japan, which relies heavily on small business for its output and employment.

**Figure 2: Liabilities generated by business failures, 1990q1 - 1997q1 (¥bn)**

A further problem has been the acute drop in international competitiveness. Between 1990 and 1995, the yen appreciated by 25 per cent in real terms. Initially, falling wealth led to a decline in imports, and the current-account surplus expanded (in yen terms) by 42 per cent between 1991 and 1994. The nominal value of the
currency was marked up accordingly. Some modest deflation subsequently offset the reduction in competitiveness that the nominal appreciation had induced. However, the transmission of nominal exchange rate impulses to domestic prices in Japan is notoriously unreliable (Marston, 1990), and deflation in the tradable goods sector was not sufficient to restore competitiveness. It has only been in the last few quarters that depreciation of the currency has allowed exports to contribute substantially to growth.

Table 1: Real and nominal exchange rates, 1990-96

<table>
<thead>
<tr>
<th>Year</th>
<th>Real exchange rate (^a)</th>
<th>Nominal exchange rate (^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>100</td>
<td>144.8</td>
</tr>
<tr>
<td>1991</td>
<td>95</td>
<td>134.7</td>
</tr>
<tr>
<td>1992</td>
<td>92</td>
<td>126.7</td>
</tr>
<tr>
<td>1993</td>
<td>81</td>
<td>111.2</td>
</tr>
<tr>
<td>1994</td>
<td>77</td>
<td>102.2</td>
</tr>
<tr>
<td>1995</td>
<td>75</td>
<td>94.1</td>
</tr>
<tr>
<td>1996</td>
<td>89</td>
<td>107.7</td>
</tr>
</tbody>
</table>

Notes: \(^a\) Figures are expressed in index form. \(^b\) Yen per US dollar. Source: APEG (1997).

Figure 3: Growth rates in two post-war Japanese slowdowns

Notes: The graph plots the year-on-year real growth rate beginning with the point at which a quarter-on-quarter fall in the flow of GDP was first recorded. By this definition, the starting points for the two slowdowns are the first quarter of 1974 and the second quarter of 1992. The Economic Planning Agency uses a more complicated formula for gauging the business cycle; its estimates suggest that the most recent slowdown ran from February 1991 to October 1993. Many private commentators believe that the actual slowdown persisted beyond that endpoint. Sources: Bank of Japan, Economic Statistics Annual, various issues; Nikkei Telecom.

As prospects for the corporate sector worsened, so did conditions in the labour market. At 3.4 per cent, the official unemployment rate may seem low, but it is underestimated since labour market practices in Japan mean that workers can be idle
or virtually unemployed without appearing to be so in the official statistics. According to McCormack (1997), if US Bureau of Labor Standards criteria were used in Japan, the rate of unemployment would be closer to 8.9 per cent. A slack labour market is also suggested by the more closely watched ratio of job offers to seekers, which has a long-run average of 0.88 but is currently at 0.75, and has recently been as low as 0.6.

Perhaps the most perplexing aspect of this slowdown is its persistence. This is not only the deepest, but also the longest, slump in post-war Japanese history. The only post-war slowdown of comparable magnitude occurred around the time of the first oil shock in 1973. As Figure 3 shows, growth rebounded much more quickly on that occasion. For that reason, and the fact that the Japanese business cycle is not synchronous with sustained growth in the rest of the industrialised world, many Japanese commentators are beginning to doubt the flexibility of their economy and the virtues of the model on which it is supposedly founded.

Implications for the Model and the Prospects for Reform

In raising concerns about the inflexibility of the economy, the 1990s slowdown draws attention to the longer-term limits to Japanese growth. Japan’s trend rate of growth has actually been easing for some time, and for reasons that are quite independent of the 1990s malaise. The data summarised in Figure 1 reveal that, after averaging 8.4 per cent between 1950 and 1970, average annual income growth per head slumped to 4 per cent in the 1970s, and then 3.4 per cent in the 1980s. Most proponents of the Japanese model missed this shift of gear. It was not until the 1970s that the epithet ‘Japan Inc.’ gained currency; it was not until 1979 that Ezra Vogel (1979) foretold the arrival of ‘Japan as Number One’; and, as Saxonhouse (1994) notes, it was not until the mid-1970s that a clutch of prominent forecasters with faith in the model forecast 10 per cent annual Japanese growth rates for the better part of 20 years.3

By the time these prophecies had been made, the Japanese economy had indeed become large. But, as growth-accounting theory predicted, by the 1970s and 1980s it was approaching its limits. The manufacturing sector had long been acquiring under-utilised resources from agriculture. It had also been generating productivity growth by absorbing new technology from abroad and it had been rebuilding the nation’s war-damaged industrial base with the most modern forms of capital. By the 1980s, the potential for any further such gains had been largely spent. As a result, Japan’s ability to achieve exceptional growth rates was considerably diminished.

Now Japan faces an additional set of constraints. Foremost among them is the country’s demographic profile. Japan is one of the most rapidly ageing industrialised economies: its working-age population will decline by a net 4.6m workers in the first two decades of the 21st century. By 2015 one-quarter of the population will be

---

3 The Economist deserves credit for having noticed Japanese growth as early as 1963, when it published its series of essays on the subject.
aged over 65, and by 2049 that proportion will have risen to one-third (see Figure 4).

The most obvious problem associated with an aging population is a tightening labour supply constraint. To deal with this problem, the Japanese parliament is currently debating reforms to the tax system which will improve the access of women to the workforce, and it has already strengthened the provisions of long-standing, but previously unenforced, equal employment opportunity legislation governing overtime. In order to encourage greater workforce participation among the elderly, social security reform is also being debated.

**Figure 4: Age composition of Japan, 1995 and 2005**

![Age composition of Japan, 1995 and 2005](image)


Besides facing a diminishing labour supply, Japan is likely to experience lower savings rates as a result of its aging population. Aging implies a rising dependency ratio, which will inflate expenditure on social security for the elderly and ultimately increase the tax burden. Furthermore, since the elderly often have a high average propensity to consume, the rising proportion of them in the population will put downward pressure on aggregate savings rates. To the extent that international capital mobility is imperfect, both of these considerations imply that Japan will have a smaller pool of savings on which to draw for future investment.4

If history offers any guidance, the productivity slowdown of recent years and the diminishing pool of savings are likely to matter more than the tightening of the la-

---

4 An alternative hypothesis, advanced by Hayashi (1986), is that the Japanese savings rate has been high because the Japanese bequeath wealth to their children. As economic growth slows, this incentive should weaken and the rate of saving should fall.
bour supply constraint. Growth accounting studies usually indicate that Japan's growth has been driven more by productivity and capital accumulation than by rising labour inputs (Kim & Lau, 1994; Drysdale & Huang, 1997; Denison & Chung, 1976).

Liberal financial markets are increasingly being seen as a way of managing the problem. Although financial deregulation has been in progress for about 17 years, it was radically accelerated in November 1996 when Prime Minister Hashimoto announced a major overhaul of the financial system. This program was simultaneously a response to the belief that the existing regulated financial structure was performing badly and an explicit response to the long-run problems presented by aging (Economic Council, 1996). The program foreshadows further deregulation of the securities industry, changes to the Bank of Japan Law, more deregulation of the system under which foreign-exchange earnings are currently repatriated, and extensive reforms of the life insurance and funds management industries.

These initiatives go against the grain of the Japanese model, every variant of which endorses financial repression. The basic argument is that the Japanese state successfully set a ceiling on interest rates on private savings accounts over much of the post-war period, in order to minimise the cost of capital to industry and so promote growth (Zysman, 1983:248). The policy was effective partly because savings were supplied inelastically and partly because favourable tax treatment of them ensured reasonable effective rates of return (Hamada & Horiuchi, 1987). If this was ever a good policy framework, it is certainly not appropriate now. The retirement of the current cohort of workers will have to be funded in some way. If social security and taxes are not to be the solution, then it is essential that the returns on private savings (including pension contributions) be maximised. That is why more efficient and competitive financial markets are currently being promoted.

The transition to a less regulated financial system is certain to deprive the state of some very important tools in economic planning. Already, financial deregulation has deprived the monetary authorities of their capacity to promote favoured industries through the application of 'window guidance'. This form of suasion was once an important means by which the authorities encouraged banks to lend to certain government-targeted industries, and it was exercised when large private banks borrowed from the Bank of Japan's discount window (a course of action to which they were regularly compelled when the financial system was tightly controlled). The practice has now fallen into disuse simply because, in liberal capital markets, banks no longer depend on the discount facility, and the leverage of the monetary authorities over the banks in relation to the direction of bank lending has lessened.

There is also the possibility that the state will eventually forfeit much of its control over the direction of capital as a result of the unravelling of the extensive system

---

5 History may not be our most reliable guide. International capital mobility is now much higher than it was in the earlier post-war years, and domestic savings ratios may therefore prove less essential to Japanese capital accumulation in the future.

6 Takeda and Turner (1992) provide an excellent summary of the process in its earlier phases.
What Has Become of the Japanese Model?

of state financial institutions. To date, regulation of the financial system has been configured to garner for the state a high proportion of personal savings: approximately 20 per cent of household savings are held through the state-run postal savings system — making it the largest system of socialised savings in the world. These funds are channelled to industries and activities which the state favours through a large network of state-operated financial institutions. Recently, the uses to which these funds are being put have been fiercely criticised from within the financial community and even from within the bureaucracy. The Minister for Health and Welfare has called for the privatisation of the postal savings system, publicly complaining that the Ministry of Finance does not earn enough on its investments. In large measure, the system performs poorly because the funds that it accumulates are channelled into industries in which Japan has a gross comparative disadvantage and which generate low rates of return (Calder, 1993:107).

This inability on the part of state-operated financial institutions to support high-growth sectors epitomises the more general tendency of the state to pursue industrial policies that retard, rather than facilitate, structural adjustment. Although industrial policy has had many successes in Japan, there are many more cases of ministries having failed to discern which industries constitute the essence of Japan's comparative advantage. The failure of key sections of the bureaucracy to support the automotive and consumer electronics industries in the early post-war period are cases in point. At the same time, industrial policy has vigorously supported industries which have very limited growth potential; the most notorious examples are the provision of subsidised state finance to mining and agriculture.

Although recent governments have defended the role of state-run financial institutions in the face of mounting opposition, there is evidence that the tide of opinion is moving against conservative interests. Prime Minister Hashimoto has conceded that the public sector ought not to be exempt from the financial reform process (The Nikkei Weekly, 27 January 1997). More tellingly, policy circles have recognised that open capital markets could assume a greater role in influencing the course of industrial development. Reforms undertaken so far include the easing of eligibility requirements for listing on stock exchanges, the creation of new over-the-counter markets for firms with a strong research and development orientation, and deregulation of the eligibility requirements for the issuance of bonds. In addition, one of the purposes of the reform to Article 9 of the Anti-Monopoly Law is to provide large firms with incentives to form small subsidiary corporations that undertake research and development.

The government's change of attitude towards industrial policy is also reflected in its formal planning process. Government plans used to identify those sectors which were to be promoted with public resources and regulation. More recently, they have acknowledged the ability of markets to identify and promote promising industries. The most recent report of the Economic Planning Agency (EPA) has earmarked few industries for support and it has stressed the importance of sweeping deregulation. Significantly, even where it did endorse the government's role in promoting certain industries, the plan identified deregulation of the operating envi-
ronnents of those industries as a priority. Most important, the announcement of a financial reform package in November 1996 has been followed with an announcement that the Hashimoto cabinet has adopted 2,823 individual deregulatory measures governing areas including energy, welfare, transport, telecommunications and real estate.

Private-sector lobby groups have endorsed deregulation. The peak Japanese employer body, Keidanren, has been actively demanding it and has insisted that the real appreciation of the yen needs to be met with measures that reduce factor prices. In many cases, this means deregulating factor markets. Significantly, Keidanren rarely calls for increased levels of protection. In fact, in late 1996 it presented the government with a list of 699 proposed deregulatory measures. Similarly, private bankers have backed demands for the dismantling of the national savings system. Apart from objecting to the fact that the system prices and allocates capital inefficiently, bankers complain that existing regulations give the government an unwarranted advantage in relation to deposit collection and so crowd them out of their principal factor market. This, they claim, is a particularly serious problem given that the banking system is experiencing extreme difficulties.

The arguments advanced by Keidanren and private bankers share a common motivation in Japan's recent economic misfortunes. Bankers argue that the current problems of the financial system need to be addressed now, and with a fundamental change of philosophy about the role of the state in the allocation of capital. Analogously, manufacturers, who face a high cost structure and a recessed domestic economy, present strong arguments that economic recovery hinges upon less state intervention in their affairs. Both sets of arguments dovetail with the longer-term pressures for economic reform. They show that while many deregulatory initiatives are motivated by the problems that Japan faces as it ages and reaches economic maturity, it is the current slowdown which furnishes reformers with a mandate for immediate deregulation.

Conclusions

The Japanese model has been strained by two simultaneous developments. The prolonged 1990s slowdown has shown that the economy is not immune from deep-seated stagnation. More fundamentally, long-term constraints on growth are increasingly apparent, and market-oriented reforms are widely accepted as the appropriate response.

These developments discredit the Japanese model in two ways. First, events have simply overtaken the forecasts it produced. Japan has failed to meet the expectations that were held for it 20 years ago by advocates of the model. By contrast, the better growth-accounting studies, which eschewed Japanese exceptionalism, have aged extremely well. For instance, the landmark study by Denison and Chung (1976) predicted a post-2002 long-run growth rate of 3.2 per cent. This fits with the EPA's recent forecast of a 3.0 per cent average annual growth between 1997 and 2000. It is certainly nearer the mark than the optimistic projections that were issued in the mid-1970s by proponents of the Japanese model.
The second sense in which the Japanese model has been discredited goes more to the heart of its assumptions. It is damaging enough that its forecasts of high and durable growth should have been invalidated so comprehensively. More serious are the failures of the institutions that were central to those forecasts. In forecasting 3 per cent annual real growth, the EPA (1997:34) warned that such a modest aspiration would be realised only if a program of appropriate deregulation and reform were implemented. Otherwise, growth would be no more than 1.8 per cent. That reform should be promoted by the very government bureaucrats who nurtured the formation of industrial policy is a telling indictment of theories which uphold the managerial capabilities of the state.

Yet despite history’s rough handling of the Japanese model, and despite the current slowdown, the outlook for Japanese growth is not entirely grim. Many recent and planned reforms will undoubtedly produce a more efficient allocation of resources and relax constraints on the availability of capital and labour. Moreover, although Japan has surpassed the West in technology in many areas, in others it is still within the frontier, and so there remains unexploited potential for further productivity growth. For instance, Japan has been slow to join the latest revolutions in telecommunications and computing. However, over the next few years these industries are certain to contribute significantly to productivity.

Although the Japanese economy may eventually recover from its present malaise, economic growth will never again be exceptional in comparison with other industrial countries. And even if it should exceed the standards set by the industrialised world to some modest degree, the differential will not be attributable to an extraordinarily active public sector. It is time to put the exaggerated claims about the Japanese model back into historical perspective.

References


Luke Gower


Nikkei Telecom, Japan News and Retrieval, Nihon Keizai Shimbun, Tokyo.


Tokyo Shookoo Research Ltd, Bankruptcies Monthly (various issues).


I am grateful to Peter Drysdale, two anonymous referees and the editor for helpful comments on earlier drafts. The usual disclaimer applies.
An Optimistic View of World Food Prospects

Ron Duncan

AnALYSES of world food prospects usually concentrate on the alleged limitations of agricultural resources and rapid population growth. The likely effects of rapid income growth in the most populous parts of the world are generally neglected, as are the continuing sharp decline in population growth rates and price responses in terms of input or output substitution.

Concerns about future food production reflect doubts about the ability of agricultural research to maintain its outstanding performance of the post-World War II period, whether because further scientific breakthroughs may not be forthcoming or because governments may not be willing to fund agricultural research sufficiently. Concern also exists about soil erosion or degradation, declining water quality, diminishing germplasm resources, and further increases in cultivatable land or water supplies.

In this article, an assessment is made of the argument that a growing imbalance exists between world population numbers and the world’s ability to feed them. It is argued that such problems as exist are mostly man-made and can be corrected. In particular, poor policies and the absence of effective property rights are responsible for many of the problems affecting the resource base.

Improving Global Food Security

The great increase in food security since World War II has gone largely unrecognised, apart from the major contribution of the so-called Green Revolution. Television images of malnourishment and death from starvation on the African continent are regrettably common; but it is seldom noted that such disasters are due more to civil war and misguided economic policies which have weakened the farming systems of those countries than to limitations of the agricultural resource base and technology. Cleaver (1993) and Cleaver and Schreiber (1992) identify production constraints such as inadequate land-tenure systems and neglect of infrastructure, which, they argue, are compounded by poor economic and agricultural policies and political instability. In global terms, however, the output of cereals (the main food source for most people) has increased by around 2.7 per cent a year.

Ron Duncan is Professor of Economics and Executive Director of the National Centre for Development Studies, Research School of Pacific and Asian Studies, The Australian National University.
since 1950, while the world’s population has grown by about 1.9 per cent a year. To put it another way: cereals output per person has increased by almost 1 per cent a year for over 40 years, when the world’s population has been growing faster than ever before. Cereal yields alone have increased by 2.25 per cent a year since 1950: faster than the rate of population growth. The Food and Agriculture Organisation of the United Nations estimates that total food production has grown by 35.4 per cent per head over the 1961–96 period in high-income countries, and by 50.8 per cent in developing countries. Within developing-country regions, Asia experienced the largest increase (70.3 per cent), followed by Latin America (31.9 per cent). In Africa, food production per head fell by around 1 per cent a year over this 30-year period (FAOSTAT Database).

Within the populous Asian region, food production performances have been extraordinary. China increased food production per head by 165.7 per cent in the 1961–96 period as the result of rapid increases in food production, with a population growth of 1.8 per cent. Indonesia increased food production per head by 69.6 per cent, even though its population grew at the reasonably rapid rate of 2.3 per cent. The increasing proportion of well-fed people in the world is largely due to the gains in Asia, where approximately 60 per cent of the world’s population live. Africa accounts for only 12 per cent of the global population, though its share is increasing because it now has the fastest growth rate.

**Waiting for Malthus**

Beginning with Thomas Malthus in his *Essay on the Principle of Population as it Affects the Future Improvement of Society* in 1798, many commentators have been willing to foresee disaster in a looming imbalance between population and food growth. More recently, Ehrlich (1968), for example, forecast starvation on a massive scale within the decade after the publication of his book, and urged strict population controls. Although the expected disasters have not happened, in recent years concern has again been rising, led in particular by Lester Brown of the Worldwatch Institute in the United States, about the allegedly adverse impact of population growth and income growth on soil, water and air: an impact that is expected to severely damage the planet’s food production capacity.

Taking an extremely pessimistic position on the spectrum of views on the world food outlook, Brown and Kane (1994) argue that a shrinking backlog of unused agricultural technology is coinciding with declining public interest in funding agricultural research; that the demands for water are pressing against hydrologic limits; that the response of crops to additional fertiliser applications is declining in many countries; and that cropland is being substantially reduced by industrialisation and ur-

---

1 Other well-known pessimists are Paddock and Paddock (1976) and Herman Daly (1980, 1996). The modern anti-natal movement, which appears to have racist and eugenic roots, uses arguments about the adverse effects of population growth on economic growth and environmental quality, even though there is no convincing evidence that faster population growth in itself reduces income growth, increases poverty or hastens environmental degradation.
banisation. Moreover, they believe that seafood production has reached its biological limits and that the carrying capacity of rangelands has been exceeded — which places an additional burden on cropland to provide increased food supplies.

In another book which has created considerable anxiety in China, Brown (1995) advances a scenario wherein China's grain production declines from the 341m tons produced in 1990 to 272m tons by 2030, basically because of competition for agricultural land from industrial and urban uses. On the assumption that consumption per head does not increase, population growth is projected to raise total consumption from the 346m tons of 1990 to 479m tons by 2030; if China's consumption per head increases to 400 kilograms — about the same level as Taiwan's today and one-half that of the United States — total consumption will reach 641m tons by 2030. The former consumption scenario leads to grain imports of 207m tons by 2030, while the latter leads to grain imports of 370m tons. The first scenario envisages China importing the equivalent of today's global grain trade, while the second would nearly treble the global grain trade: an outcome which Brown sees as inconceivable.

**Future Food Demand**

In evaluating these and other less pessimistic forecasts of the world food outlook, attention should be paid not only to the perceived constraints on the supply side but also to whether reasonable judgments are being made about the future demand for food.

Alex McCalla, Director of the World Bank's Agriculture and Environment Department, has said that:

> Everyone agrees that the world's population will exceed 8 billion people by 2025 ... everyone agrees that world food supplies will have to more than double by 2025, because of increases in income and urbanisation in addition to population growth. Given this widespread agreement on the needs or demand side of the equation and its magnitude ... why is there so little agreement on the ease or difficulty of generating the supply to meet that demand? (McCalla, 1994:1)

But is this increase in demand so certain, even if we disregard the usual neglect of what might happen as a result of changes in food prices in the various scenarios? The demand for food is dependent on population size and on the effect of income changes on food demand; but population size becomes the more important determinant as incomes increase. At very low income levels, 80–85 per cent of household income may be spent on food; so, as incomes increase from these low levels, food demand tends to grow very quickly. As incomes increase further, however, the marginal propensity to consume food falls. The consumption growth of the basic staple declines fastest. For example, the income elasticity for rice in China is

---

2 China's grain production achieved a record 500m tons in 1996.
now believed to be negative (Ito et al., 1989; Peterson et al., 1991). At the income levels of the high-income countries, consumption growth per head of food changes very little. The world’s cereal consumption per head has not increased appreciably since 1978. In developing countries, the average has not increased since 1984. Given the rapid income growth achieved by a large proportion of the people in developing countries, 3 the slowdown in consumption growth rates indicates that the most rapid phase of cereals consumption has already happened for most of the world’s population. In developing countries which have experienced sustained income growth, dietary patterns have shown significant changes, with shifts away from staple sources of caloric supplies, such as rice or maize, towards wheat, and from caloric sources to protein sources such as meats, fruits and vegetables (see Mitchell et al., 1997). These dietary changes will continue as incomes rise, but the decline in the rate of growth of total food consumption per head can be expected to continue. In China, for example, real GDP grew substantially faster during the 1980s than during the 1970s, yet consumption of cereals grew by an average of 2.3 per cent a year during the 1980s compared with 5.2 per cent a year during the 1970s.

As food consumption growth slows with increasing incomes, population plays a more important role in food demand. But the population growth rate is also slowing rapidly: from 2.06 per cent in 1965-70 to 1.74 per cent in 1985-90. The World Bank (1992) projected that it could fall to between 0.57 and 1.41 per cent by 2020-25: a wide range that reflects the uncertainty surrounding population projections, mainly about the speed of the fertility slowdown in the demographic transition (Bongaarts, 1995). 4 The reduction in fertility has been comparatively rapid in the fast-growing East Asian countries. Presumably, if Sub-Saharan Africa could generate sustained economic growth, population growth there would slow faster than demographers are presently forecasting. Moreover, its food consumption would increase faster than it has in recent decades, and then slow. Under the World Bank’s low population growth rate, the world population would reach 7.6 billion in 2025. The increase in the demand for food by 2025, as determined by population and income growth, is therefore not as certain as McCalla assumes.

How Limiting are the Supply Constraints?

To meet future food demand, it appears that supply will not need to increase as fast as it has over the past 40 years. But, as we have seen, some argue that even a future rate of increase of around 1.7 per cent a year, as expected by the major forecasting agencies, will be frustrated by resource constraints. How realistic are the fears that Brown and Kane articulate?

**Diminishing water supplies.** It is widely agreed that achieving increases in water supplies does indeed pose problems, in terms of both the higher cost of additional

---

3 In 1960-90, the real per head GDP of developing countries increased on average by 162 per cent.
4 The total fertility rate is the number of births a woman would have by the end of her reproductive years if she experienced age-specific fertility rates of a given year.
supplies and the declining quality of existing supplies (Rosegrant, 1997; Rosegrant & Ringler, 1997). Groundwater supplies are being depleted in many regions as pumping exceeds replenishment, and rapidly growing household and industrialisation uses compete with agricultural uses.

But water losses and wastage in agriculture are high because, as Rosegrant (1997) points out, water is often available free or at highly subsidised prices. However, raising the price of water to cover its true cost (private and social) is usually politically difficult. As well, property rights are often not well or easily defined, as in transborder (between provinces or countries) situations; and until such difficulties are resolved water use will usually be inefficient. Developing the appropriate institutional mechanisms presents enormous challenges, particularly with regard to water resources shared among countries. But these challenges are being tackled, as exemplified in Australia by the formation of the Murray-Darling Basin Commission and the progress being made through the Commission in devising ways to share the basin waters more efficiently and to reduce the basin’s severe salination problem. An international agreement has recently been drawn up on sharing the waters of the Mekong River, and discussions are taking place over sharing the waters of the Jordan and the Nile.

**Land availability and land degradation.** Buringh and Dudal (1987) estimate that less than one-half of the world’s land area suitable for crop production is currently being used for this purpose. Much of the presently uncultivated area, however, is located in remote areas in Africa and South America. World cropland has hardly increased since 1960, while land under pastures and meadows has increased by around 30 per cent since 1960. If yield increases can deliver the slower food production growth likely to be needed in the future, the additional cropland requirements will be minimal. These projected increases in crops include the demands for feeding livestock, which implies slower increases in pasture areas.

Crosson and Anderson (1992) argue that the various reports on the loss of arable land through erosion or chemical degradation were exaggerated; indeed, the losses claimed run counter to the increased yields realised from these soils over sustained periods. Rosegrant and Ringler (1997) agree. From his analysis of long-term data on soils in China and Indonesia, Lindert (1996) finds that, of the important ingredients of soils, organic matter and nitrogen do appear to have declined on cultivated lands in both countries, while total phosphorus and potassium have generally increased. Alkalinity and acidity have not worsened, nor has the topsoil layer become thinner. Though China’s soil organic matter and nitrogen have declined, its crop yields do not appear to have been adversely affected: presumably, Lindert argues, because nitrogen fertilisers have been applied.

---

5 Rosegrant and Ringler (1997) state that water-use efficiency in irrigation in much of the developing world is only in the 20-40 per cent range.

6 Worldwide there are around 200 shared river basins, most of them shared between two countries.
Lindert identifies three aspects of economic development which may result in improved soil quality. First, if we take all soil-farming feedbacks into account, the shift in food demand away from staples toward legumes and animal products is likely to replenish soil nutrients. Second, development brings cheaper capital and clearer property rights, which improve conservation. Third, urbanisation and industrialisation raise the productivity of soils at the urban fringe.

In arguing that China’s cereals production will fall well below its demand by 2030, Brown (1995) draws two parallels. First, in countries that, like China, began to industrialise when they were already densely populated (the examples he uses are Japan, Republic of Korea, and Taiwan), farmland was largely converted to industrial and urban uses. Second, most of the world’s major fishing stocks have been severely depleted because of the growth in population and incomes (particularly in Japan, where consumption of seafood products is comparatively high); in a similar fashion there will be a reduction in agricultural production in China because of over-use of agricultural land. But the logic of both arguments is faulty. If agricultural land is converted to other uses in the absence of government intervention, the country’s welfare is increased, and the rise in incomes earned in these other uses can be used to purchase (import) the needed agricultural products. Given its high incomes and low ratios of land to capital and land to labour, it was clearly in Japan’s interests to specialise in non-agricultural products. And although its grain production has fallen, Taiwan is a large exporter of other, more highly valued, agricultural products which it is better placed to produce than grains. As in Taiwan, China’s comparative advantage in agricultural production is shifting away from land-intensive crops such as grains towards more labour-intensive and higher-value crops such as fruit and vegetables (see Feng Lu, 1996).

The depletion of fish and other marine-based stocks stems from the absence of effective property rights to marine resources and the consequent common access problem, resulting in over-exploitation (for discussion, see Ibeduru, 1995; Williams, 1996). But it is clear that, even with effective marine property rights and management controls to ensure optimal harvest rates, the world’s demands for seafoods cannot be met from unfarmed resources that allow no increases in productivity. Responding to the increasing prices of seafood products, the farming of fish and crustaceans has boomed, backed by scientific breakthroughs in breeding and farming methods.

The correct conclusion to draw with respect to Chinese agriculture, and agriculture more generally, is that farmers must be given secure property rights to land so that farming can be efficient and sustainable. As well, if agricultural land becomes scarcer and its price rises, other means of farming involving land-substitution practices will be devised. But, unlike with seafoods, whose price increases triggered the search for other forms of production, there has been no sustained increase in prices of agricultural products signalling the need for a major increase in research effort. In fact, agricultural prices have continued to trend downwards in real terms, at least as measured in terms of wages.
The problem of farmers' insecure property rights to land needs to be addressed in many developing and transition countries where productivity growth is urgently required for enhanced food security. Discussions with World Bank staff stationed in Moscow suggest that the Russian Federation's poor recent agricultural performance is basically the result of the absence of long-term security of land tenure for individual farmers. By contrast, the extension of reasonably long-term access to land for individual farmers in China and Vietnam was a key reason, together with market-based prices, for the increase in their productivity (see McMillan et al., 1989). In China, the original 15-year leases were extended for another 30 years in 1995. To allow scope for improved productivity through increased farm size, the leases should also be made transferable.

It is also clear that, in many developing countries, more secure property rights to land — whether freehold or long-term leasehold tenure — and cuts to the punishing taxes frequently imposed on agriculture are needed to prevent further soil degradation and even to improve soil quality. In most cases, the most appropriate management of land and water resources will be achieved by institutional developments (the creation of property rights and of markets for trading in those rights) which internalise the external costs associated with their use. This may well mean that marginal land will no longer be farmed (or not farmed as intensively), or that water use in agriculture will be reduced in response to higher prices. But, overall, resource allocation and social welfare will be improved. The impact of any consequent reductions in resource use should be more than offset by the improved productivity from the institutional developments. More secure access to land and water should lead to greater investment in land, including investment that promotes soil and water conservation, which in turn should lead to higher and more stable output.

Prospects for continued yield increases. The pessimists argue that crop yields have slowed over the past decade or so, for reasons such as the declining response to increases in fertiliser applications, and slower development of improved varieties. A different view, argued above, is that yields do not need to grow as fast as in the past, because demand growth is slowing. But how likely is it that slower rates of cereals yield growth, of around 1.5 per cent or even 2 per cent a year if required, can be achieved over the next 30-40 years?

Yield increases can be achieved either by extending the use of known technology or by the development of new technology such as improved crop varieties, improved fertilisers, or improved farming practices. Large gaps between on-farm and experimental yields exist in many developing and transition economies due to the lack of application of existing technologies. Closure of these gaps should be greatly

---

7 Jacobsen et al. (1996) discuss the creation of effective property rights within a communal setting which led to much more effective use of soil and water resources and reduction of external costs from soil and water run-off. Hanstad and Li Peng (1995) report on an experiment in the auctioning of transferable use rights to some of the 33m hectares of 'wasteland' in China and Luliang Province (this is uncultivated but reclaimable land which was not distributed to householders during decollectivisation). It is claimed that the secure access to use rights will reduce soil erosion and increase forest cover.
enhanced by improved institutions and policies, particularly more secure property rights to land, lower taxation of agriculture, and liberalisation of output and input (particularly fertiliser) markets. As well, there is ample room for improving the scope for yield-increasing new technologies as theoretical maximum yields — determined by photosynthetic potential, land quality, length of growing season and water availability — are many times higher than actual yields on average (Plucknett, 1995). The challenge is for research to provide the knowledge needed to reduce this gap.

Despite concern about the loss of biodiversity from developmental activities such as forest clearing, very substantial stocks of germplasm of the major food crops are held by research bodies in various countries (Wright, 1996). Enough germplasm is available in storage banks, quite apart from what might be available in situ, to support future conventional plant breeding research and biotechnology applications. Research applying biotechnology techniques holds the potential to improve crop yields through introducing higher plant resistance to pests and diseases, higher tolerance to adverse weather (such as drought) or soil conditions (such as salinity) and improved responses to fertilisers. Plant breeding can thus substitute for declining availability of good quality arable land, just as research into soil nutrition can increase the availability of arable land (as it has in Australia).

Research in the improvement in fertiliser responsiveness should have a high pay-off as levels of application of fertiliser in relatively favourable farming areas in Asia are now quite high and response rates appear to be slowing. Crop productivity can be raised without increased fertiliser use by improving nutrient uptake efficiency and nutrient balance (Rosegrant & Pingali, 1994). Improved efficiency of fertiliser use is also needed to reduce the impact of fertiliser run-off on water supplies. This is a problem in Western Europe and parts of other high-income countries; but, except for intensively cultivated areas of East Asia, developing countries use too little fertiliser, which reduces soil fertility and increases erosion (Rosegrant & Ringler, 1997). This under-utilisation of fertiliser is often attributable to fertiliser import restrictions, which provide privileged interests with monopoly rents.

Conclusions

Concerns about an impending or even distant global imbalance between population growth and food supplies are exaggerated. The major problems in the food supply system are either man-made or can be corrected through institutional developments. Pessimistic commentators generally ignore the effect on food consumption growth rates of the substantial increases in incomes achieved by a large proportion of the people in the developing world. This, together with their rapidly slowing population growth rates, means that future growth in cereals consumption will not be nearly as fast as it has been during the past 40 or so years — the period of most rapid growth of the world's population. As a result, food supplies do not need to grow as fast as they have in recent decades. Moreover, the concerns about future production growth are generally not warranted: or, if warranted, the causes and proposed solutions are misplaced.
Increasing water supplies will be difficult and costly; but much can be done to make the use of existing supplies more efficient. Water is too often unpriced to farmers; despite the political difficulties, this has to change. Sharing of water supplies across countries also challenges policy-makers to devise institutions which achieve the most efficient utilisation of water supplies.

Establishing long-term, secure access to land for farmers is the most urgent need in many developing and transition countries and will make the greatest contribution to their food security. With secure access, farmers are more willing to invest in the land and adopt improved technologies, leading to higher productivity and reduced soil loss and degradation.

Besides market failure reflecting the lack of appropriate property rights in land and water, extensive policy failures in many developing countries adversely affect agriculture through excessive taxation of agricultural output, either directly or indirectly (such as through over-valued exchange rates or import restrictions on fertilisers). Again, reforming these policies will be resisted by special interests, such as urban populations that enjoy subsidised food prices.

It is, of course, uncertain whether research can continue to generate the necessary knowledge to keep yields increasing by extending the production capacity of land through substitution for natural resources. In a sense, land ceased to be a limiting constraint in the post-World War II period as a result of scientific discoveries. In view of the base of scientific knowledge which has been established in the 20th century, there are no grounds for pessimism about the potential for further scientific discovery. While the scope for private agricultural research has increased with the development of property rights in new technology, particularly biotechnology, there remains an important role for government funding of national and international research to increase agricultural productivity. But decisions about such funding should be based on objective analysis of the costs and benefits.

References


Ron Duncan


Feng Lu (1996), 'China's Grain Trade Policy and Food Trade Pattern', RSPAS, Australian National University, Canberra (mimeo).


The author is grateful to two anonymous referees for their helpful comments.
‘Judge & Co.’: Judicial Law-Making and the Mason Court

Geoffrey Lindell

It has been said that ‘judges were the traditional lawmakers’ before ‘they were overtaken by the onslaught of democratic theory and supplanted by parliament’. The onslaught followed in the wake of Jeremy Bentham’s withering criticism of judges’ role in the development of the law. Nevertheless, it is now widely accepted that judges still make law but in a special and more limited sense than the political branches of the government. Justice Michael Kirby (1983:59) has often reminded us of Lord Reid’s (1972:22) famous remark that the declaratory theory of the common law is based on a ‘fairy tale’ which is no longer believed. So the issue, in Australia as in other parts of the world which apply Anglo-American common law principles, is not so much whether they make law, but in what sense, and in what circumstances.

The decisions of the Australian High Court during 1986-95, when Sir Anthony Mason presided as Chief Justice, saw dramatic changes to judge-made law. It was almost inevitable that decisions like the Mabo case would excite controversy about the legitimacy of the role of judges in making law. The recent publication of a book (Saunders, 1996) on the achievements of the Mason Court and the role of appellate courts of appeal in Australia and elsewhere makes it a timely occasion to revisit this debate.

Areas of Change

Among the Mason Court’s most important constitutional decisions were those which established a constitutionally guaranteed freedom of political communication, the cases which concerned the separation of Commonwealth judicial power from legislative and executive power, and the Mabo case. Developments occurred in relation to

Geoffrey Lindell is Reader in Law at The University of Melbourne.

1 Sturgess and Chubb (1988:2-3), who referred to Jeremy Bentham’s 1809 work Elements of the Art of Packing. Bentham coined the term ‘Judge & Co.’ referred to in the title of this article.


3 This is a collection of essays which were originally delivered by distinguished jurists and lawyers from Australia and elsewhere at a highly successful conference to mark the retirement of Sir Anthony from the Court. The essays in the first part of the book serve as an extremely useful record of the achievements of the Court during this period.

4 Australian Capital Television Pty Ltd v the Commonwealth (1992) 177 CLR 106 (‘ACTV case’, as regards the freedom of communication); Chu Kheng Lim v Minister for Immigration and Ethnic Affairs
s.92 (guarantee of freedom accorded to interstate trade) and s.117 (prohibition of discrimination on grounds of residence in other States). As well, there was a steady and decisive expansion of the rules of procedural justice (formerly known as 'natural justice') in the field of administrative law. This has been seen as a response to the previous increase in the range of powers and discretions vested in administrative agencies, resulting in the establishment of a 'unifying principle that decision makers whose decisions could cause loss of any kind should adopt fair procedures'. This was thought to be 'justified and appropriate as a proper check against oppressive government power' (Finklestein, 1996:62).

The developments on procedural justice included the controversial use made of international agreements as a means of enlivening the application of the same rules even though the agreements were not given the force of domestic law by the parliament.

Important changes occurred in the criminal law as regards the rights of defendants (for example, to legal representation in serious criminal trials); doctrines of estoppel; and unjust enrichment and restitution. It has been said that those changes illustrated the concern of the Court with the concepts of fairness or fair dealing or reasonableness in the relationships of citizen to government and citizen to citizen. An underlying view of reasonableness showed itself also in the trend towards making all cases in tort depend on intention or negligence (Jackson, 1996:24; Priestley, 1996:99-100). This involved the reversal of a number of long-standing cases including a case which had stood for 126 years in favour of strict liability for the escape of dangerous things.

An important contribution made by the Mason Court to the development of other areas of the law which govern dealings as between citizens was the treatment of fair dealing and good faith. This was illustrated by the increasing application of equitable doctrines to contracts and commercial law (see Renard, 1996). The beginnings of these developments pre-dated the period under review and were also accompanied

---

(1992) 176 CLR 1, 27 (protection from arbitrary arrest because of the inability of the parliament to vest judicial power in non-judicial bodies).

5 Cole v Whitfield (1988) 165 CLR 360 (s.92) and Street v Queensland Bar Association (1989) 168 CLR 461 (s.117).

6 For example, Annett v McCann (1990) 170 CLR 596; Ainsworth v Criminal Justice Commission (1992) 175 CLR 564.

7 Minister for Immigration and Ethnic Affairs v Teoh (1995) 183 CLR 273. The case is seen by some as undermining the important democratic principle that the executive branch of government should not usurp the role of the elected representatives of the people in altering the law of the land. But see Walker (1996) and Evans (1996).


9 Burnie Port Authority v General Jones Pty Ltd (1994) 179 CLR 520 overruling Rylands v Fletcher (1868) LR 3 HL 330 which was said to by Mr Jackson (1996:24) to have fallen under the 'steamroller of negligence'.
by legislative changes (such as the enactment of the Trade Practices Act and corresponding State and Territory legislation) which gave the courts greater powers to grant remedies against unfair dealings, especially in the field of consumer transactions, while apparently falling short of adopting the American concept of good faith as an essential element of the law of contract.

Australian courts have historically enjoyed the jurisdiction possessed by the English Court of Chancery to set aside contracts on the grounds of fraud, misrepresentation, breach of fiduciary duty, undue influence and unconscionable conduct. However, what made the application noteworthy under the Mason Court was the greater willingness of the courts to exercise the jurisdiction and to continue to develop the doctrines which regulate its exercise. This had a very significant impact on the law of contract. The High Court’s approach to unconscionable conduct has been to prevent the enforcement of contractual bargains where this would lead to an unfair or unconscionable result. It has also enabled the courts to enforce the promises or reasonable expectations of parties to a commercial transaction, notwithstanding any technical difficulties which might otherwise have prevented this taking place.

These developments prompt some reflections on the era under discussion and the state of judicial law-making by the time Sir Anthony Mason retired from the Court.

Achievements of the Mason Court

By the time Sir Anthony Mason became Chief Justice, the process of ending appeals to the Privy Council had been completed as a result of the Australia Acts 1986, as had the establishment of the mechanism which ensures that all appeals heard by the High Court now lie in the discretion of the Court. That said, it is difficult to avoid the impression that the Mason Court favoured a degree of judicial activism. Many of the developments which occurred bear the hallmarks of Sir Anthony Mason’s judicial philosophy (which itself underwent considerable change during his long period in office as a judge). Sir Anthony rarely found himself in dissent.

A progressive attitude. The Mason Court was progressive at a time of great and accelerating social and economic change (Saunders, 1996:2). Its willingness to embrace technological change is illustrated by its decision in McKinney v The Queen when the majority changed the rules relating to the admissibility in criminal trials of confessions

---

10 A leading illustration of the latter development was the relaxation of the rules relating to consideration and privity of contract as essential requirements for the enforcement of a contract: Waltons Stores (Interstate) Ltd v Maher (1988) 164 CLR 387; Trident General Insurance Co Ltd v McNiece Bros Pty Ltd (1988) 165 CLR 107.

11 The less activist approach which he formerly supported is illustrated by the cases decided in the early 1980s, discussed by the Chief Justice of the South Australian Supreme Court, John Doyle (1996:90-1); see also Kirby (1996). The change may perhaps be attributed to a heightened disillusionment with the effectiveness of the legislative process as a means of effecting necessary changes in the law.
obtained from persons under interrogation while in police custody, by treating them as unsafe if they were not recorded on videotape. The same willingness to respond to changes in community perceptions is illustrated by *R v L* when it was boldly asserted that, if it ever was part of the common law that marriage constituted continuing consent to sexual intercourse, it was no longer so in modern times. Yet the Court failed to embrace the new approaches to the rules of private international law recommended by Law Reform Commissions and academic writers. Significantly, this was one of the few areas where Mason CJ lacked the necessary support of the rest of the Court for departing from the traditionally accepted rules.

**An accessible style.** The Mason Court abandoned the ‘formalistic’ and ‘legalistic’ style of judicial reasoning in favour of a more accessible style, even if this runs the risk of exposing the Court to greater criticism when it relies on policy considerations which are no longer hidden by the veil of legalism. This development seems to indicate that the Court might now have accepted the suggestion made long ago, namely, that ‘the ground rules for principled (judicial) decision-making are exactly the same as those for the *intelligent* discussion of any issue’ (Evans, 1976:68).

**An Australian approach.** The Mason Court is thought to have developed a distinctly Australian approach to its development of the law. The heavy if not exclusive reliance on English sources of law has now given way to a more varied willingness to consider sources of law in other countries for non-binding guidance and assistance. The development began before the elevation of Sir Anthony as Chief Justice and represents the achievement of a kind of national maturity which is the delayed but inevitable result of the formal attainment of Australian political and judicial independence.

**Protecting the individual.** For the author, perhaps the most important aspect of the Court’s work during this period has been its concern for protecting the individual against the abuse of private and public power. Sir Anthony Mason and the Court as a whole obviously felt that the Court has a very significant and appropriate role to play in dealing with the rights of individuals: a role that does not necessarily presuppose the existence of a judicially enforceable Bill of Rights. His Honour had suggested that the

---

13 (1991) 174 CLR 379. A more recent, if not even more striking illustration of the dynamic view taken of the nature of the common law resulted from the need to bring the law of defamation into conformity with the Constitution (and in particular the implied freedom of political communication). This was achieved by the reformulation of the common law defence of qualified privilege in *Lange v Australian Broadcasting Corporation* (1997) 71 ALJR 818 (*Lange case*).
14 Although it was unnecessary for the decision, Brennan, Dawson, Toohey and McHugh JJ reaffirmed the need to establish liability for interstate torts according to both the law of the place where the alleged wrongful act occurred and the law of forum in *McKain v R W Miller & Co (SA) Pty Ltd* (1991) 174 CLR 1. This was so despite the Court’s decision in the earlier case of *Breavington v Godleman* (1988) 169 CLR 41 where his Honour elaborated his support for the preferable and less mechanistic ‘proper law’ approach.
15 For further discussion, see Finn (1996).
degree to which a decision was 'determinative of the rights or interests of an individual' was a leading criterion for determining whether a decision should be the subject of judicial review (Mason, 1989:124).

**Equitable rules.** So far as private law is concerned, it is questionable whether the courts should extend the reach of equitable rules in relation to commercial transactions where the parties can be expected to have access to legal advice and enjoy equal bargaining power, leaving aside the transactions which involve small business and consumers. There is the further need to strike an appropriate balance between fairness, on the one hand, and on the other, certainty and predictability in commercial transactions. Some will also question the ability, as well as the need, for judges to fix standards in the market place for such parties. They will ask why the same parties should as a rule be able to use a valuable public resource — the court system — to arbitrate their commercial disputes.

**Public law.** In the field of public law, the Mason Court become increasingly occupied with restrictions on legislative power which are not concerned with the federal nature of the Constitution. As a former Commonwealth Solicitor-General, Sir Maurice Byers QC, observed, if anything characterised the Mason Court it was the implicit rejection of an aspect of the landmark decision in the *Engineers* case which strongly affirmed that the ballot box was the only answer to potential abuse of power exercised by the majority in parliament. This is consistent with Sir Anthony Mason's view that 'our evolving concept of the democratic process' was 'moving beyond an exclusive emphasis on parliamentary supremacy and majority will' and as embracing 'a notion of responsible government which respect(ed) the fundamental rights and dignity of the individual...’ (Mason, 1987:163).

**Mabo.** The *Mabo* case raises interesting parallels with the landmark decision of the US Supreme Court in *Brown v Board of Education* regarding the desegregation of public schools. The protection accorded to native title consists of the inability of the Commonwealth to legislate for the acquisition of native title without the provision of 'just terms' (that is, compensation), and also the inability of the States to legislate in respect of native title otherwise than in accordance with the terms of the Racial Discrimination Act 1975 (Cth) and Native Title Act 1993 (Cth) by reason of s.109 of the Constitution.

One rationale that is sometimes advanced for the developments which began with *Brown v Board of Education* and culminated in the activism of the Warren Court in the United States in the 1960s is that the court should intervene to protect individuals...
and groups who are denied real access to the political process. Perhaps a similar explanation can be advanced in relation to the High Court's solution to the native title question. The pursuit of this approach does achieve very valuable results at least in the short term and the decision of the Court in *Mabo* certainly did have the effect of breaking a difficult political and legislative log-jam on the Aboriginal land rights question. But to adopt, in the words of Mr Jackson QC (1996:25), 'reasoning' which 'did not follow earlier perceptions of the effect of European settlement' comes at a considerable cost, as is illustrated by the torrent of virulent criticism unleashed against its decision in that case and also the even more virulent (but with far less justification) criticism unleashed against the succeeding case in *Wik*.\(^{19}\) The fact that, as Sir Anthony Mason (1996:114) observed, the Australian High Court did for the indigenous inhabitants of Australia only what the courts in the United States, New Zealand and Canada did for their indigenous peoples is unlikely to calm the voice of the critics in Australia. The adoption of the rationale adverted to above also has the potential to make the courts a potential dumping ground for issues which are too hot for politicians to handle.\(^{20}\) In the author's view, it is extremely doubtful whether courts should act to upset traditional understandings of the law in advance of public opinion or where significant sections of the community are divided on the need for a change in the law.

*Section 92.* The praise accorded to the decision in *Cole v Whitfield* is well deserved (Zines, 1995). The case adopts an interpretation which probably and more closely reflects the intentions of the framers when they resolved to provide that 'trade, commerce, and intercourse among the states ... shall be absolutely free' in s.92 of the Constitution. This was done in a unanimous judgment, 'the most important parts' of which were said, by the present Chief Justice, to have been written by Sir Anthony Mason (Brennan, 1996a:13). In that judgment the Court adopted a free-trade (federal) approach to the interpretation of that section (aimed at the elimination of protectionism in a common market). This had the effect of abandoning the misplaced *laissez faire* view of the section which had prevailed since the *Bank Nationalisation* case was decided in 1949 when legislation to nationalise private banking was held to violate s.92 despite the absence of any suggestion that the legislation discriminated against interstate banking.\(^{21}\) The Court was able to clarify the interpretation of one of the most litigated sections of the Constitution even if the unanimity of the Court was disturbed shortly afterwards in the application of the new test in relation to the taxation of interstate trade and commerce.\(^{22}\) As Emeritus Professor Zines (1995:18) correctly points out, the effect of the new test will be to increase the scope of both Commonwealth and State power to regulate and control trade and commerce. If

---

\(^{19}\) *Wik Peoples v Queensland* (1996) 71 ALJR 173.

\(^{20}\) Adapted from the language used in Sturgess and Chubb (1988:123), who, however, refer to the tendency of modern legislatures to leave certain issues to be determined by the courts as a reason for justifying judicial creativity.

\(^{21}\) *The Commonwealth v Bank of New South Wales* (1949) 79 CLR 497.

\(^{22}\) *Bath v Alston Holdings Pty Ltd* (1988) 165 CLR 411.
so, it will not be the first time that the Court has interpreted the Constitution in a way that expands governmental power at a time when the governments then in office show the least signs of exercising that power. This occurred, for example, with the effect of the Engineers’ case on the scope of the Commonwealth’s legislative powers. The conservative Bruce Government showed little interest in realising the full potential of those powers during the 1920s (Sawer, 1956:329; Zines, 1992a:75-6). The same can obviously be said about the ability of governments to nationalise banks and other forms of trade and commerce at the present time. But the Court’s new approach will at least have the effect of restricting the destructive potential of s.92 to invalidate legislation which deals with the control and taxation of trade and commerce including such matters as the marketing of primary products and transportation of goods and persons.

Professor Zines (1995) has rightly praised the Court’s tendency to prefer substance over form in its new-found willingness to give greater effect to constitutional restrictions on legislative power (such as s.117 of the Constitution). The earlier tendency to give restrictions on power a restrictive, narrow and formal operation was probably symptomatic of the lingering influence of the British doctrine of parliamentary supremacy.

The Court’s willingness to depart from the previous interpretations of s.92 was not matched by its reluctance to disturb the wide view taken of the meaning of excise in s.90 under which the States are excluded from imposing a wide range of taxes on the production, manufacture and distribution of goods. But this may be due to the policy predilections of the majority who viewed the purpose of s.90 as helping to ‘create a Commonwealth economic union and not an association of States each with its own separate economy’. A majority of the present Court subsequently adhered to the wide view of excise but in a way that destroyed the practical significance of certain kinds of franchise fees on tobacco and liquor which the States had previously been able to levy under an anomalous exception to the wide view.

**Implied rights.** The ACTV case has attracted much popular interest at least for the result it achieved in recognising the existence of some kind of constitutional guarantee of the freedom of political communication which restricts legislative power and is implied as an indispensable element of the representative democracy that is recognised in the Australian Constitution. Perhaps one of the reasons for its general popularity was its attractiveness to the media for the effect it had, as a result of subsequent decisions, in ‘reforming,’ on a uniform basis throughout Australia, the law on defamation so to facilitate the discussion of political matters. Attempts to achieve reform in

---


24 Ha v New South Wales (1997) 71 ALJR 1080 per Brennan, McHugh, Gummow and Kirby JJ; Dawson, Toohey and Gaudron JJ dissenting.

this area by means of cooperation between Australian government and parliaments have proved to be singularly unsuccessful.

More recently, in the *Lange* case the Mason Court decisions on defamation were accepted as authority for deciding that the law on defamation as previously interpreted failed to conform with the freedom in question. But the law was brought into conformity with the Constitution in a different way from that suggested by the earlier decisions, although this seems to have produced substantially similar results to those achieved in the earlier defamation cases even if those results do not stem directly from the operation of the constitutional freedom.

What has concerned some legal commentators is not the utility of the results described above but rather the judicial vehicle used to bring them about. Judicial implications from the structure of the Constitution, for example in relation to federalism and the separation of judicial from legislative and executive powers, are not new. What was new about the process of implying restrictions on legislative power from the representative nature of the governmental institutions created by the Constitution was the fact that such implications had the potential to defeat the acknowledged intentions of the framers of the Australian Constitution not to adopt a Bill of Rights. The orthodox understanding was that in that respect Australia remained more closely aligned with British notions of constitutional law, under which as Dawson J was to point out 'the guarantee of fundamental freedoms' did not lie in 'any constitutional mandate but in the capacity of a democratic society to preserve for itself its own shared values'.

This gives rise to serious doubts about whether the judges enjoy a democratic mandate to give effect to an implied Bill of Rights, as well as the usual concerns about the use of implications as a vehicle for judges giving expression to their own subjective values.

Some of these concerns may be alleviated by Sir Anthony Mason's acknowledgment that in view of the decision of the framers to eschew the American model of a Bill of Rights 'it is difficult if not impossible to establish a foundation for the implication of general guarantees of fundamental rights and freedoms'. The implication he supported was confined to strengthening the workings of representative democracy. This would help to perfect the integrity of the process by which decisions are made rather than determine the outcomes which emerge from that process.

Not all judges were or would be content to limit the process of implication in this way. Quite apart from the inherent potential of a freedom of political communication to embrace other important freedoms such as freedom of movement, association and assembly, the concept of representative democracy might be seen to extend to other

---

26 *Lange* case (1997) 71 ALJR 818. But the present Court failed to confirm the application of the constitutional freedom to discuss political matters which are not relevant to the federal level of government.

27 ACTV case (1992) 177 CLR 106 at 183.

28 Ibid at 136.

29 As Professor Zines (1991:34-5) has indicated, 'provisions concerned with the establishment and maintenance of democratic processes appear, to some degree, to shade into those which are thought
basic rights which are not restricted to ensuring the genuine and effective nature of the election process. As well, some judges have in the past suggested that equality is necessarily implied from the Constitution even though the framers explicitly refused to include this right in the Constitution: which is true also of the implied constitutional inability of the Commonwealth parliament to create retrospective criminal offence.

Admittedly, the chances of a majority accepting the implication of equality seem greatly diminished given the Court's rejection in the recent case which involved the removal of Aboriginal children from their families. Moreover, a majority of the Bench as presently composed seems to be opposed to the use of general concepts such as representative democracy as free-standing principles which operate independently of those provisions, when those concepts are only partially recognised in the express provisions of the Constitution; that is, as a reason for invalidating legislation which does not otherwise breach the express provisions of the Constitution. Reference has already been made to the decisions of the present Court which may be taken to have consolidated rather than extended previous developments; but generalisations can be hazardous given the surprising case of *Kable v Director of Public Prosecutions (NSW)* where a majority of the Court has begun to extend some aspects of the separation of judicial powers case into the State judicial sphere.

**Concluding Remarks**

It has been said that the end of the Mason Court era coincided with a vigorous debate about the limits of judicial creativity in Australia. It was also said that the debate clearly reflected the reputation earned by the Mason Court especially during the latter part of the former Chief Justice's tenure as a court of final jurisdiction which gave effect to a conscious philosophy of judicial activism (Sackville, 1997:145-6). Many leading members of the Australian judiciary have acknowledged that judges make law as a result of the inevitable choices which confront them in the decision-making process (see desirable by people who emphasise, not merely democratic structures, but liberty of the individual and the protection of the individual'.

Although the suggestion in relation to equality was supported for different reasons by Brennan, Deane, Toohey and Gaudron JJ in *Lecth v The Commonwealth* (1992) 174 CLR 455, the case has not been taken as establishing its correctness. The same can be said in regard to the support expressed by Deane, Toohey and Gaudron JJ in *Polyukhovich v The Commonwealth* (1991) 172 CLR 501 as regards the suggestion in relation to retrospective criminal offences. Both cases are discussed in Zines (1997:418-20 and 210-12, respectively) and Lindell (1994:33-7).

*Kruger v The Commonwealth* (1997) 71 ALJR 991 per Brennan CJ, Dawson, Gaudron, McHugh and Gummow; Toohey J dissenting. Three of the majority judges rejected the implication without qualification but Brennan CJ's remarks can be confined to laws for the Territories enacted under s.122 and Gaudron J supported a limited constitutional guarantee of equality.


(1996) 70 ALJR 814 where a majority of the Court held invalid State laws which provided for the judicial authorisation of the preventative detention of persons who were thought to pose a threat to the lives of others (Toohey, Gaudron, McHugh and Gummow JJ: Brennan CJ and Dawson J dissenting).
for example Doyle, 1996; Sackville, 1997; Kirby, 1983). The present Chief Justice of the High Court said:

In modern times the function of the courts in developing the common law has been freely acknowledged. The reluctance of the courts in earlier times to acknowledge that function was due in part to the theory that it was the exclusive function of the legislature to keep the law in a serviceable state. But legislatures have disappointed the theorists and the courts have been left with a substantial part of the responsibility for keeping the law in a serviceable state, a function which calls for consideration of the contemporary values of the community.\(^{34}\)

The notion of ensuring that the law is kept in a ‘serviceable state’ is particularly applicable to what might be termed ‘lawyers law’ in the areas of law such as torts and contracts. Parliaments may well assume that the courts will develop in an orderly manner the law in those areas safe in the knowledge that any unwanted results of that development can if necessary be reversed by overriding legislation. The assumption here is also that courts do not change the law on matters which divide the community. Even in the area of ‘lawyer’s law,’ however, it is important for the courts to develop techniques to ensure that the expectations of persons who plan their affairs in reliance on previous case law are not defeated. The difficulty with the judicial acknowledgment of law-making by judges is, however, that, as Professor Brian Galligan has correctly pointed out, the judiciary may have failed to prepare the public for that acknowledgment.\(^{35}\) Public understanding would not be assisted by the acceptance of a literal and oversimplified version of the doctrine of the separation of powers.

Sir Gerard Brennan has attempted to confine the acknowledgment of judicial law-making to the development of the common law (as distinct from the interpretation of the Constitution)\(^{36}\) and also to stress, as others have, that judges make law in a very different way from that of the political branches of the government. He emphasised that ‘the development of the substantive principles of law by the High Court must be the outcome of the application of the judicial method to the cases that come before it’ (Brennan, 1996b:264; emphasis added).

The attempt to confine the acknowledgment of judicial law-making to the development of the common law and so deny its application to the interpretation of the Constitution is curious. While undoubtedly deriving some theoretical support from democratic considerations, it will be difficult in practice for a court not ‘to make law’ and thus have regard to policy considerations, in any area of the law including constitutional law, at least when those consideration are used in the same special and con-

---

34 *Dietrich v The Queen* (1992) 177 CLR 292 at 319 per Brennan J as he was then.
35 Quoted by Doyle (1996:85). If so, Kirby J should be seen as an exception.
36 *Theophanous v Herald and Weekly Times Ltd* (1994) 182 CLR 104 at 143.
fined sense that will be explained below. This is because of the necessary brevity and open nature of constitutional provisions which call for interpretation.

Sir Gerard Brennan's attempt to emphasise the use of the 'judicial method' is more persuasive. For him, the scope of judicial policy, as distinct from political policy, is informed by precedent and disciplined by analogy; and these factors confine the scope for discretionary judgement (Brennan, 1996a:13). Sir Anthony Mason stressed the place of the strong traditions of consistency, coherence and continuity in the orderly development of the law. Judicial decision-making was seen as principled and reasoned, in contrast to political decision-making which involves compromise and expediency. Moreover, the need to resolve disputes between the parties before the court necessitates closer attention to the interests of individuals since it is concerned with the particular, whereas political decision-making is much more concerned with the general. The differences also go to procedure and method in that judges are bound to hear the interested parties and give public reasons for their decisions. They are also required to deal with problems raised by litigants while politicians are free to decline to deal with problems raised for their decision (Sturgess & Chubb, 1988:346-7).

A final factor is that courts can decide only the cases that come before them and cannot on their own initiative decide issues in the absence of litigation which involves their determination. As well, in Australia the ability of federal courts to render advisory opinions is restricted. The significance of these factors should not be exaggerated since, as a final court of appeal, the High Court has a high degree of control over the cases which it chooses to hear under the special leave process. In addition judges can signal their interest in certain issues in a way that may encourage legal advisers to recommend the initiation of litigation which will raise those issues. Some constitutional cases also have a suspicious similarity with advisory opinions, given the willingness of the High Court to entertain in some cases challenges to the validity of legislation before it comes into force.

If judicial law-making is now accepted by judges of Australian courts, or at least those at the appellate level, some difficult issues remain regarding the way this role is to be performed in the future, as was clearly demonstrated by Doyle CJ (1996) of the South Australian Supreme Court. His Honour believed that the work of the Mason Court would come to be regarded as marking a significant development in Australian legal thinking and judicial techniques. It has prepared Australian courts for a new approach to their law-making function. But awaiting its successors were the tasks of defining more clearly the limits between the law-making role of the Courts and that of parliament; clarifying and expounding the judicial techniques which support the law-making function; and refining the procedures of the Court to enable the function to be carried out in the best possible way. The warnings are timely and he is not the only judge to have issued them (see for example Sackville, 1997).

37 For a typical example see Zines (1992b:54-9) on how it is almost impossible for judges not to have regard to functional considerations in determining the characterisation of laws in the distribution of powers between the Commonwealth and the States.
In relation to the function of courts to keep 'the law in a serviceable state', Chief Justice Doyle (1996:93-4) has pointed to the significant distinction between the need for changes generally and the need for those changes to be brought about by judicial action. Searching questions are raised as to whether the reasons given by the High Court in such cases as Mabo sufficiently addressed that distinction. This was so in the latter case despite the fact that it is seen by some as coming very close to 'the boundary between the appropriate roles of the legislature and the judiciary' (Jackson, 1996:25) and also that the practical impact of the decision was such that an urgent legislative response was required since it was widely agreed that the impact of the decision could not be left to be sorted out in accordance with common law principles on a case-by-case basis. The legislation in question was to have very significant political, social and economic consequences (Doyle, 1996:93).

The procedures adopted by the Court to carry out its law-making function will need to address the tendency of judges in the modern era to resort to community values and perceptions as a reason for changing the law. They will also need to address the associated need for more systematic materials to be placed before the courts in relation to such matters. Changed social, political and economic conditions are legitimate factors which call for changing judge-made law. Yet the use made of such values is obviously not free from danger, especially as some judges could smuggle their own subjective values under the guise of perceived community standards, and arriving at the commonality of such values is difficult in an increasingly multicultural society. Sir Anthony Mason is probably right to reject any calls for the introduction of surveys to gauge contemporary values (1996:114); but Doyle CJ (1996:96-7) is surely right to insist on more systematic materials being placed before the Court for this purpose.

In a recent case the Court may have failed to keep the law up to date when it refused to accept the right of patients to obtain copies of their medical records despite Canadian judicial developments which affirmed the existence of such a right, and despite the acknowledgment by two members of the Court that a majority of the community would support such a right. A liberal approach may also need to be adopted in relation to the exercise of the power to allow for the intervention of third parties. Another technique which deserves greater attention is the possibility of a court engaging in prospective overruling so as to ensure that when previous cases are overruled the effects of the establishment of any new rules are confined to the future and to the immediate parties in the litigation which gives rise to the new rules. Its full acceptance probably calls for the demise of the lingering remnants of the declaratory theory of law. The technique, however, seemed to receive short shrift in the recent Ha case on s.90 where it was rejected by both the majority and the minority on the ground that prospective overruling was inconsistent with the exercise of judicial

---

38 Breen v Williams (1996) 70 ALJR 772. See also Sackville (1997:159).
40 The existing law on the matter was discussed by both Brennan CJ and Kirby J in Levy v Victoria (1997) 71 ALJR 837 at 843-6, 873-4 respectively.
power, despite the absence of any discussion of the tentative steps that may have been taken in that direction before the decision of the Court in the *Ha* case.

In 1921, a famous American judge, Benjamin Cardozo, lucidly analysed the question of judicial law-making in a way that fully and freely accepted its existence. But he accepted that for practical reasons, if nothing else, adherence to precedent should be the rule and not the exception. Cardozo (1978:112-13) summarised the process of judicial law-making as requiring courts to have regard to four factors: (i) logical progression (mainly inductive rather than deductive as involving law-making by analogy); (ii) historical development and evolution; (iii) the customs, morals and traditions (including the accepted standards of ‘right’ conduct) of a society; and (iv) social utility or welfare. Courts and the legal community generally have yet to improve on this test in elaborating the limits of judicial law-making.

The challenge for the modern judge remains, in the words of Justice Michael Kirby (1997:11), to ‘find where the line lies in a particular case, at a particular time and place’. As indicated by Professor Saunders (1996:8), the most significant issue for the judiciary here and around the world ‘is the difficult and delicate task of finding an appropriate balance between the functions of the courts and the other arms of government’ particularly in age ‘in which the assumptions of majoritarian representative democracy have been eroded from both inside and out without the full ramifications of the change being, as yet, fully explored or understood’. That said, ‘the legacy of the Mason Court is a useful base from which to start’ to resolve such issues.

References


42 *Bropho v Western Australia* (1990)171 CLR 1 (but note the refusal in that case to adopt the technique by Brennan J as he was then); *McKinney v The Queen* (1991) 171 CLR 468. See also Sackville (1997:159); Keith Mason (1989); Constitutional Commission (1988:424-3).


Australia's Constitutional Convention
1998

George Winterton

The Convention to be held in Canberra during 2-13 February 1998 to debate the republican issue will be the sixth official Australian Constitutional Convention, or the seventh if the unofficial People's Federal Convention held at Bathurst in November 1896 is included. The term 'convention' — literally, a 'coming together' — has generally been employed in Australian politics to denote a meeting convened for the purpose of drawing up or amending a constitution, other governmental meetings being denominated 'conferences'. This usage has a long pedigree.

Antecedents

The English assemblies which restored Charles II to the throne in 1660 and declared the abdication of James II in 1689 were called 'Convention Parliaments' rather than true Parliaments, since they had not been summoned by the monarch. But the true laboratory of the Constitutional Convention is the United States. Indeed, R. R. Palmer (1959:214) considered the framing of a written constitution by a Convention 'the most distinctive work' of the American Revolution.

The first 'Convention' to frame a constitution was the 'Hartford Convention' of 1639 which produced the 'Fundamental Orders of Connecticut', considered 'the first written constitution to establish a popular government' (Edel, 1981:2). The English Convention Parliament of 1689 was paralleled by two revolutionary Conventions, comprising delegates from each town, held in Massachusetts that year consequent upon the overthrow of Governor Sir Edmund Andros (Jamieson, 1887:8-9). These revolutionary gatherings of 1660 and 1689 provided the model for constitution-making by several elected State Conventions called in the wake of the 1776 Declaration of Independence (Edel, 1981:3). In this way, 'the wild scion from the woods [the revolutionary Convention] was domesticated in the garden of the Constitution and made to subserve the purposes of regulated life' (Jamieson, 1887:11-12).

In several States the Convention acted also as the legislature, but that was not the case in New Hampshire (1778, 1781-83) and Massachusetts (1780) in which popularly elected Conventions were called for the specific task of framing a consti-
tution, clearly setting the pattern for later Conventions. Moreover, these States were distinctive also in submitting the constitution framed by the Convention to the electors for approval, with Massachusetts widening its franchise on that occasion to include all adult males (Hoar, 1917:4-7; Warren, 1937:347; Beer, 1993:309-10).

Undoubtedly the most important Constitutional Convention was the federal Convention, which met from May to September 1787 in Philadelphia, supposedly to revise the Articles of Confederation, but which ended up drafting a completely new constitution. The 55 delegates from twelve States (Rhode Island was not represented) were selected by their State legislatures, and their handiwork, the United States Constitution, was submitted for approval not to the electors themselves but to popularly elected State Conventions, a procedure which 'sought to reconcile wide popular approval with rational deliberation' (Beer, 1993:312).

The methods of amendment specified by the United States Constitution comprise a complex series of alternative procedures reflecting in various degrees the foundational principles of federalism, representative government and popular sovereignty, although direct popular approval is not included. Amendments must be proposed either by a two-thirds majority of both Houses of Congress or by a Convention called by Congress at the instance of two-thirds of the State legislatures. The former has been the sole method employed; although frequently advocated, no federal Convention has been held since 1787 (see Weber & Perry, 1989). Proposed amendments must be ratified by either the legislature or Conventions in three-quarters of the States, the appropriate mode being stipulated by Congress (US Constitution art V). To date, only one amendment has been ratified by State Conventions: the 21st (1933), which repealed the 18th.

One advantage of a national Constitutional Convention is that it would enable amendment proposals to be sent to the States for ratification without passing through the sieve of Congress. President Lincoln, for example, advocated this route in his first inaugural address (1861) because

it allows amendments to originate with the people themselves, instead of only permitting them to take, or reject, propositions, originated by others, not especially chosen for the purpose, and which might not be precisely such, as they would wish to either accept or refuse.

However, these pleas have so far fallen on deaf ears. But while the Constitutional Convention has been little employed at the federal level, the opposite is true of the American States where, as the Australian constitutional framers well knew, constitution-revision Conventions have been a frequent phenomenon since 1776 (see for example Deakin, in Australasian Federation Conference, 1890:248).

The first Australian federal ‘Convention’ was a meeting of representatives of the seven Australasian colonies and Fiji held in Sydney in November-December 1883. It led to the establishment of the Federal Council of Australasia, which New South Wales and New Zealand never joined. At the instance of Victorian Premier James Service, the gathering ‘grandiloquently called itself a convention’ (Garran, 1958:44),

No one has disputed the denomination of the next Convention, the National Australasian Convention held in Sydney from 2 March to 9 April 1891. It was attended by seven representatives from each Australian colony and three from New Zealand, all elected by their respective parliaments. American precedents — both federal and State — had been cited by Alfred Deakin in successfully moving for such a Convention a year earlier at the Australasian Federation Conference in Melbourne (Australasian Federation Conference, 1890:246-8). The Convention produced a draft national Constitution which is a close ancestor of the present Australian Constitution. But John La Nauze (1972:78) employed dramatic overstatement in asserting that 'the draft of 1891 is the Constitution of 1900, not its father or grandfather'.

When the States, especially New South Wales, failed to advance consideration of the 1891 draft Constitution, the federal movement hit the 'doldrums' (Garran, 1958:99) from which it was rescued by the Premiers' decision to convene a second, popularly elected, Convention and submit its handiwork to the electors for approval before transmission to Westminster for enactment. In Stuart Macintyre's (1993) apt continuation of the aeolian metaphor, with the decision to convene a popularly elected assembly 'the wind of public participation began to fill [the federal] sails'. As John Hirst (1996:42) has noted, the Corowa (1893) formula, which was endorsed at the January 1895 Premiers' Conference in Hobart — a popularly elected Convention to frame a Constitution to be approved by the electors before enactment at Westminster — was 'an amazing process ... Quite un-British'. But to what extent the post-Corowa federation movement can be considered 'popular', and whether federation was 'brought to life by the people', as Sir Robert Garran (1958:101) claimed, may be debated (de Garis, 1993; Macintyre, 1994; Macintyre, 1997:25-7).

Another Popular Convention?

The 1897-98 Australasian Federal Convention, which essentially produced the present Commonwealth Constitution, was the only Australian Convention to be popularly elected and the only one whose efforts were crowned with success in the sense of seeing its proposals implemented. The two 20th-century Conventions, those of 1942 and 1973-85, comprised politicians: Commonwealth and State representatives in the former, and Commonwealth, State, Territory and local government representatives in the latter. But none of the former's proposals, and few of the latter's, have been adopted at referendum. Indeed, as is well known, only eight of the 42 proposals put to referendum have been carried.

1 However, very few of the 1973-85 Convention's resolutions were submitted to referendum (in 1977 and 1984). The 1942 Convention was not primarily concerned with constitutional amendment, but
This rather dismal record has led to suggestions for reform both in the requirements for adopting a constitutional amendment and in the process whereby proposed amendments are formulated. The former include proposals to allow the States (and possibly also citizens), as well as the Commonwealth parliament, to initiate proposed amendments, and to liberalise the requirements for passage by requiring approval by a majority of electors in three, rather than four, States in addition to approval by a national majority of electors (see House of Representatives Standing Committee on Legal and Constitutional Affairs, 1997:57). However, since none of these proposals (which raise broad and important issues) has been implemented, they need not be considered further here.

Of greater interest here are proposals to alter the process whereby constitutional reform is considered. Apart from the two non-popularly elected Conventions (1942, 1973-85), extensive constitutional review has been undertaken by a Royal Commission (1927-29), a Joint Commonwealth Parliamentary Committee (1956-59) and a Commission assisted by five committees (1985-88). Yet, despite the high quality of the resulting reports, they have virtually nothing to show in the way of actual constitutional amendment. Hence it is understandable that some have wondered whether the successful formula of 1897-98 might not be revived, although a major difference, of course, is that proposed amendments must be approved not only by the electors, but must first secure the support of the Commonwealth parliament, or at least the House of Representatives and, therefore, the government. Moreover, it is difficult to assess how important popular election was to the success of the 1897-98 Convention, for its membership of leading colonial politicians was not significantly different from that of the 1891 Convention whose delegates were elected by the colonial parliaments. Nevertheless, a popularly elected Convention has been advocated by highly informed commentators such as Sir Robert Garran, Sir Robert Menzies and, more recently, R. D. Lumb (Garran, 1958:209-11; McMillan et al., 1983:357; Lumb, 1992, 1993-94). The concept was supported in earlier years by W. M. Hughes and Earle Page; the Hughes Government introduced a Bill for a partly elected Constitutional Convention in December 1921, but it was not proceeded with (Lumb, 1992:59-60).

Arguments in favour of a popular Convention include, first, that the election focuses public attention on the relevant issues and the proceedings of the Convention. In other words, it is a significant educative process. As an editorial in the Age noted on Australia Day 1993, a popularly elected Convention "would engage public imagination, broaden public involvement and raise public consciousness. It would rather with a temporary reference of powers by the States to the Commonwealth pursuant to section 51(xxvii) of the Constitution during the war and for post-war reconstruction.

2 Strictly speaking, approval twice with a three-month interval by either House suffices as a precondition for submission of proposed constitutional alterations to the electors by the Governor-General. But since there is no reason to consider the Governor-General's power as a reserve power, the approval of the government and, therefore, effectively of the House of Representatives, is necessary for a proposed alteration to be submitted to referendum: see Winterton (1983:212 n.152).
enjoy the sort of legitimacy that is lacking in government-appointed or self-appointed bodies examining constitutional change'. Even the non-compulsory postal ballot to elect half the delegates to the February 1998 Convention has raised public awareness of the republican debate, although it is too early to assess the educative value of that experience.

Second, the general public probably lacks confidence in its understanding of the issues, but is concerned to see them receive proper consideration, which the deliberations of a Convention especially elected for the purpose may satisfy. The ten days set aside for the 1998 Convention to debate the republic should alleviate public anxiety in this respect.

Third, a popularly elected Convention should reflect the diverse views of the electorate and their relative preponderance. Since proposed constitutional amendments must be approved at referendum, there is obvious value in exposing them to consideration by such a representative body. For similar reasons, it has occasionally been suggested that the members of a Constitutional Convention be chosen by random selection, like a jury (Mueller, 1996:315-16). Such an assembly might be 'the most likely to take into account the long-term interests of all future citizens and agree to a compromise' (Mueller, 1996:316). But, unlike the far larger samples selected for reliable public opinion polling, the membership of a Convention would seem to be too small for accurate reflection of the various shades of public opinion. The 1998 Convention will derive little benefit from this representative aspect of a popularly elected Convention because only half its membership will be elected and even those members were not elected by the compulsory voting which will presumably apply to the ultimate section 128 referendum.

Moreover, there are negative aspects to a popularly elected Convention, apart from the cost of the election (around $50m). First is a concern that, as an American commentator noted, 'particularly in this age of television and mass-media dominance' the Convention will be 'filled with pop singers and athletes who, among other deficiencies, lack expertise on constitutional matters' (Mueller, 1996:315).

Such concern has been voiced regarding the candidates for election to the 1998 Convention, with Greg Craven reported as complaining that their quality was 'a matter of deep national embarrassment', especially when compared with the Conventions of the 1890s:

Where is Barton, where is Deakin, where is Griffith, where is Playford, where is Andrew Inglis Clark, where is George Reid, where is Henry Parkes, where is Henry Higgins, where is Isaac Isaacs and where is Josiah Symon, where is Sir John Forrest? It goes on and on and on and on.

(Rintoul et al., 1997)

However, this anxiety appears unwarranted. While the constitutional expertise of leading candidates for the Convention may well be inferior to that of 1897-98 delegates such as Barton, Deakin, Isaacs and Higgins (Griffith and Inglis Clark did not attend that Convention), no one would propose popular election to assemble a
body of experts in constitutional law, or any other subject for that matter. But the overall quality of leading candidates is no cause for disquiet. Constitutional expertise can be provided by the Commonwealth government, and indeed there is considerable political, legal and historical expertise among the appointed delegates and parliamentary representatives at the 1998 Convention.

A more significant concern is that the division of opinion in the community will be reflected among the Convention delegates, leading to deadlock in its proceedings. The 'boycott' of the American federal Convention of 1787 by leading 'anti-federalists' such as Patrick Henry undoubtedly facilitated consensus at that (non-popularly elected) Convention (Mueller, 1996:323). Thus, a distinguished Canadian constitutional scholar, Edward McWhinney (1981:33), has remarked:

For its most effective operation, a constituent assembly would seem to require to be elected against a background of an already existing, and continuing, societal consensus as to the nature and desired direction of fundamental political, social and economic — and hence constitutional — change.

For this reason, a Constitutional Convention on the republic would have had a greater prospect of achieving consensus if it had been preceded by an indicative (not legally binding) plebiscite on the question whether Australia should sever its constitutional links with the Crown, and perhaps also on the appropriate method of selecting a republican head of state. Assuming a pro-republic vote in the plebiscite, the Convention would then have confined its attention to the details of republican government within the constraints imposed by the results of the plebiscite, and there would have been little point in monarchist candidates standing for election. Such a course of action was widely advocated. The Australian Labor Party (ALP) and the Sydney Morning Herald urged an indicative plebiscite followed by the framing of a proposed constitutional amendment by a Commonwealth parliamentary committee; indeed the ALP and the Australian Democrats introduced a Bill for a plebiscite in the Senate, the Plebiscite for an Australian Republic Bill 1997. New Zealand provides a model in this respect, MMP having been adopted following an indicative plebiscite (called an 'indicative referendum') among several electoral systems which led to the enactment of legislation which came into effect upon approval by the electors at a referendum.

---

3 The term 'plebiscite' (rather than 'referendum') is employed in this context to indicate its legally non-binding nature. Whether a majority of electors in four States as well as a national majority would be required for an affirmative result should be specified in advance in the legislation providing for the plebiscite and prescribing its consequences.

4 By commentators including Sir Zelman Cowen, Sir Anthony Mason, Cheryl Saunders, Charles Sampford, Malcolm Turnbull, Peter Howell, Christopher Pearson, Helen Irving, Frank Brennan and Crispin Hull, and by the Australian, the Age and the Canberra Times.
However, the contrary view — that the public education provided by a Convention is a necessary prelude to any plebiscite — also carries considerable weight, and was probably best argued by Bob Ellicott (1996):

Doubt, suspicion, an absence of the important detail, in the hands of skilful campaigners, could easily lead people to vote no in an early plebiscite when their ultimate informed view might well be quite different... The holding of some form of convention is probably the best way of informing people of the issues involved and the merits of the solutions available.

In any event, the Howard Government ignored calls for an indicative plebiscite from the ALP, the Australian Democrats, the Australian Republican Movement and others, and instead honoured its 1996 election promise by establishing a Convention, half popularly elected, to consider not only the form and timing of an Australian republic, but also the preliminary question of whether links with the Crown should be severed.

The 1998 Convention

The 1998 Convention will comprise 152 delegates of whom half were popularly elected in a non-compulsory national postal ballot held from early November to 9 December 1997. About 47 per cent of eligible voters participated. The ballot was conducted on State/Territory-wide electorates with State representation roughly proportional to population. The election was heavily contested, notwithstanding a $500 non-refundable fee payable by each candidate. In New South Wales 174 candidates contested 20 seats; Victoria 158 for 16 seats; Queensland 129 for 13 seats; Western Australia 59 for nine seats; South Australia 38 for eight seats; Tasmania 23 for six seats; Australian Capital Territory 16 for two seats; and Northern Territory 12 for two seats (Sydney Morning Herald, 10 October 1997). Many candidates stood as part of a monarchist or republican ‘ticket’, but many others stood as independents. The ballot paper allowed voters to choose either a group or a number of candidates equal to that State or Territory’s representation at the Convention. Members of the Commonwealth, State and Territory parliaments and appointed delegates were ineligible for nomination as elected delegates and conse-

---

5 The States’ seats are essentially proportional to the size of their Commonwealth parliamentary representation (combined House and Senate membership).

6 It is difficult to see why the deposit was non-refundable. If the purpose of the deposit was to discourage candidates with little prospect of election from standing, there is no reason why usual electoral practice could not have been followed, providing for return of deposit to successful candidates and those receiving a specified proportion of first preference votes – 4 per cent in Commonwealth elections: Commonwealth Electoral Act 1918 (Cth) section 173. If the Commonwealth simply wanted to recoup some of the cost of conducting the ballot, this would seem a relatively trivial benefit (relative to governmental finances), even considering the large number of candidates, for the cost in discouraging citizen participation in government.
quently ineligible for election (Constitutional Convention (Election) Act 1997 (Cth) section 25).

The non-elected members comprise 36 delegates appointed by the Commonwealth government and 40 members of Australian parliaments. The former include a youth delegate from each State and Territory (five women and three men) and 28 others (16 men and 12 women), both republicans and monarchists or at least anti-republicans, including three former viceregal representatives, the present and former chairs of ATSIC, local government representatives, and two bishops. Their appointment was announced on 31 August 1997. The parliamentary representatives include 20 from the Commonwealth parliament and 20 from the other Australian parliaments. The Commonwealth parliamentary representation is roughly proportional to party strength in the parliament: twelve government members, six Opposition, one Australian Democrat and one independent. Each State has three parliamentary representatives: the Premier, the Leader of the Opposition, and another member nominated by the Premier (a Minister in all States except South Australia and Tasmania, where the leaders of the Australian Democrats and the Greens, respectively, were nominated). The two self-governing Territories are represented by their Chief Ministers.

The Convention’s composition — essentially half popularly elected, one-quarter parliamentary and one-quarter reasonably broadly based appointed delegates — would seem to auger quite well, for several reasons. First, it has long been considered appropriate that the forum for debating constitutional reform should reflect the constituencies which must ultimately be persuaded to approve proposed amendments: the Commonwealth parliament and the electors, both nationally and in four States. This suggests that a Constitutional Convention should include Commonwealth and State parliamentarians together with delegates representing a wide range of community interests and viewpoints, whether or not popularly elected. Thus, in 1921 the Hughes Government introduced a Bill (ultimately not proceeded with) to establish a Constitutional Convention comprising 75 popularly elected delegates and 36 delegates, half of whom would be nominated by the Commonwealth and half by the States, although Prime Minister Hughes expressed willingness to dispense with the nominated portion as advocated by Country Party leader Earle Page (Lumb, 1992:59-60). A similar proposal by three commentators, including Gareth Evans, in 1983 envisaged a Constitutional Convention comprising 32 popularly elected delegates and 80 parliamentarians, including 16 from the Commonwealth, eight from each State, two from each self-governing Territory and two local government representatives from each State (McMillan et al., 1983:366-7).

Second, each of the three groups participating in the Convention should contribute a different dimension to its proceedings. The popularly elected delegates will primarily bring passion and commitment to the various shades of public opinion on the republican issue; the parliamentary representatives embody political experience and, one hopes, a facility for compromise, with the Commonwealth government presumably determined to prevent the Convention it proposed and established from descending into acrimonious deadlock and shambles; and the ap-
pointed delegates may offer a less passionate range of community views together with some legal and governmental experience. They may exert a moderating influence on the passionate commitment of the elected delegates; but this may represent an unduly optimistic perspective.

The Convention will debate whether or not Australia should become a republic; which republican model should be put to the electorate to consider against the status quo; and in what time frame and under what circumstances any change might be considered (Howard, 1997:3061). The Howard Government has undertaken to put to referendum under section 128 of the Constitution any proposal endorsed by a ‘consensus’ at the Convention and, if no consensus eventuates, to hold an indicative plebiscite on the republican issue. In either case, a vote on the republic issue is promised ‘by the end of the year 2000’ (Department of the Prime Minister and Cabinet, 1997:1).

The government has not defined what it considers a ‘consensus’, although the responsible Minister, Senator Nick Minchin, has suggested that it means ‘not simply a bare majority, but a clear majority’, declining to be more specific (Taylor, 1996). In view of the presence at the Convention of substantial numbers of both proponents and opponents of the republic, a consensus either for or against severance of constitutional links with the Crown appears unlikely. However, there is a greater prospect of broad agreement on the issues of the ‘how and when’ of an Australian republic. Indeed both Prime Minister Howard and Convention chair Ian Sinclair have undertaken to work to that end (Gordon & Hawes, 1997; Millett, 1997).

The Convention’s outcome will, to some extent, depend upon the agenda and procedures adopted at the Convention, on which the delegates should have the opportunity to vote. No more than three of the Convention’s ten days should be allocated to consideration of the initial question ‘whether or not Australia should become a republic’ and, on the other issues, broad principles should be debated rather than detailed proposed constitutional amendments. It would be calamitous for the Convention to become bogged down in debating the detailed language of proposed constitutional provisions governing the method of appointment or removal of the Head of State or in attempting to codify the powers of the office. That task should be assigned to expert drafters at the conclusion of the Convention.

The Future

The performance of the tripartite February 1998 Convention will be closely analysed for years, since it is likely greatly to influence the future direction of Australian constitutional reform. If it breaks down in deadlock, the combination of popularly elected delegates, parliamentarians and appointed citizens is unlikely to be repeated.

---

7. In August 1996 the Western Australian Commission on Government recommended the establishment of a Convention to review the State Constitution, half the delegates to be popularly elected and half appointed by Parliament, to enable ‘an appropriate balance [to] be struck between expertise and community support’: Commission on Government (1996:108).

On the other hand, if broad agreement is reached on a suitable model for republican government, such a Convention will probably be employed for wider constitutional debate.

The Howard Government initially envisaged that the Convention would consider several issues in addition to the republic, including four-year Commonwealth parliamentary terms; the federal balance of power, including the external affairs power; admission of new States, especially the Northern Territory; and constitutional recognition of local government. Further issues mentioned were the qualifications for membership of the Commonwealth parliament, appointment of High Court justices, and State-initiated constitutional referenda (Minchin, 1996:3-4). However, while there is no difficulty in assigning any number of issues for consideration by an appointed Convention including parliamentarians, the position is more complex when the Convention is popularly elected or includes a popularly elected component. Candidates will presumably stand for and against each major issue to be debated at the Convention, but will have no particular expertise or representative point of view on other issues. Hence, the representative function of a popularly elected Convention will be most effective if the Convention is assigned a relatively small number of subjects.

However, some commentators have suggested that a Convention be elected periodically to consider appropriate constitutional reform. Twenty years ago, R. D. Lumb (1977:101) urged that such a gathering be convened 'at regular intervals'; and Cheryl Saunders (1994:119) has more recently suggested that 'regular constitutional reviews, elected or otherwise' be conducted every ten years. But frequent constitutional review could undermine the sense of stability and permanence necessary for effective constitutional government. Hence others have suggested constitutional review at less frequent intervals, such as every 25 to 30 years (see Edel, 1981:9; Mueller, 1996:325); indeed, the Polish Constitution of 1791 expressly required the convening of an 'extraordinary constitutional diet' every 25 years (art VI, in Blaustein & Sigler, 1988:75). Periodic constitutional review is undoubtedly necessary for, as Thomas Jefferson — 'certainly [no] ... advocate for frequent and untried changes in laws and constitutions' — eloquently noted,

[L]aws and institutions must go hand in hand with the progress of the human mind. As that becomes more developed, more enlightened ... and manners and opinions change with the change of circumstances, institutions must advance also, and keep pace with the times. We might as well require

9 Cf. the proposal by Representative Ebenezer Herrick of Maine in 1824 that 'amendments [to the US Constitution] be considered every tenth year but be prohibited in the intervals between decades' (Edel, 1981:9).

10 One lesson Brian Galligan (1995:132) draws from Australia's dismal constitutional reform record is that 'progressive elites and federal governments should stop pestering the Australian people with referendum proposals that are unnecessary or that attempt to centralise power in Canberra'.

11 Some might say this will be the fare at the February 1998 Convention!
a man to wear still the coat which fitted him when a boy, as civilised society to remain under the regimen of their barbarous ancestors. (Letter to Samuel Kercheval, 12 July 1816).

However, constitutional review is probably best undertaken, not at some fixed interval, but when a considerable weight of public opinion acknowledges that the issue requires attention. Moreover, the method of review should be carefully adapted to the subject under consideration: for example, an issue of particular interest to the States, such as federal financial relations or the potential republicanisation of State governments, requires greater State parliamentary participation than one in which the States have little direct involvement, such as four-year Commonwealth parliamentary terms or the qualifications for membership of the Commonwealth parliament.

The tripartite constitution of the February 1998 Convention is a novel experiment; as has been well said, it is a ‘most unconventional convention’. Its performance will be closely studied both in Australia and abroad for, if the combination of popularly elected representatives, parliamentarians and government-appointed citizens proves an effective one, it may well provide a model for similar bodies in other countries contemplating constitutional change, such as Canada, New Zealand and even the United Kingdom. Australia has tried government-appointed commissions, parliamentary committees and parliamentary Conventions, all to no avail in the sense of actual constitutional alteration. The tripartite February 1998 Convention may well represent, to borrow President Lincoln’s aphorism in his Annual Message to Congress of 1 December 1862, the ‘last best hope’ for Australian constitutional reform.

References
Australasian Federation Conference (1890), Official Record of the Proceedings and Debates, Government Printer, Melbourne.


Department of the Prime Minister and Cabinet (1997), Republic - Yes or No?, Canberra.

---

12 In an unpublished paper by Mr James B. Kelly of Mt Tambourine, Queensland, from which the author has derived considerable benefit.


*The helpful comments of two anonymous referees are gratefully acknowledged.*

*Editorial note: In the popular election of delegates to the 1998 Constitutional Convention, republican candidates won a majority (56.4 per cent) of the total votes cast and a majority in four States. Of the 76 elected delegates, 27 were monarchists, 27 were affiliated to the Australian Republican Movement, 19 were republicans with other affiliations, and two were of unknown affiliation. The 76 appointed delegates comprised 17 constitutional monarchists, 30 republicans, and 29 undeclared.*
"Excellent, diverse and wide-ranging."
— IRVING LOUIS HOROWITZ, Rutgers University

Transcending the all-too-common politicization and superficiality of public policy research and debate, The INDEPENDENT REVIEW is the interdisciplinary, quarterly journal devoted to political economy and the critical analysis of government policy. Edited by Robert Higgs, The INDEPENDENT REVIEW is superbly written, provocative, and based on solid scholarship.

Ranging across economics, political science, law, history, philosophy, sociology, and related fields, The INDEPENDENT REVIEW boldly challenges the politicization and bureaucratisation of our world, featuring in-depth examinations of current policy questions by many of the world’s outstanding scholars and policy experts. Undaunted and uncompromising, this is the journal that will pioneer future debate!

"The Independent Review is of great interest."
— C. VANN WOODWARD, Pulitzer Prize-Winner, Yale U.

"The Independent Review is distinctive in badly needed ways."
— LELAND YEAGER, Professor of Economics, Auburn Univ.

"The Independent Review is an exciting new journal."
— JOHN DIGGINS, Prof. of History, City U. of New York

"... does not accept the pronouncements of government officials at face value, nor the conventional wisdom over serious public problems."
— JOHN MACARTHUR, Publisher, Harper’s Magazine

"The Independent Review avoids faddish pyrotechnics while seeking objectivity and high standards."
— TIMUR KURAN, Prof. of Econ., U. of Southern California

"A great team, economic logic and tough analysis."
— BRUCE YANDLE, Former Executive Director, F.T.C.

In Recent Issues:
The Ambitious, Accommodative Adam Smith
— SALIM RASHID
Preventing Banking Crises in the Future
— GEORGE G. KAUFMAN
What is Living and What is Dead in Classical Liberalism?
— CHARLES K. ROWLEY
The End of Welfare and the Contradiction of Compassion
— STEPHEN T. ZELIAK
The International Monetary System in Retrospect
— LELAND B. YEAGER
The Origins of the War on Drugs, 1984-1989
— BRUCE L. BENSON AND DAVID W. RASMUSSEN
Culture, Politics, and McCarthyism
— IRVING LOUIS HOROWITZ
Market-Based Environmentalism and the Free Market
— ROY E. CORDATO AND PETER J. HILL
Why the Great Depression Lasted So Long and Why It Ended
— ROBERT HIGGS
The Case Against Psychiatric Coercion
— THOMAS S. SZASZ

Feature Articles
Book Reviews
Special Features
Figures and Tables
160 Pages per Issue
March, June, Sept. & Dec.
Annual Index
Plus more . . .

Individual Subscriptions:
$27.95/1 year, $52.95/2 year
Institutional Subscriptions:
$49.95/1 year, $94.95/2 year
International orders:
Add $28/subscription shipping

Order Toll-Free (U.S.)
800-927-8733
Phone: 510-632-1366
Fax: 510-632-1366
review@independent.org
http://www.independent.org

The INDEPENDENT INSTITUTE
Dept. 8B
134 Ninety-Eighth Avenue
Oakland, CA 94603 U.S.A.

Look for The INDEPENDENT REVIEW on better newsstands and in bookstores!

Reviewed by David Preston

Michael Jones is a senior lecturer in management at The University of Canberra and author of *The Australian Welfare State*, whose first edition was published in 1980. His new book on social welfare in New Zealand is a carefully crafted text which is likely to feature on university reading guides for some time. As a summary of key issues involved in the contentious area of welfare reform, it is notable for its breadth of coverage and its balanced presentation. Jones has an encyclopaedic knowledge of the literature in the field, and of the sometimes ambiguous research evidence which confronts those who want clear and simple answers.

The bibliographies which follow each chapter are impressive, with New Zealand and Australian material set in a wide international context. Jones is careful also in his citation of evidence to represent a diversity of views. At times, the attempted coverage of issues seems a little ambitious for a book of just under 200 pages, even when restricted to cash benefits. However, the breadth of the references cited makes this a good basic text for those who wish to pursue topics in greater depth elsewhere.

The text is divided into three main sections. The first picks up several key themes whose implications are sketched out here and in some cases followed up in more detail later. Jones identifies the unusual nature of the New Zealand welfare system (and the related Australian model) with its predominant reliance on a single tier of tax-funded flat-rate benefits, the key distinction being between the highly targeted income-testing regime for the working-age population and the non-contributory universal pension for the elderly. He sets this in the context of the wide variety of models operating around the world. He also stresses the contrast between the limited scale of welfare reform so far undertaken in New Zealand and the radical measures taken in some sectors of the economy. Most of the more detailed discussion in this section deals with reform issues related to the working-age population.

Jones’s discussion of options and issues is more thoughtful than some other recent tracts in the field. In particular, he picks up the ‘four types of welfare intervention’ analysed by Deborah Mabbet, as well as Esping Anderson’s more frequently cited ‘three worlds’ or models. Mabbet’s approach identifies also the hidden redistribution effected through trade protection and cross subsidisation. The
demolition of this system as part of the economic restructuring of New Zealand has left many previously employed workers with limited job skills in a difficult transitional situation, and created a new type of dependence problem for the welfare state.

Although he sets New Zealand's problems clearly in an international context, Jones is cautious about the applicability of other countries' approaches to resolving them, particularly through social insurance. He does not see any easily importable solutions, but notes that New Zealand's system 'became trapped in its past' (p. 43) when confronting the problems of the 1970s. But despite its title, Jones's book is not a pamphlet with a set of final answers.

The second section of the book deals with poverty and inequality. At times it seems to move away from the main themes of welfare reform, but the focus is restored as Jones picks up most of the key themes in the recent poverty debate, and relates them to the key issues that a reform strategy must deal with. He notes the ambiguous nature of most definitions and measures of poverty in developed countries, and also the lack of good statistics in New Zealand on how individuals' economic situations change over time. He also draws attention to the high ratio of benefits to wages, although the comparison is with equivalent social assistance elsewhere rather than with what is provided under contributory earnings-related schemes. Jones notes the recent trend towards greater inequality in income distribution in most OECD countries; and he cites the key economic factors identified by other academics, notably technology, trade competition, demand shifts and migration inflows from poorer regions, together with the interface problems and dynamics of the welfare systems themselves and of the groups they cater for. He raises but does not resolve the issue of the extent to which current trends are transitional rather than structural.

The third section of the book deals with issues of an aging society in New Zealand and elsewhere. Jones focuses on the major problems this poses for the design of pension systems and the cost of other support systems for the dependent elderly, such as health and rest-home care. Jones suggests that New Zealand, in contrast to the stricter Australian model, 'treats the aged as high class citizens' (p. 174). He notes that whereas New Zealand used to have a low proportion of elderly in the population by developed country standards, it will have a very high proportion of elderly in the 21st century. The former situation allowed the luxury of a generous tax-funded universal pension system at an early retirement age; but the future poses significant adjustment problems since New Zealand's current policies are 'a recipe for severe intergenerational conflicts after 2030' (p. 161). Jones's proposed reforms for the state-funded New Zealand Superannuation will come as no surprise to those who have followed the recent debate. They include lower pension-wage ratios, higher pensionable retirement ages, and means tests for assistance. These are essentially the same options as those advanced in mid-1997 by the government's Periodic Reporting Group in 1997, though the perspective is more pessimistic.

Jones also deals with another recent policy preoccupation in New Zealand: the development of a fully funded contributory retirement pension system. He judges
that Australia's move in this direction has been, on balance, a desirable development, despite its design problems and especially the cost of its tax concessions, which in 1994-95 amounted to the equivalent of 71 per cent of Australian pension outlays. On balance, Jones supports compulsory provision, and consider that 'If New Zealand rejects a funded individual retirement system it will probably have to make its retirement benefit system less attractive' (p. 179).

The conclusions Jones reaches on suggested future directions for the New Zealand welfare state as it applies to the non-aged are unsurprising and carefully nuanced. He notes that the shift to a more internationally competitive economy in New Zealand has had 'limited impact on the scale of social dependency' (p. 183), and that 'many of the non-aged dependent seem strangely insulated from broader economic changes and expanding employment opportunities'. However, he sees 'no simple solutions' (p. 184).

Rather than proposing specific policies or programs for the non-aged dependent, Jones gives indicative support for several broad directions of change. A key underlying theme is the idea of the 'developmental state' rather than the 'welfare state', the former of which involves policies which equip people for self-sufficiency rather than supporting passive dependence. This leads naturally towards making work a much larger ingredient in the policy mix. But Jones is less specific about how increased work requirements for a variety of beneficiaries is actually to be implemented. He discusses time limits on benefits, without reaching a clear conclusion. He notes that the Swedish model of active manpower policies focusing on training and job provision 'is expensive' (p. 186). He suggests that New Zealand should look at more significant work requirements for sole parents, without being very specific about what these should be.

In essence, Jones hands the problem back to the New Zealand authorities to solve. He ends with a hopeful endorsement of the capacity of New Zealand governments and its policy advisers to successfully address the key issues, and suggests that New Zealand 'once it gives the welfare state enough critical attention, will undoubtedly develop a reform strategy suited to its history, values, and particular social and economic circumstances' (p. 188).

Let us hope that he is right.

David Preston is a former General Manager of the New Zealand Social Policy Agency. In 1997 he was a member of the design team for the New Zealand Compulsory Retirement Savings Scheme proposal, and Technical Adviser to the Independent Referendum Panel.
As a relative newcomer to local government from a research background, I relished this opportunity to review a book dealing with the issues and controversies surrounding reform in the sector. Local government has turned out to be a mixture: atomised industry with attendant employment and corporate governance concerns; a community advocate with a culture of grass-roots activism; and a participant in some of the institutions of government striving for greater recognition as the third sphere of government in the Australian federation.

Local government is experiencing an apparently unparalleled period of change and self-examination. The editors of the publication under review have brought together a collection of essays seeking 'to document the sweeping changes which are occurring in the local government sector and place them in the wider context of the transformation of the Australian economy'. They recognise that this is a challenging task given the 'paucity of material on the metamorphosis of Australian local government and the idiosyncratic dispersion of existing sources' (p. vii).

The contributions are organised into three parts: structure and processes; reform and renewal; and the future. The first part gives an overview of the institutional basis of local government as a self-standing entity, its relationship with the other spheres of government, its management ethos and its governing legislation. For these themes in particular, the genesis of the publication during the period of the establishment of the Coalition Government and after the more predictable patterns and relationships of the Labor administrations was unfortunate. Several authors alluded to the possibility of a change of direction under a new federal government; some of their conclusions have had to be relegated to the class of might-have-beens.

Neil Marshall’s introductory article embodies an optimistic assessment of local government as a player in the federal compact based on observations of growing confidence, increasing representation in national decision-making forums and the crowning achievement of an Accord in 1995 with the Commonwealth. He could not, of course, have foreseen that by the time of the collection’s publication in February 1997, the Accord would have become completely irrelevant to Commonwealth policy and its replacement effectively abandoned. As well, over the intervening period, a National Commission of Audit questioned the need for a ‘direct’ financial relationship between the Commonwealth and local government; the Council of Australian Governments, incorporating local government, seems increasingly unlikely to continue as a peak intergovernmental forum; and the Building Better Cities and Regional Development programs have become victims of a prevailing mood of fiscal rectitude.
To me, having experienced at first hand only a period of relative decline in relations with the Commonwealth, the optimism of the authors could appear misplaced. Were their assessments merely a reflection of better times for local government, they might have shown what can be achieved in the 'right' environment. But the real question for local government is whether effective intergovernmental relations can ever improve in the long run through the vagaries of the political cycle.

Ralph Chapman, in his contribution on intergovernmental relations, argues that the role and sophistication of local government will improve in the federal system. He claims that the Accord was not about funding, which would have been ephemeral, but about the growing confidence of local government. He identifies national constitutional recognition as the critical element in the process of assured legitimacy and relevance: a view enthusiastically advocated by most people involved in local government.

Ed Wensing provides an alternative perspective on the kind of lasting change needed to make local government less dependent on the vagaries of politics. For him, greater influence and relevance have come from the grass roots, through the interaction of measures such as the now ubiquitous general competence powers which increase autonomy and flexibility, with councils playing the role of integrated managers in their local areas. He argues that councils have been able to 'assume a stronger leadership role in local affairs, accept additional or modified functions and be more proactive in relations with State or Commonwealth agencies' (p. 102).

Also in the first part of the book, Judy McNeill explores the history of local government from humble and mixed beginnings and traces its development as the third sphere of government. Her analysis of the plethora of complex functions it performs is vitiated by a paucity of data. The Australian Bureau of Statistics data enumerate, in the finest detail, outlays on functions such as public order and safety, welfare and housing, and community development, each of which may include at the local government level up to ten individual services, none of which predominates. The Australian Local Government Association continues to press for consistent and meaningful data on the basis of which the exact contribution and focus of local government can be assessed. These numbers would surely highlight local government's contribution to national and regional goals, for example in environmental programs or community care, assisting local government's claim as a true federal participant in these areas.

The second part of the book contains assessments of the various reform drives. It is interesting that few of the current reform themes included are internally driven. The contribution on 'reforms' in community and cultural participation fits awkwardly into the list, which should have included environmental management.

Of the reforms affecting local government, the most controversial and disruptive for councils have been amalgamations and competition policy. Chris Aulich's contribution on the latter is in my view the most original and valuable in the collection. He presents a useful survey of the issues and classifies the theories underlying the different types of competitive reforms and their likely outcomes, including their chances of increasing the efficiency or effectiveness of service delivery. His discus-
sion of policies applying to local government under national competition policy (NCP) and competitive tendering makes it clear that local government has over many years embraced competitive reforms in the normal course of events and as a consequence of attempts to improve services and/or reduce their cost of delivery. Aulich argues that compulsory competitive tendering (CCT), as mandated in Victoria and under consideration elsewhere, is a not entirely logical or desirable variant of contracting practices already well-established in local government. The chapter also serves to dispel the persistent (especially in Victoria) belief that CCT is either a response to or the ultimate expression of NCP.

Anne Vince's contribution on amalgamations is less focused on theory and principles than is Aulich's. The relevant points are nevertheless brought out in the case studies of Victoria, Tasmania, Western Australia and New Zealand. It has always been clear that the Victorian amalgamations, and the events which led up to them, were more a political response to the likelihood of resistance to reform whose gains have been exaggerated to balance the high costs. Vince concludes that 'the Victorian experience is also evidence that poorly planned, hastily executed amalgamations which do not involve intense consultation with councillors, staff and communities of amalgamating councils can result in long-term organisational problems which adversely affect service delivery' (p. 160). The account of the Tasmanian experience points to the administrative success of amalgamations conducted in the opposite manner to Victoria's: with consultation, certainty and transparency. But in neither case, nor in Western Australia or New Zealand, can it be shown that amalgamation, whether forcibly or 'organically' achieved, results in better local government in any sense of the word.

The lone contribution in part three of the collection, by Peter Self on the future of local government, brings together many of the strands of the earlier authors and views the various changes in a broad context. Two themes emerge from this treatment. One is that there is the potential for much to be lost were local government required increasingly to emulate the State and federal spheres in drives for efficiency and small government, especially where local government remains local and sensitive to community needs. The other is that the current crop of reforms may not be the greatest forces for change in local government in the longer term: they are likely to be easily surpassed by the demands of demographics and the scope of future roles and responsibilities.

Local government clearly displays a great diversity of forms and functions. The editors have met the challenge of finding contributions with a linking theme and appealing to a lay readership.

Chris Bell is Policy Manager (Finance and Microeconomic Reform) of the Australian Local Government Association.
A Multidimensional Framework for Assessing CTC

Graeme Hodge, Contracting out Government Services: A Review of International Evidence, Montech Pty Ltd, Melbourne, 1996

Reviewed by Stephen Rimmer

This book's genesis lies in a project undertaken for the Local Government Research Foundation of South Australia and the author's doctoral research program. It is based primarily on a review of 245 Australian and international studies, published and unpublished, of competitive tendering and contracting (CTC), especially those dealing with its use in local government.

The author, a Course Leader in the Faculty of Engineering at the Royal Melbourne Institute of Technology, argues that CTC, as a form of privatisation, is implemented to achieve a narrow range of objectives, such as enhanced efficiency, lower costs and smaller government. These objectives are then compared to the author's 'multi-dimensional' framework for assessing 'privatisation activities'. This includes some 30 economic, social, democratic, legal and political performance indicators, ranging from 'investment levels' and 'demonstrated reduction in corruption' to 'violations of professional standards' (pp. 10-11). With this wide range of measures, the author seeks to assess both the direct and the indirect social, economic, political and legal impacts of CTC.

Hodge's findings are generally similar to those of other recent and wide-ranging studies of the impact of CTC, such as Rimmer (1993) and the Industry Commission (1996). Thus, he finds that the use of CTC usually — although not always — leads to cost reductions (typically of between 9 per cent to 14 per cent) and that it is not usually associated with reductions in service quality. The author speculates that CTC may lead to an increased risk of corruption (he cites the US Defence Department) and also a reduction in accountability.

The book concludes with a discussion of possible limits to the use of CTC, such as a lack of competition for local government contracts. It argues that decision-makers should use wide-ranging cost-benefit analysis of both the direct and indirect impacts of CTC. It also discusses various issues that could be addressed prior to CTC, such as benchmarking and establishing the cost of service provision.

The book is ambitious in its scope and covers a lot of ground. But at times the discussion is superficial and based on unsubstantiated assertions. For example, the author claims that:

There is considerable evidence to suggest that privatisation is an ideology or a belief structure which is pursued by some governments as an end in itself. Even in the United Kingdom the economic and market rationale for Thatcher's reforms came several years after the reforms were initiated. Of further importance is also the fact that to the extent that privatisation is an ideology, then it is possible that two different reviewers — even when both
are reviewing the same empirical findings — may well come to quite different conclusions. (pp. 4-5)

Yet none of the allegedly ‘considerable’ evidence for these assertions is cited.

The book also displays confusion about key terms. For example, after noting various definitions of privatisation and acknowledging key conceptual differences between ownership, competition and regulation, the author claims that any of these reforms can be described as ‘privatisation’. Hence, there is a ‘need to view contracting out and competitive tendering as part of the privatisation movement (p. 4). But there are important distinctions between CTC and privatisation. For example, under CTC governments retain control over and remain responsible for service provision. Indeed, many reforms designed to increase competition or review regulations that impact on business have nothing to do with privatisation.

Hodge argues that a wide range of performance indicators should be used to measure the direct and indirect impact of CTC. But the analysis often ignores the indirect impacts of CTC. For example, the author claims that CTC leads to overall reductions in employment, and rejects the Industry Commission’s (1996) conclusion that, while CTC may lead to a reduction in employment for provision of a particular service, it is likely to lead to an increase in overall economy-wide employment levels (p. 29). But here he has simply failed to distinguish between first-round effects of CTC on employment and subsequent ‘flow-on’ effects. The initial reduction in public-sector employment for a particular service will be offset to some extent by increases in private-sector employment resulting from the transfer of employees to, and the employment of others by, external contractors. Savings from CTC can be used to increase the quantity or quality of service provision, increasing employment for that service. Furthermore, savings from CTC can be used to reduce levels of taxation, so lowering business costs and encouraging economy-wide employment.

Despite these shortcomings, the book’s review of the literature, the largest yet undertaken, makes an original and useful contribution. While it broadly confirms the results of other recent reviews of the impact of CTC, it also enhances our understanding of the process in some areas, such as its impact on service quality.

References


Stephen Rimmer is a Director at the Office of Regulation Review, which is part of the Industry Commission. The views expressed here are those of the reviewer and not necessarily those of the Commission.
East Asia’s Achilles Heel


Reviewed by Eric Jones

The targets of this book are claims about the respective merits and likely futures of East Asian and Western economic, social and political systems. These claims are that the West is sluggish and in social decline whereas East Asia is meritorious and unstoppably successful. The propositions come together in the utterances of East Asian politicians and bureaucrats who excoriate the West, vaunt the superiority of ‘Asian values’, and perceive the relationship between the two as zero-sum. The terms they use are deeply resentful. Within the English-speaking West, home-grown illiberal and anti-capitalist intellectuals have engendered a mood in which such criticisms are accepted almost at face value. This has been made easier because few people who contribute to public debate seem to understand the nature of the spontaneous order underlying Western economies or display much grasp of comparative history.

Fads have not helped. We have witnessed extremes of both Japanolatry and dismay at Japanese competitiveness, and now seem in the midst of shifting from ‘China fever’ to fear of China. Meanwhile, Western liberalism’s tradition of self-criticism, which is valuable, even vital, because it encourages self-correction, has focused attention excessively on the West’s flaws at the expense of its virtues. We have also all heard the ‘declinists’ gloat that the United States is finished — though the good news may be that the success story of American economic restructuring is at last starting to be recognised.

The perverse and self-interested assaults on the West from outside have seldom, if ever, been contested as comprehensively as in the present book. It is a brave book in the context of the current asymmetry which makes it safe to deprecate Western failings whereas mention of the oppressions and horrors elsewhere automatically brings charges of cultural imperialism. In the teeth of this onslaught and for fear of losing a share of the Wealth of the Indies (so to speak), the defence of incipient democracy in Asia is close to being abandoned to what Christopher Lingle calls a ‘pre-emptive cringe’. The indispensable freedom of expression so hard-won in the West is similarly being surrendered by publishers willing to truncate their lists for fear of losing Asian sales. Western leaders and businesses are being flattered by the ‘New China Lobby’ and the like into ignoring not merely ethical failings but the actual risks of doing business in countries where law and hence the enforcement of contracts is unreliable.

Lingle, an economics professor at Case Western Reserve University, Cleveland, will have none of this. He mounts a spirited defence of liberalism and free markets within an argument that the institutions of East Asia will prove incapable of sustaining high rates of economic growth. His writing is slightly repetitious but always temperate.
and lucid. He meets fire with cold water, approaching what he sees as the worm in the apple of the ‘East Asian miracle’ from so many angles that in the limited space of a review one must pass over most of them. I shall emphasise here his central theme about the suppression of information markets and what he thinks this means for the region’s prospects.

The case for ‘Asian values’ refers to three arguments: that these values caused, or substantially contributed to, economic growth in the region; that they protect willing populations against Western decadence; and that positive freedoms are more important than negative freedoms, or at any rate should come first. Let us take these points in turn. First, some of the values, such as hard work, thrift and concern for education, undoubtedly helped growth but are hardly specific to East Asia. It is improbable that they were a prime cause of growth; moreover, some are less cultural values than responses to particular institutional arrangements such as limited public welfare. Lingle draws attention instead to the role of the contingencies of the periods during which high growth began: for example, the non-repeatable removals of bureaucratic barriers, raids on non-renewable resources, and the importation of catch-up technology from the West.

Second, we do not know how willingly regional populations accept restrictions on access to the Western media because they cannot register their preferences. Censorship denotes an unwillingness to test the matter. This is a vexed subject. Many Westerners may actually approve of bans on pornography, feeling that if contemptible behaviour can be watched the unwary may take it as a fashionable thing to imitate. But there is a contrary risk in concealing such behaviour. The remainder of society is prevented from understanding and guarding against the trends shaping it, which also applies with force to political censorship. Anyone tempted by censorship should read Milton’s Areopagitica. There must be no prior restraint on publishing, says Milton: leave it to the courts. We should nevertheless recognise that a too-ready threat of prosecution actually adds up to prior restraint because it induces self-censorship.

Third, what of the ‘cruel choice’ between positive and negative freedoms? The ‘Asian values’ principle is that freedom from want — that is, positive economic freedom — is the more important. There is something in this; George Stigler admitted that F. A. Hayek underrated the freedom that prosperity has brought ordinary people, even in the presence of state coercion. The objection remains that prosperous authoritarianism is still authoritarian. There is no sunset clause, nor any means of enforcing one, that would guarantee a progression from positive towards negative freedoms: that is, towards freedom from arbitrariness. The individual has no ultimate security of property or contracts: a point made ardently in this book.

Optimists may think that positive freedoms should come first because they are beneficial in themselves; and once society is well-to-do there will indeed be a progression towards demands for a free press, political pluralism and independent law. By contrast, societies with political freedom and independent law — India, for example — have not always been super-successful in making the reverse transition from negative (civil) freedoms to positive (economic) freedoms. Which order is better? It may depend on our time preferences. After all, India has now begun to deregulate.
ther, without negative freedoms any society may slide back. In the absence of free debate — a lack which Lingle portrays as East Asia's greatest defect — rulers may always revert to making ill-judged decisions.

Lingle stresses the informational failings of authoritarian societies. Political markets suppress information flows and distort incentives. They skew the allocation of talent, drawing able people into bureaucratic occupations, rent-seeking and ersatz capitalism, away from risk-taking inventiveness and genuine entrepreneurship. Political markets reinforce the tendency of practices like guanxi (networking) to exclude competitors. Signals are muted, competition reduced, and misallocations of investment likely. How, then, to explain the growth rates observed in East Asia these past 30 years? His answer is that this was the accumulation phase when inputs were cheap and manufacturing for export was the obvious route to wealth. The institutions involved will, however, be too inflexible to elicit the creativity needed by economies increasingly based on services. The Tiger economies are reaching early maturity. It is Latin America and Eastern Europe which are now entering the high-growth phase that the Tigers enjoyed in the 1960s.

Lingle undermines the triumphalism which promises that the next century will be Asia's; but the deeper questions are how soon, and with what upsets, growth rates may level out. Will rent-seeking increase when the pie grows more slowly? Will the Mandate of Heaven be stripped from current regimes, only to see them replaced by others no more tender towards civil rights? Alternatively, will high incomes evolve a pluralism flexible enough to manage prosperity? The picture Lingle paints is of a system incapable of encouraging enough entrepreneurship or technological creativity, let alone allocating investment capital efficiently. He endorses the view that the Overseas Chinese middle classes have been co-opted by authoritarian regimes but believes outside pressures will force change. He thinks that global competition will oblige the region's institutions to adapt.

In statements to the press on a visit to Australia during mid-1997, Lingle implied that troubles in Japan and Thailand were the start of this vexed transition. That is possible. Alternatively, it may be a little like reading the floods and landslips of the same period as evidence of global warming. Economics has a poor record of forecasting the outcome of competition between systems; and in any event the Tiger economies are buffered by the considerable reserves they have built up during the past generation. Yet Lingle's case that their growth phase has not been the harbinger of an Asian century is informed, intelligent, internally consistent and clearly stated. Anyone thinking of investing in East Asia and everyone concerned about the region's political future would be well advised to take it into account.

*Eric Jones is a Professorial Associate at the Melbourne Business School.*
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)

Power to the Artist:
The False Promise of Moral Rights

Daniel Clode

In March 1997, Australia's Commonwealth government announced the introduction of 'moral rights' into Australian copyright law. According to the Minister for the Arts and the Attorney General, Australian artists could now feel assured that 'the Government is empowering them' (Alston & Williams, 1997).

‘Moral rights’ were first devised by French jurists as intellectual property rights for creators of artworks. These rights were conceived, not as economic in nature, but as protecting the spirituality and sentiments that artists encapsulate in their creations (Ricketson, 1987:456).

There are several categories of moral rights. The ‘right of paternity’ (also called the ‘right of attribution’) entitles artists to be known to the world as the creator of their works and to prevent others from falsely claiming authorship. The ‘right of integrity’ allows artists to sue anyone who modifies, distorts, or makes derogatory use of their creations. The ‘right of disclosure’ gives artists discretion over whether and when their work is to be divulged to the public. The ‘right of withdrawal’ allows artists to recall works from the public domain. The ‘right of repentance’ enables artists to modify, retract or disavow the messages originally expressed in their works. The ‘right to prevent excessive criticism’ conceivably gives mediocre artists plenty of opportunities to sue.

Daniel Clode is the political cartoonist for the Australian Medical Association and an Arts/Law undergraduate at The Australian National University.
Moral rights are intended to remain with artists even after they have sold their artworks. Artists are thus to retain a say over how those artworks are displayed, labelled or published by future buyers.

However, it is argued here that moral rights legislation is unnecessary. In Australia, numerous legal remedies under common law, the Trade Practices Act and the Copyright Act are already available to artists who believe their moral rights have been violated by misuses of their artworks. Incomplete though this protection may be, there is another remedy which can be tailored to meet an artist’s every concern and which prevents disputes in the first place. The greatest power artists possess lies in their freedom of contract, allowing them to forge terms of trade with their customers which resolve any anxieties about the paternity, integrity, publication or recall of their creative works. Moral rights legislation threatens to erode the level of coordination which such voluntary agreements can deliver.

Artists who are sensitive about alterations to their creations, or incomplete attribution, may insert appropriate conditions into their contracts of sale. Second and subsequent purchasers of an artwork may not be bound by such contracts, but the price paid by the first buyer can include a compensation premium. In this way, artists may insure themselves against affronts to their aesthetic tastes, while bearing the costs of these private preferences. By contrast, indiscriminate legislation forces such costs on to other artists who may not share those concerns.

**Approaches to Protecting Moral Rights**

There are three ways to protect moral rights. The opting-in approach allows artists to protect themselves primarily through contract, with the common law remedies of tort and equity as fallbacks. This system has operated in Australia for most of the 20th century. The opting-out approach uses legislation with waiver provisions; this is the option proposed in the Copyright Amendment Bill 1997. The third approach, involving perpetual, inalienable and unassignable rights, is found in the French Code. These different approaches are set out in Table 1.

More than 60 countries currently employ the latter two approaches for giving legislative force to different combinations of moral rights (CLRC, 1988:32). The Australian Copyright Amendment Bill 1997 proposes to protect only the rights of attribution and integrity. ‘Artists’ who can lay claim to moral rights include painters, writers, poets, musicians, composers, movie directors and even producers. I therefore use the terms ‘artist’ and ‘artwork’ in this broad sense.

Australian advocates of moral rights legislation usually argue that Australia must honour its international obligations under the Berne Convention for the Protection of Literary and Artistic Works. In fact, signatories to this Convention are entitled to rely upon domestic legal traditions, such as defamation law. Countries may therefore employ significantly different approaches, and are not confined to copyright law (Ricketson, 1987:475). Indeed, Australia adopted this option when it became a convention signatory in 1928, by choosing to rely on existing common law remedies.
### Table 1

**Approaches to protecting moral rights**

<table>
<thead>
<tr>
<th>Moral right</th>
<th>Concern to be remedied</th>
<th>Opting-in (existing remedies)</th>
<th>Opting-out (Copyright Amendment Bill 1997)</th>
<th>Compulsion (French law)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paternity/ attribution</td>
<td>False claim of authorship or failure to acknowledge actual creator.</td>
<td>Passing off; s.52 of Trade Practices Act.; defamation; ss.190-192 of Copyright Act; equitable relief.*</td>
<td>ss.192-194; 195AA-AG, AN-AO, AQ.</td>
<td>Originals, copies and quotes must bear the name or pseudonym of the artist.</td>
</tr>
<tr>
<td>Integrity</td>
<td>Modification, distortion, parody or derogatory use of a creation.</td>
<td>Passing off; defamation; ss.35(4)-(5) &amp; ss.55(2) Copyright Act; equitable relief.*</td>
<td>ss.195AH-AK, AP, AR-AS</td>
<td>Artist may sue for any change to artwork or setting, even if not the owner.</td>
</tr>
<tr>
<td>Against excessive criticism</td>
<td>Parody, mockery or severe criticism.</td>
<td>Defamation; ss.41 &amp; 42 of Copyright Act; equitable relief.*</td>
<td>Criticism of art is actionable, even if artist’s reputation is not harmed.</td>
<td></td>
</tr>
<tr>
<td>Disclosure</td>
<td>Discretion to determine when and whether work is divulged to public.</td>
<td>Detinue; conversion; breach of confidence; restraint of trade; defamation; equitable relief.*</td>
<td>Artists can refuse to release artwork, even when contrary to contractual obligations.</td>
<td></td>
</tr>
<tr>
<td>Withdrawal/ repentance</td>
<td>Discretion to withdraw work from public; modify, retract, or disavow messages in artwork.</td>
<td>Detinue; conversion; defamation; equitable relief.*</td>
<td>Retraction or modifications allowed upon change of convictions.</td>
<td></td>
</tr>
</tbody>
</table>

*E.g. unconscionability and estoppel.

### Inalienable Rights or Tradable Preferences?
The Copyright Amendment Bill 1997 was promoted as a ‘long overdue’ protection of the most ‘basic rights’ for artists (Alston & Williams, 1997). This rhetorical allusion to natural rights or divinely ordained entitlements was not reinforced by the legislative drafting. The Bill provided for the waiver of supposedly ‘inalienable’ rights to reputation as if these were economic choices that may be contracted away. This contradiction is the defining feature of the opting-out regimes that operate in
leading common law countries, including the United Kingdom, Canada and the United States.

During a Senate Committee inquiry into the Copyright Amendment Bill 1997, industry bodies argued that they could not afford to operate without waivers, since artists might veto modifications to their works (SLCLC, 1997). Such vetoes would render useless many of the art stocks that advertising agencies have acquired over the years. Modifications could not proceed without compensation, consent forms, or court action. Against this, peak organisations of artists argued that if waivers were allowed under statute, then employers would write them into all standard contracts. They also argued against the waiver provision on the ground that waiver processes were ‘economic’ in nature (though they simultaneously sought economic forms of redress to compensate for alleged breaches of moral rights). But if moral rights exist to protect personal reputation and aesthetic concerns as opposed to economic interests, it is odd that they should endure for 50 years after an artist’s death, concurrent with copyright. The estate of an artist cannot know with any certainty which uses of an artwork would have upset the artist, or when the artist would prefer to exercise a waiver of rights.

Without waivers, moral rights threaten art investors with major uncertainty, especially in the film and advertising industries where a chain of artists contribute to the final work. For example, many local films depend upon investment from the Australian Film Finance Corporation, which requires producers to obtain distribution contracts equivalent to 40 per cent of their budget. But broadcasters will usually not agree to buy distribution rights unless they can edit films to fit television time slots. If disputes over integrity rights deter broadcasters, many investors will not commit to new local production (Williams, 1997:219).

Yet even with waivers, moral rights impose major costs. Under such a system, arthouses, publishers, magazines and art syndicates need to implement internal monitoring systems and codes of conduct on the use of art, with all attendant expense. The paperwork burden can be massive, especially given the proposed retrospective application of moral rights to existing works (s.195AZL). Where an artist refuses to consent to changes to an artwork he has sold, the work must either be withdrawn or restored to its original condition.

The Australian bill allows for waiver of moral rights in the course of employment, prior to the creation of artworks (s.195AZG(3)). However, where artists work as contractors rather than employees, s.195AZG(2) limits waivers to specific works and may necessitate consent forms at every stage of production, including completion of advertisements and subsequent variations (AANA, 1997:465). These costs must ultimately impact on independent artists in the form of fewer work opportunities. The s.195AZG(3) mechanism for waivers over future works concedes that the pre-contractual negotiations are an appropriate point at which future uses of artwork can be anticipated. Why then restrict waivers to particular contractual situations, disadvantaging independent artists? Why not allow artists full freedom of contract, under which they can bargain terms covering the integrity and paternity of their art-
works? The opting-in system thus generates fewer costs and distortions in the art industry than the opting-out model.

Exploitation or Fair Bargains?

In common law jurisdictions, the most compelling argument mounted for moral rights legislation draws upon notions from English equity law (Martin & Bick, 1983:84-6). The argument is that struggling artists suffer from an ‘inequality of bargaining power’ which prevents them from negotiating fair terms under contract to govern the future uses and attribution of their work; only blanket legislation can ensure all artists against exploitation. This argument was the most prominent feature of community submissions to the Senate inquiry in support of moral rights.

Artists are said frequently to obtain bad deals because of a ‘take it or leave it’ approach by buyers. Yet, according to an AGB-McNair survey of 950 artists, of those artists engaged in contracts with galleries, agencies or managers, three times as many were satisfied with their contractual arrangements as were unsatisfied or uncertain (Throsby & Thompson, 1994:79). When asked how they might improve their promotion, the most often cited method was ‘more time/effort by artists themselves’ (31 per cent). Only 11 per cent felt that a new agent was the solution (Throsby & Thompson, 1994:38).

These statistics do not include sales outside an agent relationship. Perhaps only good artists get agents, fair deals and satisfactory incomes. However, 1991 census data on artistic professions provide little evidence of widespread exploitation (ABS, 1997:41). Among arts workers, 29.6 per cent earn $16,000 or less, compared with 29.3 per cent in the total Australian workforce. True, 24.4 per cent earn between $16,001 and $25,000, significantly fewer than the 30.4 per cent in the general economy. But arts workers enjoy a greater share of high incomes than other professions: 46 per cent earn over $25,000, compared with 40.2 per cent in the workforce as a whole. If gross incomes were adjusted to account for expenses incurred at work, the overall position of artists could be said to be only slightly worse than for the wider population. It should be noted that incomes do not reflect the job satisfaction and flexibility inherent in an artistic lifestyle.

The few cases of artists in poverty seem to have little to do with exploitation by art purchasers. A more salient cause may be found in the government subsidies which provide false flattery for sometimes undeserving artists and which entice some individuals into unsuitable careers on the very margins of economic survival. A significant 32 per cent of artists have applied for Australia Council grants in the last five years and 17 per cent have received grants at some stage in their career. Of those gaining grants, 21 per cent regarded their grants as a ‘stimulus to continue’ in the arts and 57 per cent felt the grants provided ‘freedom from financial worries’ (Throsby & Thompson, 1994:97-8). These Australia Council figures indicate only a portion of dependence on grants since they do not measure how many have also sought or obtained funding from other Commonwealth agencies, States or local

---

1The median level of arts-related expenses is $3,500 (Throsby & Thompson, 1994:26).
governments. The levels of dependency exposed in these 1993 survey findings are consistent with a 1988 survey which found that 39 per cent of artists believed that ‘buying time’ was the most important purpose of grants, while only 9 per cent viewed grants as a vehicle for ‘experimentation and innovation’ (Throsby & Thompson, 1994:49).

According to Browning (1995:162), over 60 per cent of the books subsidised by the Australia Council’s Literature Board sold no more than half their print run. It is little wonder that some incomes in the arts are particularly low, when governments are encouraging unmarketable talent into professional practice. These policies have contributed to a chronic labour oversupply, leaving only 24 per cent of all artists with regular income, 20 per cent on semi-regular income and 51 per cent on intermittent income (Throsby & Thompson, 1994:73). The oversupply of labour, rather than power imbalances and unfair deals, appears to be the biggest cause of any income depression in the arts.

**Legal Attrition or Market Sanctions?**

The market provides artists whose works have been abused with a less costly tool of retribution than court remedies, which can be both expensive and uncertain in outcome. Buyers who abuse artworks will acquire a negative reputation and lose the best artists. We might describe this as the ‘sanction of opprobrium’. The impact of reputational losses are asymmetric (Epstein, 1984:967), to the advantage of artists. For instance, the ratio of painters to galleries makes it easier for the former to monitor the latter than vice versa. If bargaining power is at all unequal in the arts, it may not be the buyer who has the upper hand.

A moral rights regime creates incentives for artists to seek legal ‘recovery’ to heal wounded egos. Lesser-known artisans have the greatest incentive to instigate moral rights actions, whereas able artists can find alternative employment to mitigate the loss of any soured trading relationship. The ironic danger is of an erosion of the creative spirit by the destructive emotions that come with litigation.

Under moral rights law, court action could be mounted over petty ‘rights’ breaches which might not otherwise provoke the opprobrium of the market. This is apparent from records of international negotiations over mutual copyright law recognition. Court action is intended to be available over ‘a literary work ... published in conjunction with numerous advertisements; an artistic work reproduced in conjunction with articles not enjoying a good reputation; a musical composition, profoundly serious and religious in tone, adopted as part of a filmed operetta’ (Ricketson, 1990:473). Submissions to the CLRC argued that lawsuits should be available to prevent inadequate lighting of art; placement of furniture in front of art; and failure to hang commissioned work (CLRC, 1988:14). In such cases, the incentive to reap a court award of money might be sufficiently strong to overwhelm the normal instinct to keep a trading relationship going.

The notion of ‘moral’ rights imputes base motives to those who may inadvertently offend the sensitivities of artists. Similarly, the pejorative concept of ‘unfair’ dismissals has been used unjustly to stigmatise many employers. The moral offen-
siveness of an investor's conduct is most accurately measured when aggrieved artists signal opprobrium by ending a commercial relationship.

Artists have the opportunity to protect their concerns by specifying in contracts how their artworks may be used. Buyers who break agreements on layout or acknowledgments will not only lose the best artists, but will also be liable for contractual damages. Artists in some fields have been very diligent in utilizing standard-form contracts over many years to protect themselves against problems of paternity and integrity (Martin & Bick, 1983:80-3). Greater use of standard contracts by art vendors would not only assist in insuring against unforeseen problems, but could also increase the availability of promissory estoppel remedies in situations involving established trading partners who operate without a formal contract.

**Protectionism or Cultural Renewal?**

In theory, moral rights legislation protects creativity as if it were a form of capital invested by artists. But in what way is the artist's creativity special? Why should not the landscape gardener claim the same legal privileges as the graphic designer? If a painter is to be compensated if one of his canvasses is touched up by a buyer, why should the house builder who is upset to see an old specimen of his craftsmanship renovated not likewise be compensated?

Even supporters of moral rights have become entangled in contradictory arguments over which professionals should claim authorship. The most heated dispute is over cinematographic works (SLCLC, 1997). Film and television often involves dozens of writers and directors, any one of whom could frighten away investors if they used their moral rights to veto script changes. In Australia's draft legislation, producers and directors were defined as authors, while scriptwriters were left out. The predictable debate that ensued over which vocation is the true source of creative genius served to expose the danger of moral rights in an industry which is essentially a collaborative venture.

The pioneers of moral rights jurisprudence were driven by juristic notions that law should strengthen over time (Stromholm, 1983:6). Indeed, the majority of submissions to the Senate inquiry into moral rights called for more extensive protection than was originally proposed (SLCLC, 1997). Despite deliberate loopholes in the draft legislation allowing 'reasonable' infringements of integrity rights (s.195AR), parodies of artworks can still fall foul of a moral rights regime. A 'reasonableness' criterion may actually encourage litigation, precisely because of the indeterminate nature of the concept (CLRC, 1988:18).

Artists are wonderfully equipped to lampoon new fashions in the arts. In the past, artistic ridicule could be requited in kind. But under moral rights legislation, the lawyer can annex the satirist's drawing board. Moral rights supporters argue that parodies can be tested for harm to the author's honour or reputation (Wier, 1992:196). But real harm to reputation is already actionable under current defamation law. If an especially vicious attack destroys an artist's following, then remedies exist allowing the recovery of economic damages. By contrast, if the artist's clientele remains loyal after an attack and that attack did not constitute defamation, then no
substantial damage has been inflicted that might justify legal action. Under copyright law, the partial replication of an original artwork is permissible in a parody, provided the secondary work does not confuse the buyers by stealing the market for the original. To quote the courts: 'Biting criticism suppresses demand; copyright infringement usurps it'.

The submission by a major law firm to the recent Senate inquiry illustrated the dangers of moral rights with examples from two French court cases:

A vibrant culture will look critically and carefully at its values and priorities. One of the more effective methods is to reassess and reinterpret its cultural heritage.

Moral rights may inadvertently stymie this cultural development ... Gustav Holst was able to restrain a modern reinterpretation of the electronic version of his composition 'The Planets'. Analogously, Beethoven's estate could have stopped Liszt's arrangements of Beethoven's work. Similarly, the Beckett estate's action to restrain Beckett's 'everyman' play 'Waiting for Godot' with a female cast robs the modern audience of an interesting reinterpretation, and firmly roots the work in a different era where the only relevant audience was 'everyman'. (Williams, 1997:227)

The Costs of Corporatism

Australian studies have chronicled few cases where moral rights were seriously abused. By 1983, there had been only one published list of known breaches, totaling 14 cases. Some of these incidents revealed 'over-sensitivity' and 'unreasonableness' by the artists involved (CLRC, 1988:26). In 1992, the Australian Copyright Council received 10,472 copyright enquiries, of which only 77 related to moral rights. Visual artists and writers made more queries than did authors of musical works or films. The main interest of visual artists was potential misuse of original works, not reproductions (Department of Attorney General, 1994:22). However, I am aware of no empirical study that catalogues the number of serious incidents each year or details whether power imbalances existed, which other legal remedies were available, and what evidential difficulties existed.

Supporters of moral rights legislation recognise that if the French laws were ever fully enforced, the community would find them 'burdensome and unworkable' (Department of the Attorney-General, 1994:31). Such laws which are tolerable only in a state of non-enforcement are probably not worth keeping. Laws which threaten to do the greatest harm to those whom they purport to assist are not even worth acquiring. The strengths of a free market should not be judged by the drama of the few abuses, but by the multitude of successes where trading partners are cemented together by goodwill. Legislative attempts to anticipate the few possible failures involve attaching to every contract obligations designed to cover all imaginable contin-

---

2 Fisher v Dees 794 F2d 432 at 438.
gencies. Thus, contracts which were intended by their authors to be simple and flexible become complex and uncertain in their effect.

The costs of moral rights go well beyond court expenses. Industry compliance will necessitate consent forms and compensation to cover thousands of existing artworks. Given that the total annual income of all artists (including proceeds from the sale of artworks) was $1.2 billion at the 1991 census (ABS, 1997:44), the possible cost of introducing a moral rights regime could total at least millions of dollars. Even more ominously, any attempt by future Commonwealth governments to transfer moral rights within collaborative ventures (such as the film industry or the information technology sector) could require payment of 'just terms' compensation to previous moral rights holders under s.51(xxxi) of the Constitution. Unless moral rights are legislatively defined as tort actions rather than as property rights, it becomes prohibitively expensive to alter their scope, even though in future there may be pressing economic, technological or cultural reasons to do so.

Ever since the Gorton Government of 1968-71, federal politicians have viewed the Australian arts community as an articulate adoptee, politically appreciative of favours. Recently, as swelling public debt has limited the scope for financial largesse, regulatory favouritism has become the new currency of corporatism. Yet strength never comes to those who sit at the table of 'infant industries' singing for a supper of subsidies and regulatory benefits. Legislation which advantages my craft only strengthens the claims of those in other realms of productive endeavour who demand an offset for my privileges. True power comes only with independence. A wise artist would say to the Arts Minister, 'Please sir, may I go without?'

References

Alston, R. & D. Williams (1997), Media Release, 4 March.


3 Smith Kline & French v Secretary Department of Health (1990) 95 ALR 87.

Senate Legal and Constitutional Legislation Committee (SLCLC) (1997), Submissions: Copyright Amendment Bill 1997, Vols 1-3, Canberra.


Throsby, D. & B. Thompson (1994), But what do you do for a living?: a new economic study of Australian artists, Australia Council, Redfern, NSW.
