

## **Distributional Implications of the Welfare State**

**James Cox**

**T**his paper is concerned with the effect of the welfare state in redistributing income away from or towards different types of household. As is well-known, OECD countries differ greatly in how much income is redistributed through the welfare state and in priorities for government spending. This paper investigates whether there is a systematic relationship between the amount of government expenditure on the welfare state and the distributional implications of the welfare state for OECD countries. (A longer and more detailed version of this paper is available from the author on request).

Most previous analysis has been concerned with the effect of the welfare state in redistributing income from high income households to low income households. By contrast, the investigation in this paper examines the effect of the welfare state in redistributing incomes between households at particular stages of life. The interest is not with how the welfare state affects households with unusually high or low incomes but with how it affects families with children or the retired in general.

There are a number of reasons for taking this approach. First, the data are of considerable interest in their own right. Not only do the data enable a more complete understanding of the distributional implications of the welfare state but they enable us to understand the extent to which the welfare states of different countries subsidise some types of households (such as single person households and sole parents) and tax other types of household (such as two parent families with children). These taxes and benefits through their effect on incentives, have implications for family formation and stability, welfare dependency and fertility. Secondly, despite the fact that it is less frequently undertaken, investigating the effect of the welfare state on population groups as a whole is in some respects easier than studying the effect of the welfare state in redistributing income from households with high current income to households with low current income. In particular, the incomes of households with unusually high or low incomes tend to be inaccurately reported in household surveys (for further discussion on this point see, for example, Behrendt, 2000; and Saunders and Tsumori, 2002). Households in general often fail to report in surveys income from social security benefits or from transactions that are not disclosed to the tax and social security authorities. These sources of income are likely to be particularly important for many low income families.

This paper therefore investigates, using data from household surveys, the effect of the welfare state in redistributing income between different types of households.

---

*James Cox is a Member of the New South Wales Pricing Tribunal.*

## Data

Household surveys have increasingly been undertaken in OECD countries in recent years. Although these surveys have weaknesses, they provide the best available basis for studying the distributional impact of the welfare state in different countries. One potential source of difficulty is differences between countries in the definition of income and demographic characteristics of the population. These differences make accurate comparisons very difficult. Fortunately, however, a great deal of effort has been made in recent years to standardise income surveys. In particular, the Luxembourg Income Study has endeavoured to make standardised data widely available to researchers, and it is the basis of the present study. The widespread use of these data make it likely that remaining instances of non-comparability of data will be identified by researchers and acted upon by the originators of the data. The final report of the Canberra Group (2001) provides a useful discussion of the problems of income surveys and possible remedies.<sup>1</sup> An account of the development and current status of the Luxembourg Income Study can be found in Smeeding (2002). Despite the efforts made to standardise the data, some problems of compatibility remain. These issues are discussed in the longer version of the paper and are mentioned below to the extent that they are relevant to the findings presented in this summary paper.

The variables of most interest to this study are measures of income and household composition. Data on income derived from market sources are available from the Luxembourg Income Study. The information on government benefits that is available from the Luxembourg Income Study is in practice limited to cash benefits. Information on government non-cash benefits (for example, in health and education) is unfortunately not available from this source. Moreover, the available information on taxes is limited to direct taxes that are paid by individuals. The income variables that are available to investigate the distributional effects of the welfare state are therefore market income (that is, non-government income) and disposable income (market income plus government benefits minus direct taxes). The differences between the amount of disposable income and market income received by a household provides a measure of the redistributive effect of the welfare state. Although this measure is less comprehensive than a measure that also includes non-cash benefits and indirect taxes, it takes account of those aspects of the welfare state that are most likely to vary between countries. Although there are important differences between countries in government spending on welfare services as a share of GDP, spending on cash benefits is the most important source of variation between countries in government spending on the welfare state.

The demographic variables considered in this study are the age of household members and household composition. Under household composition, households are first classified into single unit households and multi-unit households (the term unit is used here to refer to the nuclear family of adults and children, and

---

<sup>1</sup> I am grateful to a referee for providing this reference.

abbreviations of it such as single person households, single parent households and households comprising a couple without children). These groups of households are then further subdivided depending on the number of adults in the household and whether or not dependent children are present. Results are first presented in this paper for a simple classification of households that does not take into account the age of the head of the household, then by the age of the household head and then finally for a complex household classification that takes account both of age and household structure.

In this paper, welfare states are ranked from small to large according to the proportion that benefits provide of total household income. This provides a method of ranking welfare states that is consistent with the data available from the Luxembourg Income Study but does not take account of spending on health and education. As will be argued below, both the total amount of redistribution and priorities for spending differ between countries. These differences are important in understanding the redistributive effects of different types of welfare state.

## **Methodology**

The main purpose of this paper is to assess the net distributive impact of the welfare state taking into account both taxes and benefits. To do this it is necessary to develop a measure of the *net* redistributive impact of the welfare state that is both easy to interpret and fairly robust, given the great diversity that exists between countries in welfare states.

The measure used in this paper is the percentage share of market or disposable income that is received by a particular group. These shares add up to 100 per cent for the population as a whole. The net redistributive impact of the welfare state can be assessed by comparing the share of disposable income received by a group with the share of market income that is received by the same group. If the share of disposable income is greater than the share of market income that is received by the group, then the welfare state redistributes towards the group. But the welfare state redistributes away from the group if the difference between these two shares is negative. These differences sum to zero for the whole population. It should be noted that there is no particular reason to expect that the amount of government benefits received by households in a country will equal the amount of direct taxes paid by households. The total amount of disposable income received by households will typically differ from the total amount of market income. Moreover, the extent of this discrepancy will vary between countries. The percentage shares approach seems to be the best way of dealing with these discrepancies.

These percentage shares measures are fairly easy to calculate and to understand. Since many errors will affect both numerator and denominator proportionately, the percentage shares may not be too greatly affected by differences between countries in the methods used to collect and report data. Moreover, the percentage shares can readily be decomposed to show results for sub-groups within the population.

As is emphasised in the rest of this paper, the differences between the share of disposable income that is received by a group and the share of market income received by the same group depends on the overall amount of redistribution that occurs in a country as well as on the structure of the taxation and social security system.

By comparing disposable to market income, this study emphasises the initial effects of benefits and taxes on the distribution of income in different countries. This procedure does not take into account behavioural responses to taxes and benefits that are likely to reduce, but not eliminate, the initial effects. For example, people are likely to respond to governments making higher benefits available by working less. Although the increased leisure is valuable in its own right, this response will reduce the gain in income as a result of the higher benefit. Since these behavioural responses depend on the initial effects, the initial effects are worth studying in their own right. (Of course, the market incomes reported in this study are influenced by the disincentive effects of taxation and the welfare state. However, the main measure of the redistribution impact used here is the differences between disposable and market income.)

## Results

Results are presented in this paper in a number of ways:

- First, for groups defined according to the simple household composition variable that does not take account of age of head of household;
- Next, for groups defined according to the age of the head of the household; and
- Finally, for groups defined according to the complex household classification variable that incorporates information on the age of the household head.<sup>2</sup>

### *Simple household classification*

Results for groups defined according to the simple household classification variable are presented in Table 1. Countries are ordered in the table according to the importance of benefits as a source of income. The countries in which benefits are least important are placed at the top of the table while those in which benefits are most important are placed at the bottom of the table. The label (such as USA 97) refers to the relevant country and the year of the survey from which results have been obtained.

---

<sup>2</sup> A referee has commented that one should think of the groups identified here not so much as different groups of people but as the *same* people at different stages of life. For example, today's parents can expect to enjoy the favourable treatment enjoyed by today's retired when they grow older. It is uncertain whether this assumption is warranted. For example, the literature on generation accounting (see Ablett and Tseggai-Bocurezion 2000 for an Australian study) emphasises that the treatment of different generations by taxation and social welfare arrangements can vary substantially.

**Table 1: Differences Between Percentage Share of Disposable and Market Income for Demographic Groups in OECD Countries**

(percentage points)

	<i>Sole Occupant</i>	<i>Single Person with Children</i>	<i>Couple only</i>	<i>Couple with Children</i>	<i>Multi-unit Households</i>
USA 97	1.35	0.81	0.35	-2.41	-0.09
Australia 94	1.42	0.99	0.57	-2.61	-0.37
Switzerland 92	2.95	0.37	0.68	-3.02	-0.97
Canada 97	1.97	0.73	0.87	-2.93	-0.65
UK 95	2.53	2.37	-1.19	-3.25	-0.46
Ireland 87	1.83	0.59	0.85	-2.20	-1.07
Finland 95	1.78	1.00	-1.39	-1.18	-0.20
Germany 94	4.55	0.59	2.53	-5.21	-2.46
Norway 95	2.99	1.52	1.40	-3.61	-2.31
Netherlands 94	3.77	1.12	-1.04	-3.69	-0.15
Denmark 97	4.92	0.64	-0.17	-3.37	-2.02
Spain 90 <sup>a</sup>	2.26	0.13	4.39	-6.67	-0.13
Belgium 97	4.30	0.42	4.48	-6.23	-2.96
Italy 95 <sup>a</sup>	3.54	-0.05	4.21	-7.07	-0.62
Sweden 95	4.90	1.93	-2.18	-4.65	-0.01
France 94	3.31	0.10	6.10	-7.79	-1.72
Average <sup>b</sup>	3.02	0.83	1.28	-4.12	-1.01

Notes: a Data on taxation are not available for these countries. The differences shown in the tables are therefore the differences between the percentage share of gross income (including benefits) and market income for the relevant groups.

b Unweighted arithmetic average for all countries shown in the table.

Source: Luxembourg Income Study

It will be seen from the table that the welfare state redistributes towards households that are sole occupants, sole parents and (in most countries) couples. The welfare state redistributes away from couples with children and multi-unit households to benefit the groups noted above. (Children are defined in the Luxembourg Income Study to be persons under 18 years.) In the statistically average OECD country, about three percentage points of income is redistributed towards sole occupants and about one percentage point to sole parents and couples. About four percentage points is redistributed away from families with children and about one percentage point of income is redistributed away from multi-unit households. Of course, no actual country is exactly the same as this statistical average.

The amount of redistribution away from couples with children and multi-unit households, and towards sole occupants, sole parents and couples, tends to be greater in large than in small welfare states. There are, however, important differences between countries. For example, the Scandinavian countries redistribute particularly large amounts to households that are sole occupants or sole parents. The Scandinavian countries, however, redistribute a lesser amount of income away from families with children than the large welfare states of western and southern European countries that particularly emphasise redistribution to couple households without children.

#### *Age of household head*

Table 2 shows that, in all countries considered here, the welfare state redistributes income from households aged 25 to 54 years to households aged over 65 years. The extent of redistribution away from younger households is greatest in those countries with the greatest amount of redistribution to the elderly. There is more redistribution from younger households towards the elderly in large than in small welfare states. The priorities for social expenditure that are adopted in different countries are also important.

The largest amount of redistribution from younger households to the elderly occurs in the western and southern European countries that place an unusually heavy emphasis on pensions for the elderly.

Households with heads aged 55 to 64 years are net contributors to the welfare state in some countries (especially Norway and Sweden) and net beneficiaries in others (especially Belgium, France and Italy). As noted above, households where the head is aged over 65 years are net beneficiaries of the welfare state in all countries and especially so in western and southern European countries.

**Table 2: Differences between Percentage Share of Market and Disposable Income for Age Groups in OECD Countries**  
(percentage points)

	<i>Age of Household Head</i>					
	<i>Under 25 years</i>	<i>25 – 34 years</i>	<i>35 – 44 years</i>	<i>45 – 54 years</i>	<i>55 – 64 years</i>	<i>Over 65 years</i>
USA 97	0.13	-0.84	-1.96	-2.42	-0.72	5.82
Australia 94	0.26	-1.36	-2.21	-2.40	0.48	5.22
Switzerland 92	-0.23	-1.34	-2.40	-2.52	-0.55	7.03
Canada 97	0.13	-1.12	-2.64	-2.70	-0.25	6.57
UK 95	0.59	-1.68	-2.50	-3.42	0.12	6.88
Ireland 87	-0.06	-2.08	-1.46	-1.29	-0.83	5.72
Finland 95	1.22	0.70	-0.98	-3.09	-0.61	2.75
Germany 94	-0.04	-3.53	-4.06	-6.00	-0.78	14.43
Norway 95	0.32	-1.32	-2.95	-4.16	-1.00	9.10
Netherlands 94	0.33	-3.18	-2.25	-2.51	1.07	6.84
Denmark 97	0.20	-1.33	-2.44	-4.19	-0.36	8.13
Spain 90 <sup>a</sup>	-0.12	-2.57	-4.54	-3.92	0.47	10.69
Belgium 97	0.01	-4.02	-5.42	-6.81	2.38	13.85
Italy 95 <sup>a</sup>	-0.08	-2.61	-5.25	-5.26	2.13	11.08
Sweden 95	0.52	-0.92	-2.87	-6.53	-1.96	11.75
France 94	0.07	-3.94	-5.99	-7.50	2.24	15.11
Average <sup>b</sup>	0.20	-1.95	-3.14	-4.05	0.11	8.81

Notes: a Data on taxation are not available for these countries. The differences shown in the tables are therefore the differences between the percentage share of gross income (including benefits) and market income for the relevant groups.

b Unweighted arithmetic average for all countries shown in the table.

Source: Luxembourg Income Study

**Table 3 Difference Between Share of Market and Disposable Income for Demographic Groups in OECD Countries (percentage points)**

	Sole Occupants			Couples without Children Head:			Couples with Children Head:				Sole Parent Family Parent:		Other Family Households			
	under 35 years	35 – 64 years	65 and over	under 35 years	35 – 64 years	65 and over	under 25 years	25 – 34 years	35 – 44 years	45 and over	under 35 years	35 and over	Single No Children	Single with Children	Married no Children	Married with Children
USA 97	-0.22	-0.53	2.1	-0.52	-1.83	2.7	0.04	-0.32	-1.34	-0.80	0.39	0.43	0.24	0.49	-0.40	-0.43
Australia 94	-0.28	-0.27	1.97	-0.83	-0.88	2.28	0.09	-0.41	-1.58	-0.72	0.55	0.44	0.15	0.42	-0.59	-0.34
Switzerland 92	-0.47	-0.16	3.70	-0.68	-1.98	3.24	-0.01	-0.47	-2.16	-0.83	0.13	0.26	0.09	-0.02	-0.45	-0.20
Canada 97	-0.26	-0.34	2.57	-0.62	-1.47	2.95	0.03	-0.40	-1.81	-0.74	0.45	0.28	0.35	0.27	-0.62	-0.66
UK 95	-0.57	0.21	2.89	-1.69	-2.51	3.00	0.07	-0.64	-1.97	-0.71	1.53	0.84	0.74	0.34	-1.01	-0.53
Ireland 87	-0.21	-0.03	2.06	-0.85	-0.16	1.85	0.27	-0.98	-1.39	-0.10	0.33	0.26	0.31	0.53	-1.34	-0.58
Finland 95	0.46	-0.22	1.59	0.15	-2.32	0.71	0.09	0.53	-0.87	-1.00	0.61	0.43	0.67	0.16	-0.57	-0.41
Germany 94	-1.05	-0.97	6.57	-2.01	-2.27	6.81	0.02	-0.88	-2.41	-1.95	0.34	0.25	0.31	0.03	-1.39	-1.41
Norway 95	-0.80	-0.50	4.3	-0.75	-2.00	4.16	0.04	-0.63	-2.16	-0.85	1.13	0.39	0.25	0.16	-1.67	-1.04
Netherlands 94	-0.26	0.47	3.55	-2.24	-1.68	2.88	0.04	-0.81	-2.14	-0.78	0.35	0.78	0.57	0.13	-0.66	-0.21
Denmark 97	-0.34	0.30	5.0	-0.89	-2.62	3.23	0.06	-0.57	-1.91	-1.13	0.40	0.25	0	0.05	-1.16	-0.66
Spain 90 <sup>(a)</sup>	-0.11	0.33	2.04	-0.53	0.10	4.83	-0.05	-1.90	-3.96	-0.77	0.03	0.10	1.48	0.35	0.54	-2.50
Belgium 97	-0.60	-0.53	5.51	-2.05	-1.03	7.55	0.02	-1.88	-3.89	-1.04	0.17	0.24	0.67	0.23	-2.07	-1.27
Italy 95 <sup>(a)</sup>	-0.23	0.09	3.67	-0.85	0.33	4.72	0	-1.64	-4.10	-1.33	0	-0.05	1.28	0.05	0.37	-2.33
Sweden 95	-0.47	-1.1	6.46	-0.96	-6.45	5.24	0.07	0.03	-2.38	-2.37	0.94	0.99	-0.01	0	0	0
France 94	-0.96	-0.99	5.28	-1.29	-0.93	8.36	0.01	-1.92	-4.01	-1.97	0.14	-0.05	0.82	0.16	-1.35	-1.3
Average <sup>(b)</sup>	-0.40	-0.27	3.70	-1.04	-1.73	4.03	0.05	-0.81	-2.38	-1.07	0.47	0.37	0.50	0.21	-0.77	-0.87

Notes: a Data on taxation are not available for these countries. The differences shown in the tables are therefore the differences between the percentage share of gross income (including benefits) and market income for the relevant groups.

b Unweighted arithmetic average for all countries shown in the table.

Source: Luxembourg Income Study



*Complex household composition*

Table 3 provides information on the net redistributive impact of the welfare state in OECD countries for groups defined according to both the type of household and (for single family households) the age of the head of the household.

It can be seen from the table that, in most countries, the welfare state redistributes away from:

- Sole occupants under the age of 65 years;
- Couples without children where the head is aged under 65 years;
- Couples with children where the head is aged 25 years and over;
- Couples with children and other adults;
- Couples with other adults (but not dependent children).

The welfare state redistributes towards:

- Sole occupants aged 65 years and over;
- Couples without children where the head is aged 65 years and over;
- Sole parent families;
- Sole parents with children and others present in the household; and
- Households made up of single people without children.

The amount of redistribution that occurs is significant — around 9.3 percentage points of income is shifted between the groups noted above in the typical OECD Country.

The table confirms that the amount of redistribution that occurs between demographic groups is greater in large than in small welfare states. There are also some important differences between countries that are worth noting.

These include the following:

- Some large welfare states (such as Germany, Belgium, France and Italy) particularly emphasise redistribution to the elderly. The welfare states of these countries have a particularly large adverse effect on the income of couples with children.
- Other large welfare states (such as the Scandinavian countries) emphasise both benefits paid to families with children and benefits to older households. Despite these differences in emphasis on the benefits side, the high level of taxation in these large welfare states results in a net adverse redistributive impact for couples with children. However, this adverse effect is less than in the first group of countries.
- Although the taxation system in France provides advantages to families (as compared to other taxpayers) these seem in practice to be outweighed by the high level of taxation in that country. The net effect of the welfare state on the incomes of couples with children is strongly negative in France.

- The adverse effect of the welfare state on the incomes of couples with children is significantly less than elsewhere in the small welfare states such as the United States and Australia. These small welfare states redistribute a relatively small amount to the elderly.
- The welfare states of the United Kingdom, the Netherlands and the Scandinavian countries particularly emphasise redistribution to sole parent families.

The strong relationship between the amount of redistribution to the elderly that occurs in a country and the extent of redistribution away from families with children is perhaps the most important result in this paper.

### **Some Underlying Factors**

This completes presentation of the main results in this paper regarding the distributional impact of different types of welfare state. The net redistributive impact of the welfare state is the sum of the separate effects of the benefits and taxation systems. These in turn depends on the demographic structure of the population and the total amount of redistribution that occurs in a country as well as on the structure of the benefits and taxation systems.

Some information about the underlying factors is presented here to enhance understanding of the distributional consequences of the welfare state. In the interests of saving space, only summary results are presented here. Further details, including the presentation of evidence, can be found in the longer version of the paper.

#### *Cash benefits*

The percentage share of cash benefits that is paid to single person households, sole parent families and couples where the head is aged over 65 years exceeds the percentage share these groups represent of the population of households. In the average OECD country, these groups comprise 49 per cent of households but receive 62 per cent of benefits. By contrast, couples under 65 years without children, couples with children and multi-unit households that include children comprise 51 per cent of households but receive 38 per cent of benefits.

There are important differences between countries in which groups receive benefits. These reflect differences in demographic structure (which are discussed below) and priorities for social expenditure. As noted earlier, the western and southern European countries pay large amounts of benefits to the elderly while the Scandinavian countries pay large amounts to younger people and families as well as the elderly. The United Kingdom and the Netherlands emphasise the payment of benefits to sole parents.

The United States directs to the elderly a high proportion of the limited amount of cash benefits that it provides. By contrast Australia, another small welfare state, pays a high proportion of benefits to younger people.

### *Taxation*

Single people, sole parents and couples where the household head is aged 65 years and over pay 29 per cent of taxation in the average OECD country but make up, as noted, 49 per cent of households.

Couples aged under 65 years, couples with children and multiple unit households that include children make up 51 per cent of households but pay 71 per cent of taxes. Depending on the precise definition that is used, around 40 to 50 per cent of taxation in the average OECD country is paid by families with children. In general, the taxation burden is largely borne by childless couples under the age of 65 years and couples with children.

There are also important differences between countries. The Scandinavian countries emphasise the taxation of benefits: sole occupants and sole parents pay an unusually high proportion of the tax bill. These groups are less heavily taxed in Australia and Ireland and also (especially for older households) in Germany (for further discussion about the taxation of benefits, see Adema, 2001). The French tax system provides some advantages to families with dependent children but nevertheless the share of taxation paid by this group is fairly high. In general there seems to be less variation between countries in the percentage shares of taxation that is paid by demographic groups than in the share of benefits that is received by them.

### *Demographic structure*

One reason why the shares of benefits and taxes received by population groups differ between countries is differences in the demographic structure of their populations.

The Scandinavian countries have a high proportion of single person and sole parent households. For example, 46 per cent of households in Denmark are sole occupants, as opposed to 30 per cent in the average OECD country. The English-speaking countries tend to have a relatively high proportion of households that are sole parents or groups of single people who live together. Ireland has a high proportion of households that comprise couples with children aged under 18 years. Spain and Italy have a high proportion of households that comprise couples (with or without children aged under 18 years) and others, including adult children. This group make up 31 per cent of households in Italy but 13 per cent in the average OECD country. Other countries fall within the extremes noted here.

### *Sources of income*

The proportion of government benefits in total household income (both market and benefit income) is 20 per cent in the average OECD country. It varies from 9 per cent in the United States to 32 per cent in France. As might be expected, the share of benefits in total household income is particularly high for older people and sole parents, and particularly low for couples below retirement age and

without children, couples with children and multi-unit households that include couples.

There is substantial variation between countries in the source of income for each demographic group. For example, 56 percent of the income of couples with heads aged over 65 years is provided by benefits in the average OECD country. However, this percentage varies from 37 per cent in the United States to 85 per cent in France.

Similarly, 52 per cent of the income of sole parent families where the parent is aged under 35 years is provided by benefits in the average OECD country. This percentage varies from 31 per cent in the United States to 83 per cent in the United Kingdom. This wide variation between countries shows that there are, in principle, many possibilities available for providing groups such as sole parents or the elderly with an income. Government benefits are not the only, or even necessarily the most important, source of income for these groups.

There is evidence from these data that people adjust to higher and more generally available benefits by working or saving less than they otherwise would have done. For example, a lesser amount of market income is received by sole parents and older people in countries that provide high benefits to these groups.

### **Some Qualifications**

The results presented above take into account government benefits that are financed through taxation. They exclude, for example, compulsory private expenditures that are required, for example, by Australia's superannuation program. More importantly, perhaps, they exclude the many ways in which people provide for retirement in countries where government programs are not extensive. It does not appear to be true that the aged are poorer in countries where governments provide lower benefits or restrict eligibility for benefits (for evidence on this point, see Börsch-Supan and Reil-Held, 1998; Disney, D'Ercole and Scherer, 1998; Rein and Stapf-Fine, 2001; Disney and Johnson, 2001). Rather, people provide for old age by working longer, saving more or through high rates of home ownership. However, the need for people to undertake such activities reduces their standard of living (or quality of life) during their working lives. The differences between the situation of families in large and small welfare states may therefore be less than the data in tables 1 to 3 would suggest.

There are, however, important differences between government and voluntary arrangements that suggest that the data in the tables tell an important part of the story. (Compulsory private expenditure, such as superannuation in Australia, should be considered to be an example of a government program rather than a voluntary arrangement.) For example, government programs in the larger welfare states over-provide for retirement. These over-generous transfers are reversed to some extent through bequests and gifts from older to younger people (for further discussion on this point, see Börsch-Supan and Reil-Held, 1998:5-6). Since not everyone receives a gift of bequest, a lower level of transfers might be more equitable. In addition, a voluntary transaction will only be entered into if it is to

the advantage of both parties. Families may decide not to increase savings for retirement if immediate expenditure needs have high priority. It is often argued that, left to their own devices, people under-provide for retirement. By contrast, governments can require people to enter into arrangements which they would not have chosen voluntarily. Finally, there is usually a close link in private retirement income programs between the amount of contributions and the amount that is available on retirement. This is usually not the case for government programs which typically involve an element of redistribution. The psychological significance of contributions differs, therefore, between government and private programs.

## **Conclusions**

This paper has examined the distributional effects of the welfare states that exist in OECD countries. In general, the welfare state redistributes away from childless couples under retirement age and couples with children towards single people, sole parents and the elderly. The extent of redistribution away from couples with children is greater in large than in small welfare states. It is greatest in those welfare states that emphasise benefits for the elderly.

Although the welfare state expands opportunities for some groups (including sole parents and the aged), it limits through taxation opportunities for other groups (such as couples with children). The direction of redistribution is not necessarily consistent with differences in need. The aged have low current incomes but would have experienced over their lifetime incomes that are consistent with economic experience of their generation. Couples with children tend to have high current incomes but also high needs<sup>3</sup>. This dual nature of the welfare state in expanding opportunities for some groups but limiting them for others needs to be considered in deciding what type of welfare state a country should have.

## **References**

- Ablett, J. and Z. Tseggai-Bocurezion (2000), 'Lifetime Net Average Tax Rates in Australia — A Generational Accounting Study', *Economic Record* 76(233):139-151.
- Adema, W. (2002), *Net Social Expenditure* (second edition), Labor Market and Social Policy Occasional Paper No 52, OECD, Paris.
- Behrendt, C. (2000), 'Is There Income Poverty in Western Europe? Methodological Pitfalls in the Measurement of Poverty in a Comparative Perspective', Luxembourg Income Study Working Paper No 258.
- Börsch-Supan, A. and A. Reil-Held (1998), 'Retirement Income: Level, Risk and Substitution Among Income Components', Ageing Working Paper 3.7, OECD Paris.

---

<sup>3</sup> Moreover, families that lack substantial assets may find it difficult to borrow against their expected future incomes. The living standards of these families will depend on current income to a greater extent than those of the aged. I am grateful to a referee for pointing this out.

Canberra Group (2001), Expert Group on Household Income Statistics, *Final Report and Recommendations*, Ottawa, <http://www.lisproject.org/links/canberra/finalreport.pdf>.

Disney, R., M. d'Ercole and P. Scherer (1998), 'Resources During Retirement', Ageing Working Paper 4.3, OECD, Paris.

Disney, R. and P. Johnson (2001), *Pension Systems and Retirement Incomes across OECD Countries*, Edward Elgar, Cheltenham.

Rein, M. and H. Stapf-Fine (2001), 'Income Packaging and Economic Well-Being at the Last Stage of the Working Career', Luxembourg Income Study, Working Paper No. 270.

Saunders, P. and K. Tsumori (2002), *Poverty in Australia: Beyond the Rhetoric*, CIS Policy Monograph No. 57, St Leonards NSW.

Smeeding, T. (2002), 'The LIS/LES Project: Overview and Recent Developments', Luxembourg Income Study, Working Paper No. 294.

*I am grateful to the Social Policy Research Centre at the University of New South Wales for making the data available to me and to Sharon Burke of the Social Policy Research Centre for undertaking the calculations on which the tables presented in this paper are based. I am also grateful for comments by two anonymous referees.*

## **Application of Equity to Child Support**

**Anthony Farr and Gary Buurman**

**T**his paper proposes that the economic criteria of horizontal and vertical equity can be applied to Child Support and may be used as standards for evaluation of the Child Support formula of the Australian Child Support Scheme. Financial modelling is used to simulate the financial circumstances of separated parents. While some previous research published in Australia has analysed the financial circumstances of parents, financial modelling has been limited to, at most, a handful of particular circumstances (for example, Scott, 1989). The only previous attempt to model a broad range of situations is an unpublished work developed by the Department of Social Security. The Joint Select Committee on Certain Family Law Issues considered this model and called for the development of an alternative model (JSC, 1994). This challenge has not been taken up to date.

This paper is limited to an evaluation of the Child Support Scheme against equity criteria and does not propose changes to the Scheme. While the Scheme clearly has other (possibly competing) policy objectives, they are beyond the scope of this paper. Sections of the paper discuss Child Support, equity principles in relation to Child Support, the model underpinning our analysis, results in respect to horizontal and vertical equity and sensitivity testing.

### **Child Support Scheme**

The Child Support Scheme in its current Australian form came into operation on 1 October 1989. The Scheme provides for the transfer of income from a parent (normally the Non-Resident Parent or NRP) to a person caring for the children (normally the Resident Parent or RP). The amount of this transfer is determined by administrative assessment using calculations legislated in the *Child Support (Assessment) Act 1989* (CSA). These calculations are generally referred as the 'Child Support formula' — though assessment is, in reality, more complex than applying a single formula. The formula considers a number of inputs including the parents' taxable income, number of children, and the amount of time parents care for the children.

The Child Support formula has been contentious since the start of the Scheme (see, for example, Carr, 1994). Generally, the debate on Child Support has consisted of calls for increasing or decreasing assessments based upon more or less substantiated claims of unfair treatment from either parent. In the opinion of some observers (for example, Wakelin, 1998), debate on underlying policy has been absent. Indeed, arguably the most complete evaluation of the scheme to date

---

*Gary Buurman is a senior lecturer in Applied and International Economics at Massey University. Anthony Farr is a computer consultant.*

(JSC, 1994) concluded that the existing formula was 'arbitrary'. This paper represents an attempt to provide a policy underpinning to the debate on Child Support by applying equity principles to Child Support.

## **Equity**

### *Background*

Economists usually invoke two equity principles or objectives known as horizontal and vertical. Horizontal equity is the principle that people in equal positions should be treated equally. Vertical equity, as applied to taxation, holds that those who have a greater ability to pay higher taxes should do so.

While both equity concepts have general appeal, they are not based on economic analysis. For example, Boadway (1979) sees horizontal equity as more a principle of justice requiring a value judgement and argues that economists assume that most, if not all, members of society would accept it. Nevertheless, economic theory recognizes a number of issues that must be reconciled before the concepts are applied. In general, these issues concern what is used to define ability to pay (equal position), relative position and equal sacrifice.

Economic theory holds that the ideal way to judge initial position (from the literature on taxation) would be to base it on the overall well-being of an individual. In economic theory this is termed an individual's utility (satisfaction). Theoretically, utility would be used as the (tax) base defining equal and relative position, with equal sacrifice defined as equal utility sacrifices in terms of tax payments (see Cullis and Jones, 1992). However, economists are reluctant to make interpersonal utility comparisons and it is impossible to use them as a measure in practice, since all personal circumstances would have to be taken into account. In contrast, reality requires a transfer system based on observable variables that can be measured.

Thus economists and policy makers rely on measures like income or expenditure. As noted by Boadway (1979), the most common measure of ability to pay is the comprehensive income base, which includes all the income of an individual regardless of its source or its use. Although most authors consider this to be the best available measure of the overall well-being of individuals, there is still room for debate. For example, if two brothers with an equal ability to earn devote different amounts of time to work and leisure, they would earn different incomes. A tax structure based on income means that the one earning the greater income pays a greater amount of tax. Hence it can be argued that income represents the outcome of peoples' decisions, not equal position, and that both brothers started with an equal ability to pay (see Rosen, 1992). There can also be problems on the consumption side as with different tax rates on beer and wine, and one brother preferring each drink. According to Ng (1979), if all factors are allowed to enter into the determination of equal circumstances, the principle of horizontal equity becomes vacuous since all individuals are different. Many economists would argue that differences in tastes or earning ability are not



sufficient to render the equity concepts useless if individuals face the same opportunity sets.

#### *Application of Equity to Child Support*

While the equity principles were originally applied to taxation, they have since been applied to a large number of areas in which equity is seen as a desirable feature. These include the distribution of welfare (see Brownlee and Macdonald, 1986), the provision of government services such as housing and the distributional consequences in specific markets (see Friedman, 1984). The authors submit that these principles can also be applied to Child Support. This argument rests in part on the general principle that fairness is a desirable property of Child Support just as it is desirable in other areas such as those referenced above. However we also find support for our submission in the legislation.

In the literature on Child Support, Sorensen and Halpern (2000) have identified inconsistencies in judicial awards as an issue of horizontal equity. However the authors are not aware of any previous attempt to apply equity principles to the two parents who are the parties to a Child Support transaction. In the Australian context, this is somewhat surprising since horizontal equity is mentioned fairly explicitly as an objective in Section 4 of the CSA. Similarly, as a result of the consideration of vertical equity in the development of the Child Support formula (Consultative Group, 1988), Section 4 of the Act suggests that it also is, at least, in line with the objectives of the legislators. Section 4 of the CSA, in part, states:

(2) Particular objects of this Act include ensuring:

(a) that the level of financial support to be provided by parents for their children is determined according to their capacity to provide financial support and, in particular, that parents with a like capacity to provide financial support for their children should provide like amounts of financial support; ...

(d) that children share in changes in the standard of living of both their parents, whether or not they are living with both or either of them.

#### *Ability to Pay*

In general, the process of applying equity principles to Child Support is straightforward. In Australia (as elsewhere), Child Support is taken from after tax income. Consequently, it is common to consider after tax income as the basis for measuring ability to pay. Other government transfers such as social security payments can also be seen to contribute to ability to pay. Hence the money available to a person after tax, social security and all other transfers (except Child Support) is the starting point for measuring ability to pay. We use the term Disposable income for this.

Much literature on Child Support goes further and sets aside a reserve for self-support. If this reserve is deducted from Disposable income, this gives a more refined measure of ability to pay by allowing for a basic cost of living. The Australian Child support formula, like most recent formulae in other jurisdictions, sets aside a self-support reserve before calculating Child Support (CSA:Section 3).

In the remainder of this paper we use the term Discretionary income for this measure. The amount set aside for self-support is described in the following section. Note that, as we use the term in this paper, Discretionary income aligns with the description above and does not allow for support of the children, which is considered separately.

## **Model**

This paper deals with separated parents who invoke the Child Support formula and each has time with the children. To simplify the discussion, they are assumed to have not re-partnered and to have no children from other relationships. Basically the model considers a single scenario in terms of the circumstances of the two parents. The model is then iterated to determine different outcomes in response to changing the important inputs such as the number of children, the percentage of nights spent with each parent and the parent's earnings from paid employment. Parameters include the costs of children data (described below), tax rates and tables of social security benefits.

All income related parameters were inflated to May 2002 in proportion to movements of the Average Weekly Earnings — All employees (ABS, 2002a). All cost of living data were inflated to 31 March 2002 in proportion to movements of the All Groups CPI (ABS, 2002b). Social security entitlements, taxation and Child Support transfers are at the level of 1 April 2002. The model does not consider transfers between parents outside the Child Support system, although they were found to be common by the Child Support Evaluation Advisory Group (1992). Such transfers mean that the transfers between parents are not (solely) based on the Child Support formula. While equity considerations could, in principle, be applied to such arbitrary transfers, that is not the focus of this paper.

The effects of differences between parents in their asset base are not considered. This is in line with the treatment in most other literature and also in all national Child Support formulae known to the authors. However, we note that differences in assets are common and affect household expenditure requirements.

## *Costs*

Our model requires as input, data on costs of living applicable to separated parents and to raising children. Unfortunately most available data deal only with intact families. However, a few Australian studies that deal with separated parents have recently become available. This paper draws heavily upon two of them.

Henman and Mitchell (2001) develop costs for households consisting of a NRP and, for part of the year, either one or two children. This research is the only

available Australian data on the direct costs of children to NRPs. Henman (2001) develops a number of different cost estimates for children in the household of the RP. Here, the data used is taken from tables D1 and D2 since these were seen as being most comparable to the Henman and Mitchell (2001) study. Both studies estimate the cost for a parent to live at a low-cost level. This level is used as the self-support reserve set aside when calculating Discretionary income.

By virtue of using the two studies described above, the current paper inherits some assumptions:

- First, the study by Henman (2001) assumes that the children reside with the RP for 365 nights per year (whereas the model deals with scenarios in which both parents have time with the children). Rather than make adjustments, we use Henman's figures.
- Second, the cost data described above deals with certain specific households. In particular, households are located in Sydney in privately rented accommodation. The single child is assumed to be a girl aged six. With two children, a fourteen-year old boy joins the girl. Larger groups of siblings are not dealt with in the Henman and Mitchell study.
- Both the Henman and Mitchell (2001) and the Henman (2001) studies derive, more or less directly, from the key study by Saunders *et al* (1998). Consequently they share many normative assumptions about what constitutes appropriate expenditure. It is unclear to what extent these assumptions impact on the available data.
- Both studies use a budget standards methodology — also known as a basket of goods approach. Consequently they aim to state what *should* be spent on children — as judged by a committee of experts on nutrition, health economics, housing, clothing needs, consumer behaviour, and household budgeting. In the studies used here, two costs are derived — a ‘low cost’ standard and a ‘modest but adequate’ standard. These standards exist independently of constraints such as income and do not vary significantly with income. By virtue of using these studies, this paper is judging Child Support with reference to what should be spent on children.
- Other studies exist which use expenditure survey data to arrive at estimates of what parents *actually* spend on their children. (For a discussion of these two methodologies see McDonald, 1990.) The approach used in the current paper could also apply to expenditure survey studies — in which case we would be judging Child Support with reference to what parents typically spend. Regrettably there is no expenditure survey data available for NRPs, so this analysis is not yet possible.

### *Operation of the Model*

The logic of the model including key assumptions is summarised below. Discussion under most of the steps is divided into method and example. Method refers to a general description of the step while the example is a numerical

illustration of the calculations. The example considers a scenario in which the NRP earns \$33,000 per year. The example also assumes two children who spend 23 per cent of their time with the NRP and RP earnings of \$24,400 per year.

**Figure 1: Demonstration of Model**

Step	Method	NRP	RP
1. Calculate the parenting payment and Newstart.	Assume RPs qualify for Newstart subject to income test.	The NRP does not qualify for any form of pension.	The RP qualifies for a parenting payment of \$3,034/year.
2. Determine tax liability and after tax income of parents.		Tax = \$6,775/ year After tax income = \$26,225	Tax = \$4,322/year After tax income = \$23,112.
3. Calculate the child support transfer.	Self-support is set at \$11,740 of NRP's taxable income. For two children, Child Support = 27% of remainder.	Adjusted Income: \$33,000 - \$11,740 = \$21,260 Child Support: 27% of \$21,260 = \$5,740	Child Support = +\$5,740
4. Calculate family tax benefit (including Rent Assistance).	Household cost data used to calculate Rent Assistance. Assumes both parents apply for family tax benefits.	Family Tax Benefit = \$4,486	Family Tax Benefit = \$7,285
5. Determine the costs of the children in each of the two households.	Uses 'low cost' figures for earnings less than AWE; 'modest' figures for earnings greater than twice AWE; and linear transition between these earnings.	Low cost range = \$6,942/ year. Henman & Mitchell (2001) data for 15%, 20% and 30% shares of time. The model interpolates to determine costs at other share times.	Low cost range = \$12,179/year.
6. Determine the Discretionary income of parents.	Allows for parent's cost of living the 'low cost' figures.	\$26,225 (after tax) + \$4,486 (FTB) - \$18,054 = \$12,657	\$23,112 + \$7,285 - \$17,743 = \$12,654
7. Determine each parent's contribution to the costs of the children.	Parent's contribution is sourced only from own income. Social Security payments otherwise due to a parent without children are deducted from Government's contribution.	\$6,942 (cost: step 5) - \$4,486 (Social Security payment: step 4) + \$0 (Social Security without children) + \$5,740 (Child Support from step 3) = \$8,196 (approximately).	\$12,179 (cost: step 5) - \$7,285 (FTB: step 4) - \$2,776 (after tax parenting payment: step 2) + \$0 (Social Security without children) - \$5,740 (child support: step 3) = -\$3,622 RP's contribution is negative as total children cost are less than Social Security and Child Support payments.

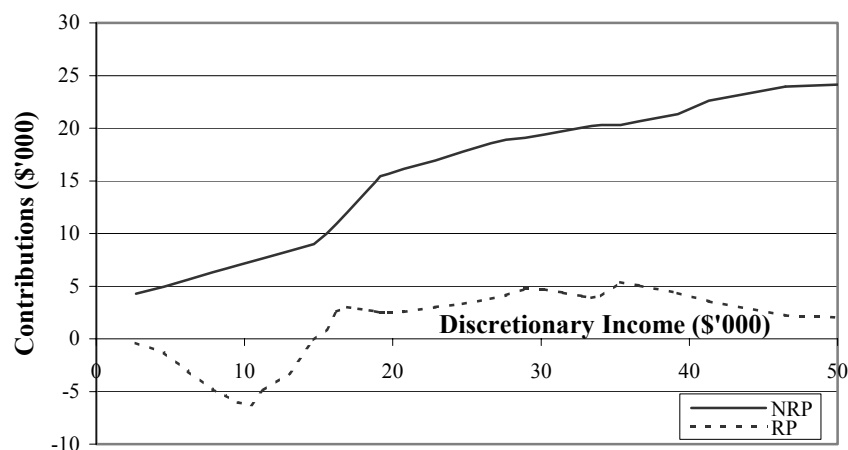
The interested reader can confirm that the above example yields points on the curves plotted in the first graph of the following section. The points relate to a Discretionary income of \$12,655 — the Discretionary income of both parents. Reading off this value on the X-axis, the corresponding value on the curve of NRP contributions is \$8,196 and the value on the curve for RP contributions is -\$3,622.

### Horizontal Equity

It is possible to use quantitative analysis to test whether the existing Child Support formula fulfils the objective of horizontal equity. We have modelled a large number of scenarios to determine the formula's divergence from horizontal equity and the sensitivity of this divergence to different parameters. Here we illustrate by providing two examples.

When examining horizontal equity in the model, we are dealing with equal position and only need to define equal ability to pay and equal sacrifice. As discussed previously, people are deemed to be in an equal position if they have the same level of Discretionary income. Equal sacrifice for horizontal equity means an equal dollar value contribution to child support.

**Figure 2: Contribution of Parents with Equal Discretionary Income, 2 Children and 23/77 Parental Percentage Time-Shares**



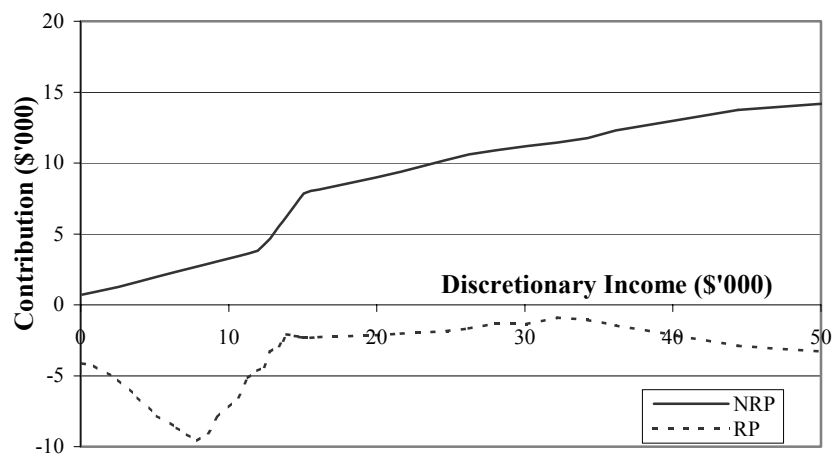
In the first scenario, we assume that there are two children whose time is split according to the standard Family Court orders where the NRP has access every second weekend and half of the school holidays. This yields a split of the time with the children of 77 per cent to the RP and 23 per cent to the NRP. The results are shown in Figure 2. The two curves represent the contributions of each parent and the distance between the lines represents the inequality in financial support. For example, given equal parental Discretionary income of \$20,000, the NRP

would be contributing \$15,806. The RP would only be providing about \$2,528. Horizontal equity would require equal contributions from the parents.

Readers may note that at high income levels, the contribution of the RP falls. This occurs because the costs of children do not increase with increases in household income in this range. Nonetheless the Child Support paid by the NRP continues to increase. Consequently as household income increases in this range, more of the costs of children are paid by the NRP leaving less for the RP to pay.

The results of our second scenario are shown in Figure 3. Here, we assume there is only one child and that time with the child is shared 15 per cent to the NRP versus 85 per cent to the RP. (The 15/85 share was chosen as it is described in the Henman and Mitchell study.)

**Figure 3: Contribution of Parents with Equal Discretionary Income, 1 Child and 15/85 Parental Percentage Time-Shares**



In this scenario, the shapes of the curves showing contributions of the NRP and RP are similar to Figure 2. Not surprisingly, the contributions required for one child are significantly lower than for two. What the reader may find surprising here is that contributions of the RP are negative for all levels of Discretionary income. (Obviously this finding cannot be generalised to other scenarios.) Essentially this is a result of the interplay of two factors. First, data from both Henman (2001) and also the preceding study by Saunders *et al* (1998) shows that the incremental cost of a second child is significantly greater than that of the first child. In contrast, both Child Support and government benefits provide significantly more for the first child than the second. The result is that, in comparison with the two children scenario, the greater Child Support and benefits for the first child exceed the lesser cost of the first child.

In both cases presented in this section, much greater financial support is demanded from the NRP than from the RP. Similar results applied in other scenarios we examined. We conclude that the current Child Support formula diverges from the objective of horizontal equity.

## Vertical Equity

We have also tested whether the Child Support formula meets the objective of vertical equity. As with horizontal equity, we illustrate with a few examples out of the many scenarios tested. Again we have varied the number of children and the time spent with each parent. Recall that we assume that equal contribution (sacrifice) is measured by an equal percentage payment out of the income measure. This is consistent with the form of the current Child Support formula. The analogous policy in taxation would be proportional rather than progressive taxation. In the following parts of this section we use proportional contributions. The final part of this section considers some other policy options for vertical equity.

Given the aim of proportional contributions, to achieve vertical equity between the RP and the NRP, both parents would have to contribute an equal percentage of the income measure towards the costs of the children. Consequently, vertical equity requires that the ratio (here termed the Sacrifice Ratio) between the RP's and NRP's percentage contributions would be equal to one. Conversely, if the Child Support formula required NRPs to contribute twice the percentage of income as RPs, then the Sacrifice Ratio (NRP/RP) would be two. Obviously a Sacrifice Ratio other than one does not meet our vertical equity criterion. The extent to which the Ratio differs from one provides a measure of the extent to which the actual Child Support formula differs from the vertical equity objective.

Figure 4 shows Sacrifice Ratios (on the Y-axis) for the two children, 77/23 percentage time-share split scenario. Here it is assumed that the RP's earnings are equal to Average Weekly Earnings for all (both full and part-time) employees, at \$35,797 per annum.<sup>1</sup> The X-axis shows varying earning levels for the NRP. Note that this is different to the graphs used for horizontal equity. In principle, any income related measure could be used on the X-axis with essentially the same result. We chose earnings since they are computationally more tractable in our model.

Two curves are plotted:

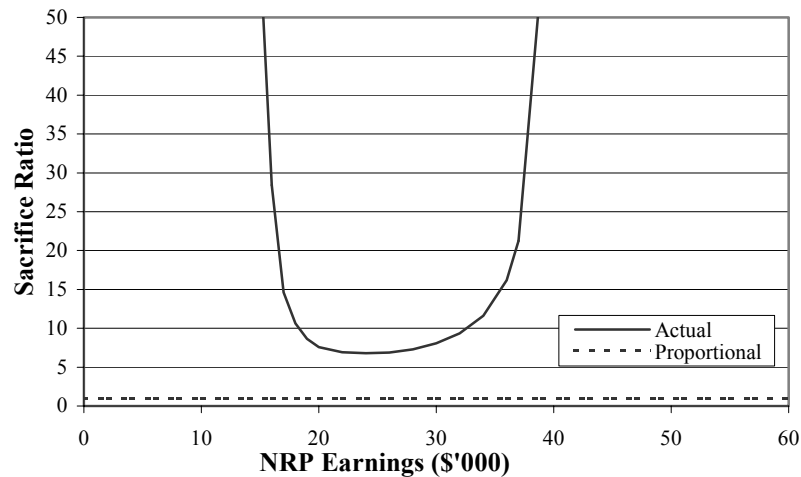
- The solid line shows actual Sacrifice Ratios for the current Child Support formula.
- The dotted line shows the Sacrifice Ratios required to meet our baseline goal (proportional contributions). As described earlier the required Sacrifice Ratio is always one so this is a simple straight line.

---

<sup>1</sup> We use AWE in part because it is less arbitrary than most other choices. Also it yields graphs which are of interest. In most cases, both RPs and also NRPs have incomes less than AWE. However where the RP's income is low, the RP's net contribution is negative at almost all levels of NRP earnings. This puts the curves 'off the scale'. Our later discussion of Figure 7 investigates situations of low RP earnings.

The distance between the lines for the actual Sacrifice Ratio and the line for our equity baseline gives a measure of the extent to which the Child Support formula fails to meet the objective.

**Figure 4: Relative Contributions for Varying NRP Earnings, 2 Children and 23/77 Parental Percentage Time-Shares**



As an example, if the situation for a NRP earning \$30,000 is examined, one finds that the NRP has a Discretionary income of \$10,601 and pays \$7,281 towards child costs. Thus the percentage of Discretionary income that the NRP contributes towards the costs of the children is:

$$\frac{\text{Contribution}}{\text{Discretionary Income}} = \frac{\$7281}{\$10601} = 69\%$$

Similarly when we consider the situation of a RP in the given scenario (two children, \$35,797 earnings etc), the model yields a Discretionary income of \$16,358 and a contribution towards child costs of \$1,393. Thus the percentage of Discretionary income that the RP contributes towards the costs of the children is:

$$\frac{\text{Contribution}}{\text{Discretionary Income}} = \frac{\$1393}{\$16358} = 8.5\%$$

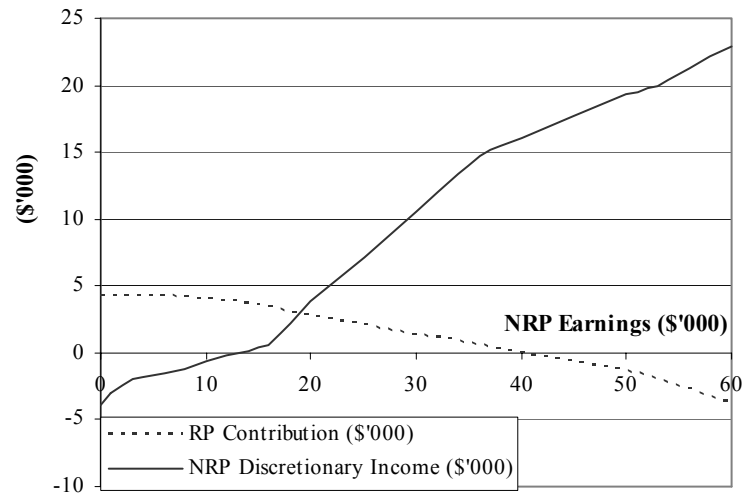
This gives a Sacrifice Ratio (Discretionary income) of 69 per cent divided by 8.5 per cent or approximately 8.1. Throughout the graph shown in Figure 4, the Sacrifice Ratio is significantly in excess of one.

Readers will note that in many cases, Sacrifice Ratios are 'off the scale' being greater than 50 to one. This occurs at both low and high earning levels. At earnings less than \$14,000 per annum, Discretionary income for a NRP is effectively zero since income (including Social Security payments) is less than

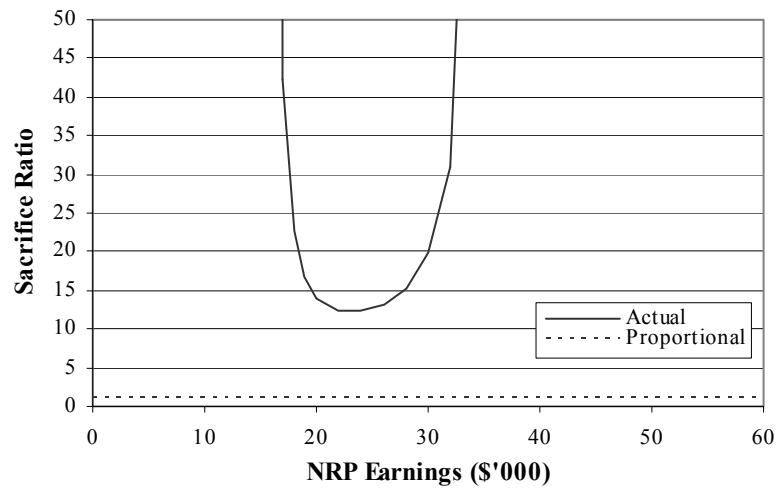


required to meet the low cost standard of living given by Henman (2001). At greater than \$40,000 the Sacrifice Ratio is again 'off the scale' — the cause being that the RP makes no contribution to child costs since Child Support together with the Government contribution amounts to more than the cost of the children. Figure 5 demonstrates this effect.

**Figure 5: RP's Contribution to Costs and NRP Discretionary Income, 2 Children and 23/77 Parental Percentage Time-Shares**



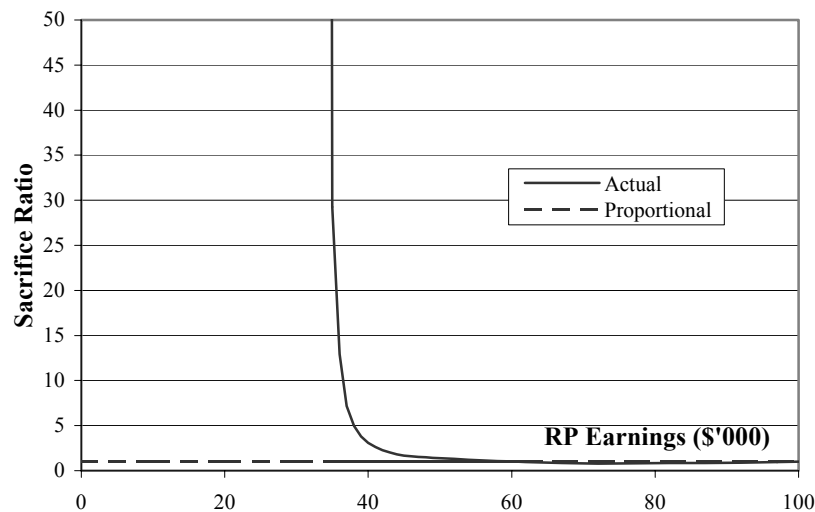
**Figure 6: Relative Contributions for Varying NRP Earnings, 2 Children and 15/85 Parental Percentage Time-Shares**



Alternative scenarios were also considered for vertical equity. Figure 6 shows similar results (to Figure 4) for a 15/85 percentage time-share split scenario, with Sacrifice Ratios that are now even higher. This occurs mainly because RPs receive approximately \$735 per year more in Government contributions when their time with the children increases from 77 per cent of the year to 85 per cent. This causes a significant decrease in the RP's contributions.

In Figure 7, the sensitivity of these results is checked for different levels of RP income. The scenario is the same as for that in Figure 4 except that we vary RP earnings on the X-axis with NRP earnings held constant at AWE. There are broadly three regions shown in the graph. Below RP earnings of approximately \$34,000 the graph is off the scale since the RP makes no contribution to child costs — those costs being less than the amount received in Child Support together with Government contributions. Where RP earnings are around \$40,000, Sacrifice Ratios are much greater than one but no longer off the scale. Where the RP earns \$50,000 or more per annum, Sacrifice Ratios are close to vertical equity. This is principally due to the fact that as RP income increases through this range, the Child Support formula reduces the contribution from the NRP. Consequently more of the child costs are borne by the RP. The combination of the reduced contribution from the NRP and the increased contribution from the RP gives Sacrifice Ratios close to one.

**Figure 7: Relative Contributions for Varying RP Earnings, 2 Children and 23/77 Parental Percentage Time-Shares**



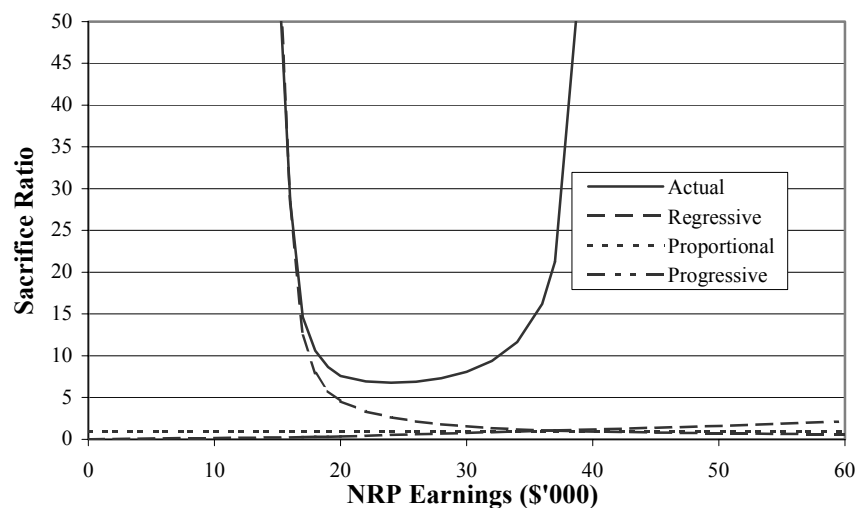
A case could be made that the parent with greater income should make a higher percentage contribution — in similar fashion to progressive taxation rates. At the opposite extreme, a case could also be made for a regressive approach — for example, that parents should each contribute half the cost of children. In Figure 8 we examine these alternative objectives.

The scenario used for Figure 8 is the same as that in Figure 4 — two children, 77/23 percentage time-share split and RP earnings equal to AWE. The same two lines plotted in Figure 4 appear again in Figure 8. As well, two further lines are plotted corresponding to the progressive and regressive goals just described:

- A regressive objective — that parents should each contribute half the cost of children. We note in passing that this would not qualify as a vertical equity policy.
- A progressive objective based upon Australia's progressive taxation rates. In particular the policy graphed is that parents should each contribute in proportion to the amount of income tax they pay.

As with Figure 4, the distance between the lines for actual Sacrifice Ratio and the line for a given objective is a measure of the extent to which the current Child Support formula fails to meet the objective.

**Figure 8: Relative Contributions for Varying NRP Earnings, 2 Children and 23/77 Parental Percentage Time-Shares: Alternative Policies**



Clearly, the current Child Support formula diverges from the progressive objective — essentially to the same extent that it does from the proportional objective. Indeed, the difference between the progressive and proportional objectives is minor compared to the divergence shown by the current formula. However the current Child Support formula does align with the regressive objective at low levels of NRP earnings. The reason is that both the current formula and the regressive objective require the NRP to contribute Child Support even when the NRP has a Discretionary income of zero. Consequently the curves

for both the current formula and the regressive policy disappear off the scale where NRP Discretionary income approaches zero. (As noted in Figure 5, this occurs where NRP earnings are approximately \$14,000.) Where the NRP's Discretionary income is significantly greater than zero, the current formula diverges greatly from the regressive objective as well.

In summary, for vertical equity, only when RP earnings are significantly more than average weekly earnings does the current Child Support formula share the costs of children in proportion to ability to provide financial support. In all other cases, the NRP is required to contribute a far greater proportion of their income than is the RP. Moreover the current formula exhibits significant divergence from other vertical equity objectives we have examined.

### Sensitivity Testing

The robustness of the findings was evaluated by varying the assumptions on ability to pay and the cost of children. Disposable income, meaning that an allowance for self-support by the parent is not deducted, was used as an alternative measure of ability to pay. In preceding sections, the self-support reserve was set to equal the low cost standard of living from Henman (2001) or Henman and Mitchell (2001). However, it could be argued that a poverty line would be a better measure and that the Henman (and Mitchell) costs are too high to be considered a legitimate poverty line — see Saunders and Tsumori (2002). Regrettably, there appears to be no generally accepted poverty line. However, to test for robustness, we examine the (extreme) case where the self-support reserve is set to zero. In this case Disposable income will equal Discretionary income. The results can equally be viewed as examining the case where self-support reserve is zero.

The cost of children in the preceding discussion was based on data from Henman (2001) for the costs of children to the RP and from Henman and Mitchell (2001) for the costs of children to the NRP and reflected judgements on what *should* be spent on children. For the sensitivity analysis, data from Valenzuela (1999) were used to estimate the costs of children to the RP. These are based on expenditure survey data and estimate what parents *actually* spend on their children. For NRPs, because there are no relevant expenditure survey results pertaining to them, the Henman and Mitchell (2001) data was used as before. This meant that NRP costs did not change with respect to income. It also meant that budget standards data (for the NRP) was compared with expenditure survey data (for the RP). In the restricted context of sensitivity testing, the methodological difficulties this involved were ignored.

The earlier findings continued to apply when Disposable income was used an alternative measure of ability to pay. In summary, if Disposable income is used in place of Discretionary income then the current Child Support formula still diverges from a Horizontal Equity objective. At moderate to high NRP earnings the formula also diverges from a Vertical Equity objective. However, for Vertical Equity, at low NRP earnings the formula only requires approximately 50 per cent

greater sacrifice from the NRP than the RP when comparison is made based on Disposable income.

The findings also continued to apply when the analysis was repeated using alternative costs of children data for the RP. In this case, however, the divergence from Horizontal equity was not so great at high incomes.

The robustness of our conclusions should not be surprising. Our primary analysis shows that the current formula diverges from equity policies by an order of magnitude. Consequently even the major shifts in assumptions that were examined here did not shift the overall conclusions significantly.

## **Discussion and Conclusions**

We have argued that the concepts of horizontal and vertical equity can be used to examine whether Child Support transfers are set at an appropriate level. To investigate this issue we have constructed a model that addresses the levels of Child Support in comparison to the levels dictated by equity principles.

The principle of horizontal equity is important because it is generally agreed that people in the same position should be treated equally. Hence horizontal equity serves as a baseline for evaluating whether a child support formula meets a basic hurdle of acceptability. This is true despite the fact that horizontal equity is only directly applicable in the (relatively extraordinary) circumstance that two parents are in equal positions. Our results show that the current child support formula yields transfers that diverge greatly from those dictated by horizontal equity. In particular the RP is treated more favorably than the NRP.

We have also evaluated the current Child Support formula against the objective of vertical equity. Vertical equity might be more relevant for actual use because it concerns parents in different income situations. Three different vertical equity policies were compared with the current formula and the formula showed significant divergence from each of them. Only in scenarios in which the RP has a high income does the formula yield transfers that approximate a vertical equity policy. In all other circumstances the RP is treated more favorably than the NRP in regard to vertical equity.

Our presentation of results in respect to vertical equity is based upon graphs of relative contributions against income. The format is useful because it summarises sensitivity to income and is readily adaptable to different assumptions and to accommodate comparison against different objectives.

Our central conclusion is that equity principles can be applied to Child Support and that they provide considerable illumination. We also conclude that the current Australian formula is overly onerous to NRPs and overly generous to RPs when compared against equity principles.

## **References**

Australian Bureau of Statistics (2002a), *Average Weekly Earnings of Employees*, Table 6302.0, AUSSTATS, Canberra.

- Australian Bureau of Statistics (2002b), *Consumer Price Index*, Table 6401.3B, AUSSTATS, Canberra.
- Boadway, R. (1979), *Public Sector Economics*, Winthrop Publishers, Cambridge, USA.
- Carr, Senator K. (1994), *Senate Hansard*, 5 December, pp. 3893-3894.
- Child Support Evaluation Advisory Group (1992), *Child Support in Australia*, Vol 2, AGPS, Canberra.
- Consultative Group (1988), *Child Support: Formula for Australia*, Child Support Consultative Group Australia, AGPS, Canberra.
- Cullis, J. and P. Jones (1992), *Public Finance and Public Choice: Analytical Perspectives*, McGraw-Hill, London.
- Friedman, L. (1984), *Microeconomic Policy Analysis*, McGraw-Hill, New York.
- Henman, P. (2001), *Updating Australian Budget Standards Costs of Children Estimates*, Policy Research Paper No. 7, Department of Family and Community Services, Canberra.
- Henman, P. and K. Mitchell (2001), 'Estimating the Costs of Contact for Non-Resident Parents: A Budget Standards Approach', *Journal of Social Policy* 30:495-520.
- Joint Select Committee on Certain Family Law Issues (1994), *Child Support Scheme — Operation and Effectiveness: Recommendations and Conclusions*, Australian Parliamentary Library, Canberra.
- McDonald, P. (1990), 'The Costs of Children: A Review of Methods and Results', *Family Matters* 27:18-22.
- Ng, Y. (1979), *Welfare Economics: Introduction and development of Basic Concepts*, MacMillan, London.
- Saunders, P., J. Chalmers, M. McHugh, C. Murray, M. Bittman and B. Bradbury (1998), *Development of Indicative Budget Standards for Australia*, Policy Research Paper no 74, Department of Social Security, Canberra.
- Saunders, P. and K. Tsumori (2002), *Poverty in Australia: Beyond the Rhetoric*, CIS Policy Monograph (57), CIS Ltd., St Leonards, NSW.
- Scott, I. (1989), 'The Child Maintenance Formula and the Costs of Children', *Family Matters* 24:23-25.
- Sorensen, E. and A. Halpern, (2000), 'Child support reforms: Who has benefited?', *Focus* 21:38-41.
- Rosen, H. (1992), *Public Finance*, 3<sup>rd</sup> ed., Irwin, Homewood, USA.
- Valenzuela R. (1999), 'Costs of children in Australian households', *Family Matters* 53(Winter):71-76.
- Wakelin, B. (1998), *House of Representatives Hansard*, 2 April, p. 2463.
- Woods, M. (1999), *The Behaviour and Expenditures of Non-Resident Parents During Contact Visits*, Research Paper No. 75, Department of Family and Community Services, Canberra.

*The authors would like to thank the referees and Stuart Birks for their considerable contributions. Thanks also to Dr Paul Henman for his input. Any errors, however, are attributable to the authors.*

## **Subsidies for Private Community Services: The Case of School Education**

**Buly Cardak and Phillip Hone**

Public policy makers in Australia face an economic and political environment with actual and potential conflicts that make policy formulation difficult. For example, there is a strong commitment on both sides of Parliament to surplus, or at least, balanced budgets. At the same time policy makers in the major parties appear to believe that the median voters in the marginal electorates are opposed to higher income taxes. These factors are coupled with wider community concern about the level of provision of government services in key areas such as education, hospitals and transport. As a consequence governments are in a policy dilemma: they are constantly under pressure to improve the provision of public services, while there are also strong expectations that they will manage the macro economy in such a way as to show budgetary restraint without raising taxes. In effect, they are being asked to provide more services without raising more revenue.

The only avenue open to governments to achieve these potentially competing goals is to be more efficient in the use of the existing government revenue base. One of the more controversial policy initiatives to deal with these competing demands has been the use of government subsidies to encourage the growth of private sector provision of community services. In this regard the Federal Government encourages people to use the private rather than public hospital systems by subsidising private health insurance. Similarly, significant Commonwealth funding is made available to non-government schools. These policies are designed, at least in part, to use some public funds to leverage greater private contributions into the provision of community services.

Both health and education subsidies have been the subject of considerable debate. While there is a number of issues and principles that are disputed in these debates, a key issue relates to how to ensure adequate funding for the public sector. Opponents argue that governments should ensure that scarce public funds are directed towards the public systems and not diverted to encourage growth of a competing system. Supporters have contended that in a political environment where higher taxes are not perceived as being feasible, the growth of the private sector is necessary to ensure that the demand on the public sector does not outstrip the supply of public sector resources.

---

*Buly Cardak is a lecturer in Economics in the Department of Economics and Finance, La Trobe University, and Phillip Hone is a Senior Lecturer in Economics at the School of Economics, Deakin University.*

In this paper we examine the economics of public funding of private sector provision of community services by examining the case of Commonwealth funding of the non-government school system.

Public schools are funded principally by state governments, while government funding of non-government schools is primarily a Commonwealth initiative. Non-government schools have received Commonwealth subsidies in one form or another since 1964 with formal recurrent grants commencing in 1970. Although the magnitude of the funding involved is not substantial in a national budgetary sense, the level of funding is growing and, most importantly, it has always been significant in a wider public policy context.

**Table 1: Government Expenditure on Government and Non-Government Schools, 1977 and 1999-2000**

	Government Schools		Non-Government Schools	
	1977 <sup>a</sup>	1999-2000	1977 <sup>a</sup>	1999-2000
Commonwealth Expenditure (\$m)	--	1786	--	2900
State Expenditure (\$m)	--	14806	--	1195
Total Public Expenditure (\$m)	9656 <sup>b</sup>	16592	1051 <sup>b</sup>	4096
Enrolments (Primary + Secondary)	2364300	2259829 <sup>c</sup>	630900	999565 <sup>c</sup>
Total Spending per student (\$)	4084	7344	1666	4139

Notes: a: Values for 1977 have been indexed to 2000 dollar terms using the CPI.

b: Education expenditure data for 1977 is not available by source (Commonwealth or State Government) in the Commonwealth Schools Commission publication 'Australian Students and Their Schools'.

c: These enrolment figures are for 2000 rather than 1999-2000.

Sources: Steering Committee for the Review of Commonwealth/State Service Provision SCRCSSP (2002): Attachment 3A (Tables 3A.1, 3A.2 and 3A.11);

Ministerial Council on Education, Employment, Training and Youth Affairs (2003): Appendix 1 (Table 27);

Commonwealth Schools Commission and Australian Bureau of Statistics (1979).



In 1977 a total of \$1051 million (adjusted to year 2000 dollars) of Commonwealth and State funds were allocated to non-government schools. This represented \$1,666 per student in non-government schools and was equivalent to 41 per cent of the average cost of educating a student at a government school. By the year 2000, total public funding to non-government schools had grown to just over \$4,000 million, with \$2,900 million provided by the Commonwealth. Total government funding was equivalent to \$4139 per student enrolled outside the government system or 59 per cent of the cost of educating a student in a government secondary school.

The aim of this paper is to explore some of the economic issues associated with this funding strategy. This analysis is essentially positive in perspective. We do not address normative issues such as the desirability of potentially transferring tax revenue to middle and high income voters. Moreover, while our analysis is consistent with the notion that the current regime of subsidies for private schooling could be a means of making increasingly scarce public funding available for the public school system or other government services, we are not arguing that budgetary efficiency is the only or even the most important objective that government is pursuing with this policy approach. Rather, we focus on the task of identifying the conditions under which the current regime of subsidies for private schooling could reduce the effective real demand on public resources because this is an important policy issue in its own right.

Given the controversy surrounding recent changes in the Commonwealth funding mechanism for non-government schools (see, for example, Maslen, 2000; Donnelly, 2001; and Sydney Morning Herald, 2002) we outline the previous and existing Commonwealth funding arrangements in the next section. We then move to identify the conditions that need to exist for public subsidies of private schools to potentially result in more public resources for students at government schools and consider if these conditions are likely to be met in the current Australian education sector. This is followed by an evaluation of the current subsidy program in terms of its effectiveness in providing choice and improving resources in both public and private schools. The final section highlights the policy implications and presents a summary of conclusions.

### **Government Subsidies for Non-Government Schools**

The two most recent schemes of allocating Commonwealth funds to non-government schools are the Education Resources Index (ERI) funding scheme, which began in 1985 and applied until 2001 when the Socio-Economic Status (SES) funding scheme was partially introduced.

The ERI funding scheme was intended to measure a school's private income and compare it to 'The Community Standard', or later, 'The Average Government School Recurrent Costs', and to subsidise schools in a way that would bring them closer to the relevant standard, thereby providing students with similar funding resources independent of schooling system. The ERI defined school income broadly to include not only private fee income but also the estimated dollar value

of volunteered services along with other non-cash donations, such as artwork. The scheme was based on a 12 category rating system, which was determined at the outset of the scheme, 1985. Ratings were not recalculated regularly and it was difficult to move to a lower ERI, which would entitle a school to greater funding. If the intention was to assist needy students attending private schools, the scheme seems to have been flawed in that it was based on school income rather than a student's family income. In any case, after a Commonwealth government review of schools funding, started in 1997, the funding system was changed in 2000. (Department of Employment, Education, Training and Youth Affairs, 1997 and 1998).

Currently, non-government schools outside the Catholic education system receive recurrent Commonwealth funding based on the SES of their respective school communities. All non-government schools had the option to continue on the ERI funding scheme and the Catholic school system — which comprises about 60 per cent of non-government school enrolments — elected to continue to be funded through the ERI framework. Non-government students comprise around 30 per cent of total enrolments, implying the SES funding scheme applies to 10 per cent of the student population. This decision is to be reviewed after four years, in 2005.

A school's SES score is based on the places of residence of students and their families and draws on Australian Bureau of Statistics Census data. The SES index used by the Department of Education, Training and Youth Affairs (DETYA) comprises three dimensions: occupation, education and household income, with extra weight given to households with school age children (family income). Because of Census confidentiality requirements, students are attributed with average occupation, education and income values for the Census Collection District in which they reside. These values take the form of a standardised index. To calculate a school's SES index, the various attribute indices of each individual in the school community are averaged and then weighted into an aggregate SES index value for the school.

For example, the occupation indices for every student in the school are averaged to provide the schools average occupation index. This is also carried out for education levels and income. In order to place more weight on households with children, the income dimension is split into overall average income and average income of households with school age children. Once these average school dimension indices are calculated, they are aggregated using the following formula:

$$\text{School SES Score} = \frac{\text{Occupation Dimension}}{3} + \frac{\text{Education Dimension}}{3} + \frac{\text{Household Income Dimension}}{6} + \frac{\text{Income of Families with Children Dimension}}{6}$$

These school SES scores are then used to determine Commonwealth General Recurrent Grants per student for non-government schools outside the Catholic education system. Recurrent grants are based on Average Government School

Recurrent Costs (AGSRC). The higher a school's SES score the lower is their Commonwealth General Recurrent Grant per student. For example, schools with a SES of 130 or greater will receive 13.7 per cent of AGSRC, the lowest level of funding. The Commonwealth General Recurrent Grant per student rises at a constant rate of 2.5 per cent for every two-point decrease in SES score. The highest Commonwealth General Recurrent Grant per student is 70 per cent of AGSRC for schools with a SES score of 85 or less. (For more details, see DETYA, 2000; and Commonwealth of Australia, 2001.)

### **Efficacy of Subsidies for Non-Government Schooling**

A possible budgetary motivation underpinning government subsidies for private schooling is that by offering the individual a subsidy for private education, individuals may contribute to private education such that the sum of the public and the private contributions exceeds the sum that would be spent if education was provided completely from public resources. Whether this is the case or not depends critically on the nature of both the subsidy arrangements implemented and the competitive market relationships that exist in the educational sector. In this section we consider how market relationships influence the budgetary efficacy of current subsidy arrangements.

These competitive market relationships have both supply and demand elements. The key aspects of the demand relationships in the educational sector relate to how responsive parents are to changes in the cost of private education and to changes in the quality of public education. First, an underlying premise of the current funding arrangement is that reduced private school fees will shift substantial numbers of enrolments from the public sector to the private sector. This responsiveness to private school charges is the price elasticity of demand for private school education. If this elasticity is relatively low, subsidies will result in few students leaving the public sector to move to the private sector.

If a private education subsidy is less than what would be spent on the child in the public sector, a subsidy scheme may operate to increase per student expenditures on the students remaining in the public sector. This is based on the assumption that the cost savings are not used for some other public expenditure or to reduce taxes. Under the Enrolment Benchmark Adjustment (EBA), the Commonwealth sought to appropriate some of the savings for other budgetary areas (Burke, 1996). This practice was modified in 2001 when the government announced its intention to direct funds that were collected under the EBA back into specific education initiatives in the public school system (Kemp, 2001). As a result, the cost effective diversion of students to the private sector can, as far as higher per student spending permits, enhance the quality of the services on offer in the public sector.

However, per student funding is a very crude measure of student funding. For example, to the extent to which higher available funds are directed towards uses not valued by the parents of potential students, higher funding may do little to enhance perceived educational quality. Moreover, the change in the nature of the

socio-economic mix of students remaining in the public system after the subsidy program may be viewed by some parents as an effective decline in the quality of the wider educational environment in the public school system. (For evidence on peer effects, see, for example, Zimmer and Toma, 2000; and Gavrira and Raphael, 2001.)

The enrolment decisions of the parents of school children are not only determined by the relative level of fees — the perception of the quality of the educational experiences offered to students in the two sectors is also likely to have a strong influence on decisions. As students switch into the private system, the marginal (in terms of willingness to switch) public education student will consider the improved public education quality as well as lower private school fees when deciding which system to attend. The extent to which an improvement in the perceived quality of the government sector induces higher enrolments in the government sector, away from the private sector, is referred to as the quality elasticity of demand for government education.

The higher this quality elasticity is, the less effective the subsidy program will be. In effect, the program has two competing and counterbalancing demand drivers. The subsidy program is based on a high price elasticity of demand for private schooling shifting substantial numbers into the private sector. This shift is offset to the extent that the private education subsidy increases per-student funding for public schools, and the extent to which this is perceived as enhancing educational quality there. Therefore there are three key questions on the demand side of the problem:

- What is the price elasticity of demand for private education?
- What will the resultant improvement in the public sector be?
- What is the quality elasticity of demand for the government sector?

Any serious education policy proposing increased private education subsidies in an attempt to improve both public and private education expenditures and outcomes needs to be based on judgments about the magnitudes of these elasticities. However, an issue often ignored when discussing education finance policy is the supply response of providers of private education to changes in subsidies and demand for their educational services.

The existing subsidy program, with its emphasis on parental choice, relies on the price elasticity of supply of private school places being relatively high — there is an implicit assumption that the private school system will take up substantial numbers without marked increases in cost structures and therefore fees. A low price elasticity of supply for private education would imply private school fee increases of a similar magnitude to the change in the subsidy, resulting in no changes in enrolments and large government transfers to private schools. These transfers may be captured by teachers in the form of higher wages or lower student staff ratios, or may be used for capital improvements by private schools, thereby enhancing their competitive edge over public schools. The activity of capitalising increases in government funding by non-government schools is confirmed

empirically in Williams (1984). The problem in such cases is that subsidies may not reduce fees, choice may not be enhanced in a meaningful way, and the only beneficiaries may be the students and families originally using non-government schools. However, it is also possible that increased government funding to government schools may not improve educational outcomes.

There has been no research published in Australia integrating these three relationships to empirically evaluate the subsidy issue. Williams (1984 and 1985) considers the issue of public funding of non-government schools in Australia and provides demand elasticity estimates (based on aggregate data and relationships) while assuming the supply side is perfectly elastic. In the US, Martinello and West (1988 and 1991) and Frey (1991) addresses the simpler, but related question of whether government subsidies of private education, in the form of a tax credit, could reduce the public education budget without making any students worse off in terms of spending. That is, they study a subsidy scheme aimed purely at reducing the overall government contribution to education with all net savings from shifting students out of the public sector into the private system being returned to consolidated revenue, in order to reduce taxes, rather than being used to enhance the public education system. This body of work emphasises how the conclusions that can be drawn on the net impact of subsidy programs are highly sensitive to relative values of the price elasticities of supply and demand for private education.

The implications can be interpreted in two ways. The first is that the expected net saving for any given subsidy will vary depending on the set of elasticities ruling at the time. The second is that for any given set of elasticities, the magnitude of the net saving will depend on the size of the subsidy. The optimal size of the subsidy (optimal in terms of maximising the net savings to the government) is positively related to the size of the elasticities of demand and supply. That is, when parents are highly responsive to reductions in fees and the private school system is willing and able to absorb large numbers of students without major increases in their fee structures, government savings will be maximised by subsidies that are set at relatively high levels. With very low elasticities the optimal initiative from a government aiming to save funds would be to tax students attending private schools or at least offer a zero subsidy. This is because with low elasticities any subsidy does little in terms of inducing students to leave the public sector for private schools and any that do, tend to drive up school fees in the private sector substantially. The results of Frey's (1991) sensitivity analysis are summarised in Table 2 below.

**Table 2: Sensitivity of Government Savings to Values for the Elasticity of Demand and the Elasticity of Supply**

<i>Price Elasticity of Demand</i>	<i>Price Elasticity of Supply</i>	<i>Optimal Subsidy to Base Private Fees (%)</i>	<i>Increase in Private School Fees (%)</i>	<i>Increase in Private School Enrolments (%)</i>	<i>Benefit/Cost Ratio<sup>a</sup></i>
0.5	$\infty$	27	0	14	1.12
0.5	2.0	2	0.4	1	1
0.5	1.0	0	na	na	na
0.5	0.5	0	na	na	na
1.0	$\infty$	78	0	78	1.44
1.0	2.0	52	17	35	1.26
1.0	1.0	27	14	14	1.12
1.0	0.5	0	na	na	na
2.0	$\infty$	102	0	205	1.67
2.0	2.0	78	39	78	1.44
2.0	1.0	52	35	35	1.26
2.0	0.5	2	2	1	1.01
3.5	$\infty$	113	0	396	1.80
3.5	2.0	88	56	112	1.53
3.5	1.0	63	49	49	1.33
3.5	0.5	13	12	6	1.05

Notes: na: with elasticities of these values there is no positive subsidy that will yield a saving for government.

a: Ratio of Reduction in Public Education Cost to Gross Subsidy Cost

Source: Derived from Frey (1991):351(Table 1).

The elasticity values represent the percentage change in quantity demanded or supplied relative to a given percentage change in fees. For example, price elasticities of demand for private schooling of 0.5 or 1.5 indicate that demand for student places in the private sector would rise by five per cent or 15 per cent respectively in response to a 10 per cent reduction in private school fees. Similarly, a price elasticity of supply of two is consistent with a 20 per cent increase in the number of places offered in the private sector in response to a 10 per cent increase in private fees. Under general conditions in a competitive market this is also equivalent to saying that a 20 per cent increase in enrolments would result in a 10 per cent increase in costs and charges.

The symbol  $\infty$  in the price elasticity of supply column indicates that under this scenario the supply of education places in the private sector is assumed to be infinitely responsive to fees charged. This means that the private sector is able to absorb all new students without an increase in student fees. Under this condition governments can always save if they select the right subsidy level. With an infinitely responsive supply side the optimal subsidy ranges from 27 per cent (with a price elasticity of demand of 0.5) up to 113 per cent (with a price elasticity of demand of 3.5). The corresponding benefit cost ratios (shown in the last column) range from 1.12 up to 1.80.

Where the subsequent growth of the private sector would increase private school fees (a price elasticity of supply of less than infinity) the optimal subsidy is lower, as is the benefit cost ratio. For example, with a price elasticity of demand of 1, a variation in the supply elasticity from infinity down to two would reduce the optimal subsidy from 78 per cent to 52 per cent and reduce the benefit cost ratio from 1.44 to 1.26.

In terms of the Australian debate, the savings measured by Frey (1991) offer us some measure of the funds that may be redirected to the remaining public education students, offering potential improvements in public education. However, most of these studies have not considered how the effect of spending these budgetary savings on enhancing the now smaller public sector — the quality elasticity of public education — would change the analysis. Williams (1984 and 1985) acknowledges the response of private school enrolments to improvements in public schools and estimates, using Australian data, the elasticity of private school enrolments to public education expenditure to be  $-0.045$ . Assuming that increased spending in schools increases quality dollar for dollar, Williams' result implies a 10 per cent increase in public education expenditure (and thus quality) would lead to a reduction of 0.45 per cent in private school enrolments. Thus the improved public education system will reduce migration to the private system, albeit a very small response, thereby reducing the potential improvements in the public system.<sup>1</sup> This is why the savings suggested by Frey (1991) should be viewed as

---

<sup>1</sup> Our assumption of a one-for-one relationship between education spending and quality simplifies the analysis. It is possible that education quality would rise by less than the change in expenditures, leading to larger predictions of quality elasticity. For example, if education quality rose by only 1 per cent when education expenditure rose by 10 per cent, we could infer from Williams' numbers that quality elasticity is  $-0.45$ .

an upper bound for the potential benefits to the public education system of private education subsidies.

Frey's (1991) results offer some support to the existing subsidy regime in Australia at the aggregate level. Average subsidies of just over \$4,000 per student (in the year 2000) could be expected to represent a subsidy on private school fees in the order of 30 per cent to 60 per cent (for private school fees of \$14,000 and \$7,000 respectively). Subsidies in this range are consistent with market conditions characterized by values of 1 or greater for both the price elasticity of demand and the price elasticity of supply. This means that the current subsidies could be expected to be effective in reducing aggregate budgetary outlays on education provided that:

- a 10 per cent reduction in private fees resulted in at least a 10 per cent expansion in the demand for private schooling; and
- a 10 per cent increase in enrolments in private schools caused private school fees to increase by no more than 10 per cent.

While one could speculate that elasticity values of at least 1 seem plausible for the Australian system on average, it is important to recall that the SES funding scheme provides markedly different subsidies to different private schools. For example, schools with an SES of at least 130 receive a grant equivalent to 13.7 per cent of AGSRC. If these high SES schools are high fee schools with fees of the order of \$14,000, this payment could represent 6 per cent or 7 per cent subsidy on actual school fees. On the other hand schools with an SES of no more than 85 could receive a subsidy of around 70 per cent of actual fees if they were relatively low fee schools with charges approximating the AGSRC.

Imposing an inverse relationship between the magnitude of the subsidy and the level of a school's SES would only be consistent with maximizing the financial effectiveness of the program if the values of the price elasticities of supply and demand were also both inversely related to the value of the SES. That is, the current funding formula would only optimise the education funding 'surplus' if low SES schools had higher elasticities of demand and supply than high SES schools. It is not clear that this relationship would hold in Australia.

While it seems the relatively rapid growth in the number of low SES schools in recent years is consistent with low SES schools having, on average, a higher price elasticity of supply than their high SES counterparts, the same may not apply to the price elasticity of demand. It is not unreasonable to assume that relatively exclusive private schools are likely to have high elasticities of demand (a substantial number of parents would be willing to move children into these schools if fees were lower) and low elasticities of supply (schools are unwilling to increase enrolments and thereby reduce exclusivity in return for higher fees or schools are at physical capacity, as waiting lists employed by many schools may suggest). From Table 2 it would appear that this scenario of high elasticities of demand (say 2 to 3.5) coupled with low elasticities of supply (say 0.5) would be consistent with small optimal subsidies of two per cent to 13 per cent. The current



subsidies for many wealthy schools probably lie in this range. However, the benefit cost ratios associated with these optimal subsidies are very low, between 1.01 and 1.05. This implies that the existing range of subsidies to wealthy schools could well be optimal and in fact have a benefit that exceeds their cost but in an environment of limited public funds, these subsidies may not rank sufficiently high in terms of benefit cost ratios to guarantee funding on purely economic and financial grounds. However, it should be remembered that many government policies are not solely driven by economic benefits. These subsidies to wealthy schools may get over the line, in terms of funding, on political grounds or some other non-economic government objective.

Complicating the issue further is the fact that private schools are quite diverse, as are their student bodies, giving rise to a wide range of elasticities even for schools with similar SES scores. For example, some religious schools may differ markedly from other schools with a similar SES. They may have a high supply elasticity where they are very willing to take on new students in order to achieve secondary goals such as the promotion of a particular religion, while demand may be very inelastic because religious education is important to parents who send their children to such schools, implying there are few parents who would decide to send their children to such schools if price was to fall. These claims are relatively consistent with the analysis of non-government schools in Anderson (1992). In terms of the sensitivity analysis in Table 2, this case could imply optimal subsidies in the range of 27 per cent to 78 per cent (assuming the price elasticity of supply is infinity and the price elasticity of demand lies in the range of 0.5 to one). Once again, these subsidy levels are probably broadly consistent with the level of subsidies offered to many poorer religious based schools. Importantly, the benefit cost ratios of 1.12 to 1.44 associated with these subsidies would appear to justify these subsidies on financial grounds.

On the issue of the quality elasticity of public education demand, one could speculate that this would be relatively low if the alternative was an exclusive private school. That is, improvements in public schools would not easily entice private school students back into the public system — the exclusivity may be very important to these parents. These claims are focused purely on the higher and more exclusive end of the private school market and are speculative. Further empirical investigation is required.

The elasticity estimates in Williams (1984 and 1985) apply to pre-1980 data and are based on aggregate data. The diverse nature of the non-government school sector in Australia implies aggregate estimation of education demand relationships will struggle to reflect the true nature of the demand relationships for the various types of non-government schools. Given the great change in public and private schools since the early 1980's (with respect to enrolments and funding), a critical direction for future research on education financing in Australia is the development of estimates of demand and supply elasticities based on recent micro-level data, as in Rubinfeld and Shapiro (1989). Estimates of the quality elasticity of the demand for public education would also be helpful. Taken

together, such data would be useful in determining equilibrium responses to policy changes in public and private education systems.

### **Program Design Problems**

Several features of the SES funding program potentially impinge on its effectiveness as a means of improving the resources of public schools. These include poor targeting of potentially mobile enrolments and the differential rates for identical students at different schools, which is related to the reliance on average school rankings.

An inevitable problem with any program aimed at enhancing choice and mobility is the inability and or unwillingness of policy makers to target the specific people who make decisions on switching their children's enrolments between sectors. This applies to both the ERI and SES systems. Not all parents are willing to consider changing educational sectors. Importantly, there is a large number of parents committed to the private education sector who have the incomes to sustain their educational choice. These educational decisions are not influenced by the subsidy, so any government funding directed at such students under the program is potential wastage in terms of providing choice or facilitating the switching of such students into private schools. It would be difficult if not impossible to ascertain who the marginal public education families are, (those that would switch from public to private schools for a relatively small subsidy), in order to target subsidies exclusively to this group. To the extent that ability to pay is the key demand factor, means testing is likely to be the most practical method of identifying these marginal families and students. If non-government schools continued to charge a uniform fee to all students despite means tested subsidies, a change in the method of delivering subsidies would be required. It is likely that subsidy would have to be paid to the parent and would be offset against the fee charged by the non-government school, thereby enhancing mobility and choice for the marginal students.

The means testing characteristics of the program implied by the use of SES go some way to dealing with the issue of targeting. However, reliance on average school SES scores rather than individual family characteristics is a shortcoming of the program. For example, all students who attend schools with high SES scores are funded at relatively low rates. This means that while parents' choices are enhanced by the subsidy scheme, high SES schools are likely to be out of reach for students and families with low SES. Assuming a willingness and capacity to expand supply on the part of existing high SES schools, a truly means tested scheme would enhance choice further and allow greater numbers of low SES students to enrol in high SES schools. For example, Barse, Glomm and Ravikumar (2000) show that means testing for education subsidies (vouchers) can increase total education expenditures and reduce inequality in the allocation of education expenditures. There is scope however for schools to recognize that enrolling students with low SES will lower the school's average SES and increase the subsidy paid for every student. In order for a high SES school to increase its

government funding, it may be beneficial to offer scholarships to students with low SES. In this sense, the SES funding scheme may offer more mobility and choice than initially thought. It would be interesting and important to study average SES of each year level in a school, as it may be possible for schools to enrol low SES students in early years to attract extra funding and exclude them in the final year (assuming they are poor performers) in order to boost university entrance outcomes. This would not truly reflect an enhancement of choice and may suggest opportunistic behaviour on the part non-government schools; see Anderson (1992) for a discussion of such policies.

Differential funding between schools under both the SES and ERI systems generates stronger incentives for parents of students in government schools to shift their children to low SES (and high subsidy) schools rather than high SES (and low subsidy) schools. If the major factor constraining parents to public schools is a low ability to pay for private schooling, the outcome of this incentive structure could be to reinforce the segmentation of schools on the basis of SES, with relatively low SES parents in the public sector shifting their children to relatively low SES schools in the private sector. There appear to be no economic grounds for the government to bias subsidies and enrolments in this way. It could be argued that this is likely to have serious social implications, stratifying students by SES, operating to strengthen the exclusivity of high SES schools, to embed a class structure and potentially to reduce social and earnings mobility. To the extent to which high SES schools are also high ERI schools the impact of the two programs, in this regard at least, seem equivalent.

### **Policy Implications**

The use of subsidies of private sector activities is both widely practiced and highly controversial. These programs can be cost effective in terms of saving public funds or providing greater resources for public sector activities. Their success in terms of budgetary effectiveness depends on the nature of the market environment in which they operate and the specific policy parameters.

A number of problems with the existing policy have been identified. An important issue is the lack of a strong link between the provision of subsidies and the potential willingness of students to shift between sectors. This poor targeting implies that part of private education subsidies are going to private school users that would never switch to the public system, implying unnecessary subsidies and, if the objective of subsidies is to enhance choice or improve the government school system, a wastage of public funds.

Simple solutions to the targeting issue are not apparent, as the identification of potentially mobile groups of students is problematic. If one accepts that lower income families have a greater unmet demand for private education than higher income families, the payment of subsidies to individuals on the basis of means-testing of student's families (essentially a means tested voucher system) rather than assessing the average SES of schools may be one way to improve the targeting of subsidies.

An important point to remember about the SES system of government funding to non-government education outside the Catholic system is that this subsidy scheme resembles a voucher scheme, because funding is provided on a per student basis and if a student switches from one private school to another, the old school loses the subsidy that the switching student entitled them to and the new school gains a subsidy for enrolling the student. This is a voucher in character if not in name.

The interpretation of the SES scheme as a means tested voucher is not as clear-cut. Means testing is typically understood as granting aid based on need, with people on low incomes receiving more aid than those on higher incomes. In the SES funding scheme, the school's average SES measures the means of an individual. A student with low SES (low income and occupation score) will receive a higher subsidy in as much as they may lower the average SES of the school they attend, thereby entitling themselves (along with the whole student body) to a higher per student subsidy or voucher. Clearly this is not what we expect from means testing. It makes it very hard for a family with low SES to have their child attend an expensive private school because the family's voucher will be the same as the voucher given to a family with the highest SES sending their child to the same school. This does not appear consistent with the governments stated objective of increasing choices available to parents although schools have the discretion to reduce fees for low SES students, for example by offering scholarships.

We conclude by noting that if market based approaches, which include subsidies, are to be successfully implemented, it is important that policymakers have a clear understanding of the demand and supply relationships for both private and public education. This requires improved estimates of demand and supply elasticities for private education. This would be a fruitful direction for further empirical research and would only improve education funding policies. The co-existence of the SES and ERI funding schemes arising from the Catholic school system opting out of the SES scheme will provide the opportunity to analyse parental responses to the different subsidy schemes and may assist in the estimation of the elasticities of interest.

Given our assumptions about policy objectives, the SES framework clearly exhibits shortcomings. Common with the ERI, the poor targeting of existing subsidies, through the use of average school rankings rather than individual means tests, makes for relatively ineffective policy. While the existing subsidy rates to private schools are of the broad order of magnitude that would be expected given the responsiveness of parents, on average, to changes in fees and the impact of higher enrolments of fees, the likely benefit cost ratios from subsidizing schools with high SES scores implies that the relatively small subsidies that are given to these schools are difficult to justify on financial grounds.

Some may argue that the various shortcomings of the SES funding scheme create a case against public funding of non-government schools. It should be emphasized that the scrapping of government funding for private schools could have serious financial implications for the whole of society. Not only will

families with students in non-government schools be adversely affected, public school students may suffer from over crowding and taxes would have to be higher in order to fund the greater public expenditures.

Alternatively, an individually means tested voucher system would improve targeting and enhance parental choice, providing more opportunities for students with low SES to attend non-government schools with high SES and fees.

## **References**

- Anderson, D. (1992), 'The Interaction of Public and Private School Systems', *Australian Journal of Education* 36(3):213-36.
- Bearse, P., G. Glomm and B. Ravikumar (2000), 'On the Political Economy of Means-Tested Education Vouchers', *European Economic Review* 44(4-6):904-915.
- Burke, G. (1996), 'School Resources', *Australian Economic Review* 116:401-8.
- Commonwealth of Australia (2001), *Gazette*, No. 1, 19 November.
- Commonwealth Schools Commission and Australian Bureau of Statistics (1979), *Australian Students and Their Schools*, Canberra.
- Department of Education, Training and Youth Affairs (2000), 'Commonwealth Programs for Schools: Quadrennial Administrative Guidelines 2001 to 2004', Canberra.
- Department of Employment, Education, Training and Youth Affairs (1997), 'Schools Funding: Consultation Report', Commonwealth of Australia, Canberra.
- Department of Employment, Education, Training and Youth Affairs (1998), 'Schools Funding: SES Simulation Project Report', Commonwealth of Australia, Canberra.
- Donnelly, K. (2001), 'Education Agenda: Beaujolais Bolshies', *Institute of Public Affairs Review*, 53(1):28.
- Frey, D. (1991), 'Optimal-Sized Tuition Tax Credits Reconsidered: Comment', *Public Finance Quarterly* 19(3):347-354.
- Gaviria, A. and S. Raphael (2001), 'School-Based Peer Effects and Juvenile Behaviour', *The Review of Economics and Statistics* 83(2):257-268.
- Kemp, D. (2001), 'Federal Budget delivers for Government Schools', Media Release, Department of Education, Training and Youth Affairs, 22 May.
- Martinello, F. and E. West (1988), 'The Optimal Size of the Tuition Tax Credit', *Public Finance Quarterly* 16(4):355-368.
- Martinello, F. and G. West (1991), 'Education Budget Reductions Via Tax Credits: Some Further Considerations', *Public Finance Quarterly* 19(3):355-368.
- Maslen, G. (2000), '“Give to rich” plan sparks outrage', *The Times Educational Supplement*, November 24:14.

Ministerial Council on Education, Employment, Training and Youth Affairs (2003), 'National Report on Schooling in Australia 2000', <http://www.curriculum.edu.au/anr/index.html>

Rubinfeld, D. and P. Shapiro (1989), 'Micro-Estimation of the Demand for Schooling: Evidence from Michigan and Massachusetts', *Regional Science and Urban Economics* 19(3):381-98.

Steering Committee for the Review of Commonwealth/State Service Provision (2002), *Report on Government Services 2002*, AusInfo, Canberra.

Sydney Morning Herald (2002), 'Spelling out the education split', 27 May 2002.

Williams, R. (1984), 'Interactions Between Government and Private Outlays: Education in Australia, 1949-50 to 1981-82', *Economic Record* 60(171):317-25.

Williams, R. (1985), 'The Economic Determinants of Private Schooling in Australia', *Economic Record* 61(174):622-28.

Zimmer, R. and E. Toma (2000), 'Peer Effects in Private and Public Schools across Countries', *Journal of Policy Analysis and Management* 19(1):75-92.

*The authors would like to thank two anonymous referees and Graeme Wells for helpful comments.*

# Buyer Power of Supermarkets

Gordon Mills

Over recent decades, the supermarket sectors in many developed economies have become much more concentrated. Nowadays, the largest chains are said to possess and exercise considerable 'buyer power'. In Australia and Britain, governments have established codes of buying practice in response to supplier complaints about supermarket conduct. For those two countries, it is timely to report on recent government initiatives, and to offer some analysis.

This paper examines the market contexts, describes and evaluates some widespread buying practices, considers the difficulties in applying general competition law, describes the provisions of the codes of practice, reports on the early and (seemingly) unimpressive experience with the codes, and suggests ways in which the governments might strengthen them.

To define *buyer power*, it seems best to adopt a direct, operational measure: a firm (or group of firms) has buyer power when, *for like transactions*, it can obtain from a supplier terms that are more favourable (for the buyer) than those available to other buyers. Besides prices, 'terms' refers to contractual conditions that relate to payment, product quality, labelling and packaging, and any associated marketing and promotional activities.

## The Market Contexts

In considering whether there is such buyer power, one issue is the proportion of the total supply of a particular product that goes to the largest supermarket chains. Accordingly, it is wise to look at the retail market comprising all 'grocery' stores, regardless of store size.

Using a comprehensive grocery-store category, the Competition Commission (2000:Table 5.6) reports that in 1998-99 the UK's five largest firms had some 69 per cent of the national market measured by turnover. Concentration is much higher in the market for 'one-stop' shopping, a market that is served only by the largest stores. For stores over 2,300 square metres in floor area, the Commission notes that those five firms had 95 per cent of that market.

The all-stores sector in Australia is more concentrated than that in the UK. For 1997-98, the Australian Bureau of Statistics (Joint Select Committee, 1999:Appendix 5) put the national share of the *three* largest chains at 62 per cent, even for a widely defined market comprising all supermarket and grocery items, together with liquor retailing.

Since then, the smallest of those chains (Franklins) has been broken up. Woolworths bought many of the one-stop stores and some of the others (Woolworths Limited, 2001). The *two* largest companies, Woolworths and Coles,

---

*Gordon Mills is Honorary Professor of Economics, University of Sydney.*

now seem to have about 58 per cent of the wide market and probably have more than 90 per cent of the Australian one-stop market. Furthermore, most 'independent' supermarkets buy through wholesalers, which have more buyer power than do individual stores.

At the firm level (in both countries), some important economies of scale and scope (especially those found in distribution) favour the large chains, regardless of whether the manufacturer or the chain undertakes the distribution. In recent years, some new companies have entered (notably, in the UK, the 'discounters' Aldi, Netto and Lidl; and in Australia, Aldi). Yet, they have had little impact on the market shares of the big chains. No entrant operates large, one-stop stores — for which the barriers to entry include planning/zoning restrictions, and difficulties in accessing suitable existing sites.

A large supermarket may sell 20,000 or even 30,000 products, which (in the UK) come from about 2,000 suppliers. In contrast, even a large manufacturer often sells most of its output to just a few of the largest supermarket chains. A smaller supplier may sell all its output of a particular product to a single chain (so-called 'solus supply') (Competition Commission, 2000:paragraphs 11.5 and 11.10).

Besides this numerical distinction, there is also a spatial difference. A single store sells its many products in a local area, where there may be little or no competition (between one-stop shops, at least, and particularly in Australia). On the other hand, the owning chain often buys its supplies on a national basis, and may solicit supplies from rival companies (Dobson Consulting, 1999:section 4.3).

Nearly all processed products sold in supermarkets are *branded*. One intention is to assure consumers of a consistent quality that is maintained over successive purchases. One effect is to enhance the market power of the brand-owner, who is usually the supplier. However, supermarket chains commonly develop their own brands. This enhances the market power of the chain, rather than that of the supplier who makes the 'own-label' product, and whose identity is not known to the consumers. Own-label strategies are considered below.

### **Buyer Power, Market Concentration and Bargaining with Suppliers**

When a few supermarket chains account for a large share of all grocery sales, then those chains together generally buy a large proportion of total output in the narrow product range of a specific intermediate market. On the other hand, the supply side (for processed products, at least) can be even more concentrated.

Even so, a supermarket company has an important bargaining chip, namely the threat to 'de-list' one or more products — that is, stop selling those products. This is likely to reduce the profits of both parties. However, a large supermarket company may buy (say) 20 per cent of the manufacturer's output while the relevant products may contribute only (say) one per cent of the buyer's turnover. Thus the *direct* proportional effect of de-listing on the buyer's profits is likely to be much less than the direct effect experienced by the manufacturer. (It is the facility to switch suppliers that gives rise to buyer power.)



Of course, there can be wider effects: if a chain delists a branded product, will customers seek the product at a rival store, and transfer other custom as well? As recognised long ago (Heflebower, 1957:279):

For many commodities consumers have no strong preference for one brand over several others ... The manufacturer of such a line is a beggar at the retailer's office.

This argument becomes stronger as retailing becomes more concentrated. There may be only a single *one-stop* store in a given local area, and a shopper who transfers to another store may incur extra travel and time costs. Provided the first store stocks a rival brand that shoppers regard as a reasonable substitute, it seems that few would transfer.

Even so, in a few cases (Kellogg's breakfast cereals is an oft-quoted example), a product may have such a strong brand reputation that consumers expect to find it in any supermarket. Then, the chain may fear that if it de-lists, a significant proportion of its regular shoppers *would* go elsewhere, perhaps leading to it losing sales of many other products. Because the brand-owner knows that too, the threat to de-list may not be credible. However, the retailer may have another weapon available, namely a threat to de-list any of the supplier's other products that have lesser brand strength (Competition Commission, 2000:paragraph. 2.454).

A large supermarket chain can employ and enhance its buyer power by introducing own-label brands. Such brands account for about 35 per cent of supermarket sales in Britain (Competition Commission, 2000:Table 6.1), but are less important in Australia. Generally, the supermarket specifies the product qualities (Burt, 2000), and can choose where to position its brand on the quality spectrum. By merely placing such products on its shelves, it can sometimes secure large sales, without incurring advertising expenses. Armed with that potential, the supermarket can negotiate low prices in sourcing supplies — perhaps from a manufacturer of a well-established brand that may be willing to supply the own-brand product for fear that otherwise the chain will obtain supply from another manufacturer, who may not have an established brand.

On its own brand, the retailer can make a good profit — in proportionate terms, at least — even when it sells at a price that is lower than that of the well-established manufacturer brand (Competition Commission, 2000:paragraphs 8.141-8 and 2.413). This applies especially when the latter price is elevated in order to fund expensive brand-name advertising. Dobson Consulting (1999:Part I) provides further discussion of conceptual aspects of supermarket buyer power.

### **Outcomes for Wholesale Prices**

Whenever there are few players on *each* side of a wholesale market, price outcomes depend on bargaining, as well as on the underlying cost and demand considerations. If trade of a particular quantity is to be feasible, the highest price

the buyer is willing to pay for that quantity must exceed the lowest price that is acceptable to the seller. Given the concentration in supermarket retailing and in (some) manufacturing, these two price bounds may diverge significantly. The bargaining serves to allocate profit shares.

In selling to small customers, the supplier may obtain a price that gives a rate of profit in excess of a normal (competitive) return. In that regard, the seller may be said to exercise market power. Then, the lowest price the seller will accept when dealing with a large chain may be insufficient to yield a normal profit *on those sales*. In such a case, it may be that most of the supplier's profits arise from sales to its smaller customers.

For the UK, some Competition Commission (2000) data, given here in Table 1, show that for some major branded goods small retailers pay large suppliers as much as 12 per cent more than the prices paid by the largest companies. Even between the *large* chains, the proportionate difference can be as much as three per cent. The Commission believes (paragraph 11.116) these differences go beyond any differences in operating costs; it concludes that the largest multiples enjoy notable buyer power even when dealing with large suppliers.

Two earlier government reports (Monopolies and Mergers Commission, 1981; and Office of Fair Trading, 1985) also found such price discrimination, but thought that the then competition between the larger chains sufficed to ensure that they passed on to their customers much, if not all, of the financial benefits. In its recent report, the Competition Commission concluded that the UK market is 'broadly competitive', but did not consider explicitly whether the large retailers still passed on such savings. As retailing concentration increases, the competitive pressure to do so will diminish. (In Australia, concentration may have now reached a critical level in that respect.)

Until recently, the Australian debate on buyer power has rested on anecdotal evidence. The Australian Competition and Consumer Commission (1999:12) notes:

Retailers are said to earn greater revenue from discounts and allowances than they do from [the standard] margins on product sales. Thus manufacturers experience considerable pressure from retailers for greater discounts and allowances.

Notwithstanding some contrary large-supplier claims, the Joint Select Committee (1999:paragraphs 5.52 and 5.54) concluded that the major chains use their buyer power to extract the best deal possible from suppliers. In addition, the Committee (paragraph 5.57) quoted the evidence of a former independent retailer:

The trading terms are remarkably different. We are dealing in an industry that works on two or three per cent profit and yet the difference in trading terms is often ... as much as five, six or eight per cent.

**Table 1: Prices Paid by Large UK Supermarket Groups to Large Suppliers of Certain Major Branded Products, 1998-99**

<i>Supermarket group</i>	<i>Market share (%)<sup>a</sup></i>	<i>No. of identified lines purchased<sup>b</sup></i>	<i>Average price paid<sup>c</sup>, as % of average price paid by Tesco</i>
Tesco	23.0	126	100
Sainsbury	18.7	122	101.6
Asda	12.2	124	102.3
Somerfield	9.8	125	103.0
Safeway	11.5	121	103.1
Morrison	3.9	118	104.6
five other multiples <sup>d</sup>	7.6	66-103	105-111
other retailers <sup>d</sup>	-	about 100	about 107-112

Notes: a Here, the market refers to the turnover of those multiples identified in Table 5.6. This is the widest of the alternative definitions considered by the Commission, and gives the lowest market shares for the large multiples.

b A 'line' is a specific branded product in a specified package size. The data come from returns from 26 large suppliers, and relate to the top five lines of each supplier. For these 130 lines, this column shows the number of lines purchased by each retailer.

c The prices averaged by the Competition Commission are here taken to be the invoice prices, net of discounts (paragraph 11.114). The averaging uses equal weights for each *included* line. Taking Sainsbury as an example, the average price index refers to the 122 lines that are purchased by Sainsbury *as well as* by Tesco.

d These groups are identified in Tables 11.10 and 11.11, respectively.

Source: Competition Commission (2000). The price data (taken from Tables 11.10 and 11.11) was provided by a panel of large suppliers. The market share information is from Table 5.6.

Following a recent attempt to collect systematic data, the ACCC's report (ACCC, 2002:Chapter 5) identifies the standard discounts and the 'funding support' provided by suppliers (ostensibly for in-store promotions and the like) that have to be deducted to arrive at the so-called 'net net' prices. It explains also that in selling to a wholesaler (as distinct from a large retail chain), the supplier

may give some funding support directly to selected individual retailers, rather than to the wholesaler.

Because of these complexities, and especially because of the low response rates to its (voluntary) survey, the ACCC urges caution in interpreting the data. For what it is worth, its report makes it clear that: (1) the survey did not find a striking correlation between size of supermarket chain and buying prices, comparable to that found by the Competition Commission in the UK; (2) suppliers did agree different net prices with different buyers; (3) 'independent' wholesalers sometimes got better terms than the two big retail chains; and (4) the two big chains got the best prices in about 65 per cent of the reported cases. As to the size of the price differences, Charts 5.63 and 5.64 of ACCC (2002) show that in 62 per cent of the cases there was a price difference that exceeded two per cent, and in 36 per cent of the cases a difference that exceeded four per cent.

Thus the extent of price dispersion in Australia is probably much the same as in the UK. In both countries, the pricing deals are complex, and it is unlikely ever to be possible to obtain precise data. It is, however, easier to obtain *qualitative* evidence on the practices of the large buyers. This is considered next.

### **Pricing and Related Practices in Wholesale Markets**

The buying practices of large supermarket chains may be considered under two general headings: prices and other terms that relate to the supply of the goods themselves (considered in this section); and other practices, including those relating to the services that are sometimes rendered by retailers to promote the sale of selected products (examined in the next section). This section looks first at pricing, and then at the other terms. It offers a *qualitative* evaluation of the economic consequences. Though some of the practices seem relatively unimportant in principle, the distributive and efficiency consequences (of any widespread use) are likely to be significant in aggregate.

It is helpful to begin by looking at contractual terms in general. The Competition Commission (2000:paragraph 2.469) notes that 'in the dynamic trading environment of [UK] grocery retailing ... full written contracts are rare'. In particular, it is not common for the written contract, if any, to specify the quantities to be supplied. The retailer then chooses and varies the quantities at short notice. However, a chain usually does promulgate standard conditions for delivery arrangements and similar matters.

Much the same general approach obtains in Australia (Joint Select Committee, 1999:section 5.104). That Committee's anecdotal evidence on other aspects of supermarket buying practices suggests that these practices are much the same as in the UK — for which the Competition Commission provides a systematic account that is relied upon here.

### *Unit prices*

When large retailers buy at unit prices that are lower than those available to small retailers, the resulting price differentials may or may not match differences in the unit costs of supply. As noted already, the Competition Commission did find that in the UK, large retailers enjoy a price advantage that exceeds the cost difference.

This price discrimination may facilitate the further growth of large firms at the expense of small retailers — an outcome that does not (necessarily) stem from differences in retailer (cost) efficiency. (Even so, the major influence on past growth must surely have been the attraction to car-owning consumers of shopping in a large, one-stop store.)

### *Temporary discounting of retail prices*

Commonly, a supermarket offers temporary price reductions ('specials') on some of the product lines on its shelves. To support such a reduction, the supplier *may* be willing to accept a reduced wholesale price. Where the parties agree on this in advance, and the supplier does not feel coerced, there can be no complaint. Yet, the Commission found that some UK chains imposed *retrospective* reductions in the agreed wholesale prices. This may disadvantage some suppliers relative to others.

### *Unilateral changes to prices and to related contractual terms*

The Competition Commission (2000) found that several of the larger UK supermarkets employed such practices. Three of these practices are noted here:

- Retrospectively seeking or imposing discounts (on previously agreed wholesale prices), especially when the retailer has surpassed a sales target (paragraph 2.495).
- Requiring or requesting compensation from a supplier when the retailer's profit on a product turns out to be less than it had expected (paragraph 2.496).
- Requiring or requesting a supplier to buy back unsold items (or simply failing to pay for them), when there was no written 'sale or return' agreement (paragraph 2.499).

Obviously, all such practices suggest the presence of buying power, which is used to secure a better outcome for the retailer at the expense of the supplier. However, some other consequences are not so obvious, and need to be argued.

After concluding (paragraph 2.458) that the largest supermarket retailers have significant buyer power, the Commission finds that each of the above practices, when carried out by such retailers, 'adversely affects the competitiveness of some of their suppliers, particularly small suppliers, and distorts competition in the supply of groceries between retailers, and between suppliers' (see paragraph 2.494(c), and similar wording elsewhere). According to the Commission such practices distort competition in the *retail* supply of groceries because they

disadvantage those retailers that do not receive the retrospective advantages, and distort competition between *suppliers* because an undue element of commercial risk, that would otherwise rest with the retailer, is transferred to the supplier.

Although many of the practices do indeed transfer some risk from retailer to supplier, the first of the three practices does not necessarily have that effect: when sales go poorly, the supplier receives a higher unit price, and this dampens the revenue variation.

When risk *is* transferred, the supplier gets a poorer deal when selling to a large, powerful retailer. Though this transfer could threaten the viability of a supplier, this seems unlikely for a large manufacturer, as distinct from a small grower.

### *Some related practices*

Without going into a detailed evaluation here, it is of interest to note several other practices that the Competition Commission (2000) found in wholesale markets in the UK:

- Delaying payments to suppliers beyond the terms in the contracts (p. 122).
- Changing quantities or product-quality specifications at less than three days' notice, and without paying compensation to suppliers (p. 123).
- Requiring suppliers to make, and pay for, changes to packaging and labelling (p. 132).
- Requiring suppliers to purchase goods or services from designated companies (p. 134).

It is hard to see how retail chains could engage in these practices without buyer power.

## **Other Practices in Supermarket Dealing with Suppliers**

These other practices can be very varied. As well as direct payments to retailers, they include supermarket pressure on a supplier to participate in in-store promotion efforts and attempts to influence the terms on which the supplier sells to other retailers. Such practices can have adverse effects on suppliers and on competition between retailers.

### *Listing fees*

A listing fee is a supplier payment that a supermarket operator may demand before agreeing to stock the product. Consider the case where the supermarket does not give any undertaking beyond agreeing to stock the product — in particular, the amount and nature of the assigned shelf-space is at the supermarket's discretion.

The UK Competition Commission (2000:paragraph 2.476) found that the larger supermarkets often charge listing fees. Sometimes the fee is required only when the supermarket first stocks a product. On other occasions, the supermarket

may demand payment to continue the ‘listing’ of a product already on the shelf. On the evidence quoted by the Commission it seems that smaller retailers do not receive such payments. Although the Commission does not say so, the presumption is that the fee is a lump-sum payment; otherwise, the transaction merely has the effect of a price discount (Tom, 1999).

The Commission finds that when retailers with buyer power seek such fees (paragraph 2.476(c)), the practice

adversely affects the competitiveness of some of their suppliers and distorts competition in the supply of groceries between suppliers, because the multiple engaging in the practice does not necessarily select the best, or most efficiently produced, product, or that preferred by consumers, but to some extent is influenced by whoever is best able to make the payment requested.

The Commission does not present any supporting or clarifying remarks. Accordingly, the claim needs examination here.

Among competing suppliers, the one ‘best able to make the payment requested’ is the company that makes the most profit — perhaps, profit per unit sold — from supply of the line. In turn, provided the retail market is reasonably competitive, this profitability reflects low costs of production (relative to product quality), or product quality characteristics, and prices, that attract consumers by providing ‘good value for money’. On these considerations, it is not clear that the Commission has established its case.

On the other hand, where retailing is not very competitive, a supermarket may be able to trade very successfully by selecting suppliers and products largely on the basis of its own profit margins, and with less regard for consumer preferences — especially when competing brands have similar qualities. Is this the situation the Commission has in mind?

There may be a further factor at work: a retailer may incur significant costs in evaluating and introducing a new product. In that case, there is an efficiency argument for seeking re-imbursement from the supplier. One effect is to transfer to the supplier more of the risk that is attached to launching a new product (Tom, 1999). Of course, this does not justify a requirement for payment of recurrent listing fees. While the cost rationale may be important, the amount of the listing fee may well exceed the amount of the one-off costs. When so, buyer power must be part of the explanation.

Despite the additional cost of seeking a listing fee, a retailer with buyer power may find the device attractive because optimal use of an extra pricing instrument can never be less profitable (and, generally, will be more profitable) than optimal use of the set of pricing instruments that is not enlarged in this way. In particular, payment of a listing fee may substitute for part of the price discount offered to large retailers. This ‘disguise’ may make it easier in practice for the large retailer to enjoy preferential terms.

In summary, the charging of listing fees rests, at least in part, on buyer power. It may serve to give large retailers a larger share, and the supplier a smaller share, of the aggregate profits that arise from serving the retail market. The supermarket's choice among rival brands may depend on supplier willingness to pay such fees. Yet, where there is little product heterogeneity between suppliers' lines, the ensuing damage to the consumers' interest may be small.

*Listing fees that buy favoured supermarket promotion*

The UK Competition Commission found that supermarket chains with buyer power made much use of several such practices. In particular, in exchange for a listing fee, the retailer awards favourable shelf location, or more shelf space, or space in more of its stores, or provides advertising or other promotional support. Here, there is especial concern if the payment exceeds the amount of the extra costs incurred by the retailer.

In the Commission's standard mantra (see, for example, paragraph 2.477(c)), the practices adversely affect 'the competitiveness of some of [the] suppliers and distorts competition in the supply of groceries between suppliers'. To explain this, the Commission again goes no further than to say that the supermarket may be induced to buy from the supplier 'best able to make the payment requested' (paragraph 2.476(c)), without (sufficient) regard for consumer preferences, product quality and productive efficiency.

Again, it is important to look at specific possible effects. When providing more shelf-space to the supplier who pays a listing fee, there is (in the short-term, at least) less space available for rival brands. If, however, the fee paid goes no further than matching the direct costs incurred by the supermarket, the practice is efficient, and there need be no adverse effect on competition.

However, an existing supermarket site may be so small as to preclude expansion of the store, and often it is very difficult to find another site. The store's 'direct cost' could then include a scarcity rent for shelf space whose supply is constrained even in the long run. This could result in the crowding-out of another brand for which a listing fee is not paid. Such exclusion could seriously reduce a supplier's profits and might even threaten the viability of weaker brands.

*Imposing conditions on a supplier's sales to other retailers*

Retailers who sell products at discounted prices sometimes allege that a large supermarket company, anxious to preserve its gross margin, tells a product's supplier that the chain will stop selling that product unless the supplier stamps out the discounting. Such pressure may induce the supplier to put pressure on the discounter, even perhaps to the extent of threatening to withhold supply.

In the UK, the Competition Commission noted one such report, but did not reach an adverse finding (paragraph 2.488). In Australia, on the other hand, the ACCC has instituted several legal proceedings alleging price fixing or resale price maintenance.



One such case concerned bread. In 1996, the ACCC alleged that a Woolworths company (trading as Safeway) had refused to accept further supplies of bread from George Weston (trading as Tip Top Bakeries) to discourage that supplier from allowing discounted pricing of its bread by small retailers at several places in and around Melbourne. In turn (as the ACCC alleged), the supplier had stopped supplies to the discounters. In May 1997, the Federal Court fined Weston \$1.25m 'for fixing the price of bread and attempted resale price maintenance'. Yet, in 2001, a court found that Woolworths had not breached the Trade Practices Act in respect of its role in these events. In 2002, the ACCC lodged an appeal against that decision. (See ACCC Media Releases 178/96, 54/97, 329/01 and 28/02.)

In another case, a court imposed a penalty of \$900,000 on Weston for attempting to promote a price-fixing agreement between competing retailers in Tasmania. The retailers were Chickenfeed Bargain Stores and a Woolworths company, trading as Purity.

The conduct in question occurred when George Weston sought to intervene in [a biscuit] price war in an attempt to minimise the amount of a claim by Woolworths to compensate it for meeting Chickenfeed's pricing. (ACCC Media Release 107/00)

Also in Tasmania, the same retailers figured in a (broadly) similar case, resulting in the Federal Court imposing penalties of \$500,000 on Colgate-Palmolive for trying to stop the advertising of discounted retail prices for toothpaste and other products that Colgate-Palmolive supplied (ACCC Media Release 119/02).

## **General Competition Law and Buyer Practices**

Both the Australian and British governments have introduced formal codes of buyer conduct, and thereby have revealed beliefs that the general legislation (the UK *Competition Act 1998* and the Australian *Trade Practices Act 1974*) is not likely to be sufficiently effective in this field. This section reviews very briefly the difficulties in applying the general legislation to supermarket buying.

While the prevailing competition law may discourage some abuses of buyer power, any action under that law requires the competition agency to investigate formally, and may require enforcement action through the courts. This can delay a decision, especially if there is an appeal. Meanwhile, subject to any interim measures that are put in place, the retailer's conduct can continue.

As presently written, the competition laws can be applied only when the buyer power is used in a way that amounts to the abuse of a dominant position (in Britain) or the misuse of substantial market power (in Australia). Even the largest retailer may not have the market share usually thought necessary for it to be found to be dominant or to have substantial market power, though this is not so likely in Australia.

Further, even if dominance or substantial market power is established, the conduct can not be prohibited unless it is found to be an abuse of the retailer's market position. Buyer power is exercised in a negotiating context, and it can be difficult to show that the retailer's practices amount to such an abuse. It is not enough that the deployment of buyer power damages the individual supplier. Rather, the conduct must adversely affect consumers — or, under the Trade Practices Act, it must have an anti-competitive purpose. The effects of buyer power on the competitive process are not necessarily adverse — sometimes efficiency is enhanced, and consumers may benefit from lower prices.

Even when a legal breach is established, it can be difficult to devise an effective remedy. Although a practice can be prohibited, the context allows the retailer to find another way of exercising its buyer power (this hampers the enforcement of buyer codes, too, as is shown later). Further, any remedy applies only in the particular case. In contrast, a code seeks to control buyer power *ex ante*, and is applicable to all the retailers in the defined sector.

Finally, actions under competition law result in considerable costs and uncertainties for the parties. For this reason, suppliers may be particularly reluctant to complain, and to provide the authorities with the documentary evidence needed to pursue a case, especially if the supplier fears retaliatory action by a large retailer to which it still wishes to sell. Again, this difficulty arises in implementing a code of practice, too.

It is no surprise, therefore, that there have been very few cases (in any country) where buyer power has been found to be an abuse, and it is hard to see how it can be otherwise. Merger policy has been used to limit the increase in retail concentration and therefore buyer power, but of course that does not control existing buyer power.

### **The Codes of Buying Practice**

In Australia, the Parliamentary report (Joint Select Committee, 1999:Recommendations 3 and 5) proposed an independent Retail Industry Ombudsman, to be appointed and funded by the government, and backed by a mandatory Retail Industry Code of conduct. In its response (17 December 1999), the Government favoured a voluntary code of conduct to foster self-regulation.

In the outcome, a (voluntary) *Retail Grocery Industry Code of Conduct* came into effect in September 2000. Its objectives (Code, section 2) are to promote fair trading, encourage transparency in dealings between supermarkets and their suppliers, and offer a simple dispute resolution mechanism, under which disputes could be referred to the new Retail Grocery Industry Ombudsman (section 10). Although the Committee had said the code should enshrine the principle of 'like terms for like customers', the Government did not agree and the matter is not considered in the code. Both Woolworths and Coles have 'subscribed' to the code.

In Britain, the Competition Commission (2000:paragraphs 2.588-90) concluded that, to stem the adverse effects of certain buying practices, the

Government should establish a Code of Practice and require each supermarket chain having more than eight per cent of a (defined) 'groceries' market to give a legally binding undertaking to comply with it (the Commission considered such a market share was sufficient to give a retailer effective buyer power). Under the code, a supermarket would put its contractual terms in writing. The code would provide for dispute resolution, with ultimate recourse to an independent third party who would report any breach to the Office of Fair Trading.

The Government accepted the proposal, and the Director-General of Fair Trading (DGFT) drafted a code, after consulting the parties. The government endorsed the *Code of Practice on Supermarkets' Dealings with Suppliers* in October 2001. In December 2001, the four companies (Asda, Safeway, Sainsbury and Tesco) that each had more than eight per cent of the market signed undertakings.

### *Details of the codes*

Though the two codes have similar provenance and substantially the same overall objectives, there are striking differences in implementation.

While both codes relate to 'groceries', the Australian code does not define the term. Instead it offers (in section 4) a precise definition of 'produce' (essentially, fruit and vegetables, meat and fish) and a very broad definition of 'product' ('made by a process of industrial transformation ... [and] sold in the [retail grocery] industry'). In contrast, the UK code defines groceries (clause 32) as 'products sold from any retail store in the UK' including food and certain other product groups, but excluding many categories such as pharmaceuticals and cosmetics, clothing, DIY products, electrical and electronic appliances and kitchen hardware.

The UK code's terms apply equally to all included products. In contrast, the Australian code has a major section on standards and specifications for 'produce' (section 5). No doubt this reflects the political concerns that gave rise to this code.

In both codes, the detailed wording tacitly recognises a reality in which supermarkets initiate contract terms and make post-contract determinations, while suppliers respond to those initiatives. The codes' ostensible purpose is to constrain such manifestations of buyer power. Moreover, the UK code obliges the supermarkets (that have given the undertakings) to train their buyers on the code's requirements.

Clause 1 of the UK code requires the supermarket to put its business terms in writing, if the supplier so asks. Australian retailers and suppliers retain 'the right to determine whether or not [a] contract is evidenced in writing' (section 6). The latter position may reflect the Australian emphasis on 'produce'. Even so, if the supermarket were to put its *standard* terms in writing, the remaining writing task (for each contract) concerns only a few parameters such as price, quantity and delivery date, and would require only a modest effort, which need not delay or otherwise impede the trading process.

The Australian code does not refer to any specific buyer practices. This

may reflect the 1998 strengthening of the Trade Practices Act (section 51AC), which deals with ‘unconscionable conduct’. The section refers particularly to situations in which the two parties have differing degrees of bargaining power and might help in dealing with supermarket buying conduct. Yet, early experience does not support such optimism.

In contrast, the UK code frowns on a lengthy list of ‘undesirable’ practices, such as demands for suppliers to reimburse certain kinds of supermarket costs, and supermarket-initiated changes in terms (even of quantities to be supplied) that are made without giving reasonable notice to the supplier. However, the list does not include retailer pressure for discriminatory pricing that benefits large chains at the expense of small retailers.

Further, the use of many qualifiers inevitably creates ambiguity and uncertainty, and may thereby reduce the code’s effectiveness. Examples include (all italics added):

- ‘a supermarket shall not, directly or indirectly, *unreasonably* require a supplier *predominantly* to fund the costs of a promotion’ (clause 14);
- a supermarket shall pay ‘within a *reasonable* time, after the date [of the invoice]’ (clause 3);
- promotions excepted, new-product lump-sum payments must reflect ‘a *reasonable* estimate by [the] supermarket of the risk [it runs] in stocking [the product]’ (clause 9).

Though the code gives formal definition to some of the qualifiers, the criticism still stands.

On the other hand, the requirements for prior notice or agreement may have some effect in reducing the incidence of arbitrary post-contract conduct by a supermarket. As noted later, this may deliver efficiency as well as distributive consequences.

Of course, code efficacy depends also on enforcement. When a supermarket and a supplier cannot themselves resolve a dispute, both codes provide for independent third-party mediation — to be carried out by the Retail Grocery Industry Ombudsman (in Australia), and by a mediator appointed by the supermarket (in the UK).

Neither of these third parties has power to impose a resolution. Both are required to report to government. In Australia, the Ombudsman has to produce an annual report and otherwise has to report to the Minister for Small Business on the operation and effectiveness of the code. In the UK, when a mediator is unable to resolve a case, the supermarket must give notice to the DGFT, who can ask the mediator for the details.

Though the direct compliance provisions are very weak, the reporting requirements might serve to inhibit buyer conduct, if suppliers prove to be willing to lodge complaints. In particular, a supermarket might wish to avoid getting a poor reputation. Further, in a concentrated supermarket sector, each large retailer

might see significant private benefit in not adding to the total number of reported, unresolved complaints, lest the government again turn its attention to the sector.

#### *Experience with the codes*

It seems unlikely that either code has had much, if any effect, so far. In Australia, the pertinent Minister reported (in Office of Small Business, 2002:30) that 'there were no formal applications for dispute resolution under the Retail Grocery Industry Ombudsman Scheme in 2001'. It seems there have been no such cases in 2002. For the UK, *The Times* (18 February 2003) reported that no supplier had made use of the dispute resolution procedures and the OFT had not received any direct complaints.

On the other hand, the press continues to report (anonymous) complaints from farmers and others about the buying practices of some retailers. *The Grocer*, recognising that suppliers may be unwilling to take the risk of initiating mediation proceedings, contacted a number of suppliers and asked whether their experience in the first six months' operation of the code was that the four largest supermarket chains had changed the way they do business. According to *The Grocer* (7 September 2002) 'The answer is a resounding no'. On 13 December 2002, the *Financial Times* reported the President of the National Farmers' Union as saying that among farmers there is a 'pervasive fear' that complaining to the OFT could lead to retailers cancelling contracts.

It seems safe to say that the code has had less impact than may have been expected by the Competition Commission. This appears to be the verdict also of the (government-appointed) Policy Commission on the Future of Farming and Food. In a wide-ranging review of UK farming, it commented that the dispute resolution mechanism of the code is perceived as being one-sided, and it recommended replacement by a clearly independent mediation service. It also proposed that the OFT should review the code's efficacy after two years' operation; in the meanwhile suppliers and other retailers should be exhorted to adopt the code voluntarily (Policy Commission, 2002:33-4).

In its response (on 12 December 2002), the Government acknowledged the imbalance of power, but explicitly refused to strengthen the code. Instead, the OFT undertook to seek views on the operation of the code, and annually to review and report on its effectiveness. It announced its first review on 17 February 2003.

### **Conclusions**

Retailer buying power can be beneficial if it countervails any market power of suppliers *and* if, under competitive pressures, the retailer shares the benefits with its customers. Clearly, large retailers have stimulated many improvements in supply-chain management; aggressive buying has forced suppliers to innovate, and to seek other efficiency gains. Yet, some buying strategies may *damage* efficiency, as for example when a buyer imposes conditions that raise supplier costs by more than the savings to the supermarket.

Further, some practices may create indirect or long-term effects that may or may not be in the public interest. Supermarkets sometimes secure gains by *ex post* opportunism that breaks contractual agreements. Such unconscionable conduct might eventually have adverse effects on market efficiency, to the extent that it undermines faith in the contractual process and the legal enforcement thereof.

Eventually, aggressive conduct by supermarket buyers may discourage supply and investment. Though it is unlikely to be in the interest of the company to go that far, the short-term incentives facing an individual buyer, whose career interest may encourage aggressive tactics, may not be in line with the long-term interests of the company.

Suppliers may respond to powerful buyers by agreeing to act together, in the hope that this will increase their countervailing power. For example, it may be possible for fruit growers to form cooperatives to establish cool stores for those products that do keep. This could give the growers some market power, stemming from scale and from an ability to ride-out short-term strategic buying.

Whether collaboration is feasible depends also on competition law. Recently, the (Australian) Competition Tribunal granted a limited authorisation to dairy farmers to negotiate collectively over the supply of raw milk to dairy processors. Where competition authorities are concerned about buyer power, they may be able to do more to facilitate supplier cooperation, without undermining government policy on collusion.

The codes of buying practice try to reduce the imbalances in bargaining strength between suppliers and the larger retailers. They depend on supplier initiatives. *If* an 'oppressed' supplier is willing to lodge a complaint, then under the UK code, it may be relatively easy for the mediator to identify any wrongs in the transaction that is the subject of complaint. Besides the distributive consequences, this might bring efficiency benefits, as for example, when it deters arbitrary post-contract changes such that the financial cost to the supplier exceeds the buyer's financial benefit.

The potential of the Australian code seems much less certain. Where the contract is not in writing, it may be very difficult to evidence any 'improper' conduct of the buyer. Furthermore, the absence of specific provisions in the code may make it less likely that the ombudsman will declare the buyer to have sinned.

In neither code will the supplier receive compensation unless the buyer thinks it prudent to accept and act upon the referee's decision. Perhaps the most important doubt concerns the willingness of suppliers to lodge complaints. In both countries, the marked absence of instances of code mediation suggests *either* that there has been no bad conduct, *or* that suppliers think the procedure will bring them no benefit, *or* that suppliers are too frightened to lodge a complaint.

The last explanation could hold whenever a buyer is able to 'punish' a complaining supplier by taking action *outside* the deal that is the subject of complaint. A buyer with market power might so act in order to establish a 'tough' reputation and hence discourage suppliers from lodging complaints.

In Australia, anecdotal evidence suggests that such conduct may be feasible when supermarkets buy produce. On 8 July 2002 (two years after the introduction

of the code), the *Sydney Morning Herald*, in a series of articles on relations between growers and the large supermarkets, quoted one grower as saying

It only takes a dominant supermarket to not purchase fruit for two days, and the market is full and over-supplied. Then they come in and buy at their leisure because the price will have dropped.

If a supermarket *is* able to withhold *all* orders for a particular fruit or vegetable in that way, how much easier would it be to refrain from buying (directly) from an *individual* supplier who has lodged a complaint about the buyer's conduct?

The extent to which the large retailers pass on to their own customers some of the financial benefits they extract from their use of buyer power depends largely on the effectiveness of retailing competition. The differing levels of retail concentration in the two countries suggest that price competition may be more vigorous in the UK than in Australia. However, this issue has not been addressed here.

In the light of experience so far, the governments may need to consider ways of strengthening the codes. An obvious approach would be to prohibit some of the practices discussed in this paper, while perhaps exempting instances where there is an objective justification for the practice. For example, a practice might be justifiable if it yields benefits to competition and consumers that outweigh any detriments to suppliers. Subject to any prohibitions, retailers and suppliers would be free, as now, to negotiate the supply terms. The effectiveness of this approach would depend upon suppliers being willing to lodge formal complaints should they believe that buyers have infringed the prohibitions.

No doubt, the competition authorities in both countries will continue to keep the sector under review. They should be willing to be pro-active rather merely responding to complaints. Both the OFT and the ACCC have authority to investigate a market to see whether market power hinders competition. Moreover, they should scrutinise very closely any merger proposals, such as the bids for (the UK) Safeway made early in 2003, that would increase retail concentration.

## References

Australian Competition and Consumer Commission (1999), 'Submission to the Joint Select Committee on the Retailing Sector', Canberra.

Australian Competition and Consumer Commission (2002), *Report to the Senate on Prices Paid to Suppliers by Retailers in the Australian Grocery Industry*, Canberra.

Burt, S. (2000), 'The Strategic Role of Retail Brands in British Grocery Retailing', *European Journal of Marketing* 34:875-90.

*Code of Practice on Supermarkets' Dealings with Suppliers* (2001), Department of Trade and Industry, London.

Competition Commission (2000), *Supermarkets: A Report on the Supply of Groceries from Multiple Stores in the United Kingdom*, Cm 4842, The Stationery Office, London.

Dobson Consulting (1999), *Buyer Power and its Impact on Competition in the Food Retail Distribution Sector of the European Union*, (Report prepared for the European Commission, DGIV).

Heflebower, R. (1957), 'Mass Distribution: A Phase of Bilateral Oligopoly or of Competition?', *American Economic Review* 47(2):274-85 (Papers and Proceedings).

Joint Select Committee on the Retailing Sector (1999), *Fair Market or Market Failure*, Parliament of the Commonwealth of Australia, Canberra.

Monopolies and Mergers Commission (1981), *Discounts to Retailers*, HC 311, HMSO, London.

Office of Fair Trading (1985) *Competition and Retailing*, Office of Fair Trading, London.

Office of Small Business (2002), *Annual Review of Small Business 2001*, Department of Industry, Tourism and Resources, Canberra, <http://www.industry.gov.au>.

Policy Commission on the Future of Farming and Food (2002), *Farming and Food: A Sustainable Future*, Cabinet Office, London.

*Retail Grocery Industry Code of Conduct* (2000), Code Administration Committee, Canberra.

Tom, W. (1999), 'Slotting Allowances and the Antitrust Laws', Testimony of the Federal Trade Commission before the Judiciary Committee, US House of Representatives, 20 October 1999, <http://www.ftc.org>.

Woolworths Limited (2001), 'Undertaking to the Australian Competition and Consumer Commission' given under s.87B of the *Trade Practices Act 1974*, <http://www.accc.gov.au>.

*Dr Martin Howe (formerly Director of Competition Policy, UK Office of Fair Trading) has contributed very greatly to the analysis presented in this paper. Gordon Mills thanks also an editor and two referees for helpful suggestions, and takes sole responsibility for judgements and any remaining errors.*



## **Access Holidays for Network Infrastructure Investment**

**Joshua Gans and Stephen King**

National Competition Policy is approaching its ten-year anniversary. In August 1993, the independent committee of inquiry on National Competition Policy released its final report (Commonwealth of Australia, 1993). This report, commonly known as the Hilmer report, set in train reforms that have changed the landscape facing Australian industry. In telecommunications, electricity, gas, ports, railroads, airports and other utility sectors, the past decade has seen industry restructuring, corporatisation and privatisation. Monopoly bottlenecks, such as electricity transmission grids and rail networks have been subject to access regimes to open these facilities to competing service providers.

For existing facilities, regulated infrastructure access is relatively easy. The key bottleneck infrastructure in telecommunications, energy and transport largely pre-dates the Hilmer reforms. Even if regulators set the 'wrong' access prices, this infrastructure was not going to go away. Over time however, as bottleneck infrastructure requires major renewal, or new infrastructure needs to be built, the effects of regulated access pricing on infrastructure investment becomes critical.

Traditionally, economic analysis of infrastructure investment and regulation has focussed on time consistency. How does a regulator commit not to effectively expropriate an investment by setting a low allowed price or rate-of-return after the infrastructure expenditure is sunk? This research has highlighted the importance of legal constraints on regulators, the design of appropriate regulatory institutions and the potential for a regulatory 'contract' to form between infrastructure owners and regulators (Newbery, 1999:chapter 2; Sidak and Spulber, 1997).

With access regulation, however, there is an additional problem for high-risk projects. Intervention to guarantee regulatory access is generally only required if a project is at least reasonably successful. If a project is a failure, for example due to low realised retail demand, then there will be little if any wholesale demand for facility access. But in these circumstances, investors bear all the project costs. In contrast, if the project is highly successful, regulatory access most likely will be sought. Even if regulators allow investors a 'reasonable rate of return' in these circumstances, unless this return fully compensates investors for the *ex ante* risk associated with project failure, access regulation will mute investment incentives. Put simply, investors will bear all the down-side risk of the investment and face a truncated up-side return due to access regulation.

The potential problem of project failure is most relevant for high-risk infrastructure investments. Augmentations to existing gas pipelines or fixed

---

*Joshua Gans and Stephen King are Professors in the Melbourne Business School, University of Melbourne*

telephone systems often carry little risk. But investments in infrastructure facilities that involve new products or service new areas may involve considerable risk; for example, the Alice Springs to Darwin rail project or the building of new gas pipelines to previously unserved regions. King and Maddock (1999) and National Competition Council (2000) discuss regulatory risk for the Alice Springs to Darwin railway.

In its recent review of the National Access Regime, the Productivity Commission (PC) paid close attention to the investment issue. Most interestingly, it investigated the notion of an 'access holiday'. This would be a period of time (in the PC's assessment, 15 to 20 years), under which an investor would be free from access regulation. The idea is that such a holiday will increase investment incentives by allowing investors a period in which to 'prove' their investment and gain profits unhindered by regulatory intervention. The concept of an access holiday (or a 'null undertaking') to both promote investment incentives and reduce regulatory uncertainty has been floated since the mid-1990s (King and Maddock, 1996). In the end, the PC fell short of recommending the use of access holidays; suggesting instead some measure to reduce regulatory uncertainty.

We believe that the case for adopting an access holiday approach is stronger than the PC has indicated. The need for an access holiday to spur investment arises from the *ex ante* inability of regulators to commit to access prices that adequately reward investors for all relevant risks. A well designed access holiday can partially overcome the problem of regulatory commitment, and represents a second-best solution to this problem, by limiting regulatory intervention for a number of years. In this paper, we consider the underlying problem of investment incentives and access regulation. We also highlight precisely what an access holiday can and cannot do. In particular we highlight important practical issues that need to be considered regarding access holidays.

### **Access Regulation and the Truncation Problem**

Why might access regulation lead to a reduction in infrastructure investment? To answer this question, consider a simple example. Suppose there is a single firm that can build a new infrastructure facility that might be subject to declaration. For example, the facility might be a cable network in a country town that is to be used for pay-TV distribution. Such an investment is likely to involve significant *ex ante* risk. While the cable firm can carry out market research to try and estimate the demand for pay-TV and other services that will use the cable, any estimates of future demand and revenues will be uncertain. The services provided by the cable may be a success, reaping significant revenue, or they may flop, leaving the investors with a large debt and a cable of little value.

For the purpose of this example, suppose that the cable network will cost \$51m to install. If the pay-TV services carried by the cable are successful, then the firm will be able to set a relatively high price and to make significant sales. It will receive (in present value terms) \$100m. If the pay-TV service is moderately successful then the firm will set a lower price and receive \$60m. If, however, the

pay-TV service is unsuccessful then the firm receives only \$20m. Suppose that the probabilities of each of these outcomes are 25 per cent, 50 per cent and 25 per cent respectively. Then the expected return to the investor is \$60m, less the \$51m cost of the investment. So the firm expects to make \$9m on average and (if risk neutral) will invest in the cable network.

Now, suppose that after building the network, the pay-TV distribution services provided by the network could be subject to regulated access. Given that the cable is already built, allowing access would appear to be socially desirable. Alternative pay-TV operators will be able to use the cable and competition in the provision of pay-TV will lead to lower prices for customers. A potential access seeker, however, can wait until it has observed whether or not pay-TV is successful before it decides whether or not to buy cable access and enter the retail market. In general, the access seeker will find it most desirable to enter into retail competition if the retail profits from pay-TV services are high. This is the situation when pay-TV is a successful product in the town. But in this case, entry by the pay-TV competitor using a regulated access price will tend to push down total industry profits. After all, that is the whole point of the infrastructure access regime: to facilitate downstream entry and competition, lower prices to customers and eliminate monopoly profits. So when pay-TV is most successful, entry by access seekers is most likely to occur, reducing the total profits of the industry and, as a result, reducing the profits of the cable investor.

In contrast, if pay-TV is unsuccessful, then entry by an access seeker is less likely. Rather, the investors in the cable are left bearing all the losses associated with their investment.

To see the effect of this *ex post* regulated access, suppose that an access seeker only finds it worthwhile to enter the market and compete in the supply of pay-TV services if demand is high in the town and, in the absence of declaration, the incumbent would receive \$100m. Further, suppose that in this situation competition from the access seeker lowers the return to the network owner to \$60m. *Ex post* this leaves the network owner receiving \$9m more than the cost of the network. But *ex ante*, the investment in the network is no longer viable. The network owner now expects to receive \$60m if pay-TV is successful, \$60m if pay-TV is only moderately successful and \$20m if pay-TV is unsuccessful. Using the same probabilities as before, this gives an expected return of \$50m, less than the capital investment of \$51m. The expected return from the investment is now negative one million dollars. Hence, if the network owner believes that its cable will be subject to a regulated access regime after it is built, then the owner will not bother to invest in the network in the first place.

This example illustrates a general principle. Access regulation is an *ex post* decision that affects the *ex ante* expected return from an investment. Further, access is most likely to be sought when the downstream returns from the investment are high. This means that *ex ante* the potential for declaration will limit the returns from an investment in those situations where it is most successful and might make socially desirable investment privately unprofitable.

This problem — where the potential for *ex post* access regulation ‘truncates’ the high end of the distribution of the possible returns to an investor in an essential infrastructure facility and so reduces the *ex ante* incentive to invest in the facility — is called the *truncation problem* by the PC.

### Can Access Prices Fix the Truncation Problem?

The potential for regulatory truncation depends on the access regulation imposed on investors *after* they make their investment. If a regulator could, *ex post*, set access prices that allow an investor an appropriate return to cover all relevant *ex ante* risk, then regulatory truncation need not arise.

To see this, consider the example above. If the regulator *ex post* was able to set an access price that guaranteed the cable investor profits of at least \$64 if pay-TV is successful, then the project would remain *ex ante* viable. More generally, if regulators can set access prices that allow investors relatively high profits when risky projects are successful, then the truncation problem can be avoided for all but the most marginal projects.

Regulators, however, do not have absolute discretion over access prices. Regulators are constrained by legislated pricing rules that limit the range of access prices that they can set *ex post*. Even in the absence of such legal constraints, regulators face strong economic and political pressures to limit access prices. After all, the essential facility has been built and *ex post* it is efficient to set price at a low level to encourage socially efficient use of the facility. Further, if the regulator did set relatively high access prices in those situations where the essential facility investment was successful, the returns received by the essential facility provider would appear ‘excessive’ *ex post*. Of course, this ignores the possibility that the investor could have lost significantly on the project. However, after the investment has been made, regulators, politicians and the public only see the actual return.

Even if the regulator could impose *ex post* access prices that fully compensated the infrastructure investor for *ex ante* risk, there are formidable information barriers in its way. Regulators do not know the true *ex ante* project risk and investors will tend to exaggerate their claims of such risk in order to boost the regulated access price. Access seekers will argue that such risks are overstated or non-existent. The regulator will almost certainly lack the information necessary to set the appropriate access price.

In summary, the truncation problem for essential facility investment is directly related to the issue of regulatory commitment. If the relevant regulator was able to commit *ex ante* to access prices that would apply *ex post* and would fully compensate for project risk, then the truncation problem would disappear. However, such commitment is unlikely in practice due to legal, political and practical constraints on the regulator.

It was argued to the PC that the truncation problem could be avoided by simply adding a ‘truncation premium’ to the access price and the PC appeared to favour such a premium in its report. Of course, if regulators could commit to an

adequate premium then the truncation problem would go away. The source of the truncation problem is the inability of regulators to commit to such access prices *ex ante*. Saying that the problem can be avoided if regulators could commit to such prices is true but trivial.

## Regulatory Instruments

The truncation problem arises because of the inability of regulators to commit *ex ante* to access prices that compensate for all relevant investment risks. But access prices are only one of the regulatory tools available to deal with essential facilities. A second regulatory tool is the right of a seeker to gain access to an essential facility. Once a facility is built, a seeker might be granted access quickly. Alternatively, access might be delayed for a period of time.

There is an interplay between the two instruments in determining the incentives of industry participants. In particular, the access pricing terms may be ‘high’ or ‘low’. Less familiar is that access rights are a matter of degree. Such rights may be ‘sooner’ or ‘later’.

In simplistic terms, the impact of these regulatory decisions on incentives can be summarised as in Table 1.

**Table 1: Instruments and Investment Incentives**

		Access Price	
		Low	High
Commencement of Access Rights	Sooner	Low provider/ High seeker incentives	Medium provider/ Medium seeker incentives
	Later	Medium provider/ Medium seeker incentives	High provider/ Low seeker incentives

Table 1 illustrates how changing the mix of access pricing and the commencement of access rights alters the balance of incentives along the vertical production chain. If access prices are relatively low and a seeker can gain access quickly then potential access seekers have strong incentives to try and enter the downstream market by gaining access to the services provided by the essential facility. But the infrastructure investor has weak incentives to build the essential facility in the first place. The investor will have little time to seize monopoly profits from the investment and will be subject to strong competition from access seekers. The converse holds if access prices are high and there is a significant delay between the time a facility is completed and the time that seekers are able to gain access.

Table 1 captures the essence of the truncation problem — regulators cannot commit not to set low access prices and if investors expect to receive low access

prices then they will be reluctant to invest in the first place. However, Table 1 also suggests a way out of this problem. If regulators are expected to set low access prices *ex post* but, at the same time, there is a commitment that any new essential facility will not be subject to infrastructure access for a significant period of time, then this will raise investor incentives. It is this *ex ante* commitment to delay access *ex post* that is the basis of an access holiday.

## Access Holidays

*What is an 'access holiday'?*

An access holiday is simply a period of time during which a new infrastructure facility would not be subject to any access regulation. During this time, the owners of the relevant facility could charge monopoly prices or deny access to competitors. Thus, during the access holiday period, the owners of the infrastructure would be free to exploit their investment in any way that they chose.

An access holiday helps to avoid the truncation problem because it can be granted before the relevant facility is built. To see this, consider the example of a cable network discussed above. The parties who are planning to build this network can approach the authorities before any significant investment has been made in the infrastructure and request to receive a period of exemption from regulated access that will begin after the project is completed. Regulatory authorities can commit to this holiday *ex ante*. After all, they simply have to state a relevant period of time for the access holiday — whether it is five years or fifty years. The investor then knows that any profits (or losses) made during the holiday period accrue to them alone.

For example, suppose in the cable network example that half of all returns (in present value terms) accrue in the first fifteen years. Remember that if the facility is subject to access as soon as it is complete, then the facility is *ex ante* uneconomic to the investors despite it being socially worthwhile. What happens if the regulator *ex ante* can commit to a fifteen-year access holiday?

Assume that after the holiday is finished, the regulator simply imposes the old access regime. As before, if the project is either unsuccessful or only moderately successful, then no party seeks access to the cable. The investors make either \$20m or \$60m with probabilities  $\frac{1}{4}$  and  $\frac{1}{2}$  respectively on the \$51m investment. If, however, the cable is highly successful, then access will be sought as soon as the holiday is complete. In the absence of any access regime, the present value of profits to the investors is \$100m. This falls to \$60m if access occurs as soon as the cable is complete. But if there is a fifteen-year access holiday, during which time half of the life-time returns from the cable are realised by the investors, then (in present value terms) the investors now receive \$50m during the holiday and \$30m after the holiday finishes. In total they receive \$80m when the project is highly successful.

The access holiday has increased the return to the investors in those situations where an immediate access regime would have undermined profits. Because the

access holiday is only finite in duration, it does not restore all monopoly profits to the investors. However, in this example, the holiday restores enough profits to ensure that the cable network will go ahead. The present value expected returns for the investors under the access holiday have risen to \$55m (that is,  $\frac{1}{4}$  times \$80m, plus  $\frac{1}{2}$  times \$60m, plus  $\frac{1}{4}$  times \$20m). Given the cable cost of \$51m, the project will now go ahead. The access holiday has enabled the regulatory authorities to overcome their inability to commit to *ex post* access prices and enabled the socially-desirable investment to proceed.

In terms of Table 1, the access holiday is simply a delay in the commencement of access rights. Given that regulators will set low access prices (at least from the investors' perspective) once the facility is built, we are caught in the left-hand-column of Table 1. The access holiday moves us to the bottom row of that table, raising the incentives to invest in essential infrastructure facilities.

It is important to note that the granting of an access holiday on an infrastructure investment does not mean there will be no access given to the relevant facility. Rather, it means that there is no *right* to access. However, it is still open for the provider to negotiate access terms with potential seekers. During the holiday period, the bargaining power of the access provider is clearly strengthened relative to a regulated access regime as the provider can choose not to allow access unless the access price is favourable enough. Indeed, in situations where the provider and seeker are not directly competing with one another or where the access seeker is more efficient in downstream operations than the upstream infrastructure owner, we would expect access to be granted resulting in private as well as social benefits.

### Access Holidays and Patents

As noted above, an access holiday is a second-best solution to the *ex post* constraints that limit regulated access pricing. In this sense, access holidays play a role similar to a patent in innovative activity. Patents encourage innovations by conferring on the inventor temporary monopoly profits. Similarly, access holidays encourage infrastructure investment by allowing investors to temporarily exploit any market power associated with their facility. Both patents and access holidays are second-best solutions in that they impose a temporary monopoly cost on society. Both an optimal patent and an optimal access holiday needs to be designed to trade-off this temporary loss with the increased incentive to invest.

The analogy between a patent and an access holiday raised some concerns, summarised by the PC in its report. A patent, like the access holiday proposed above, is a fixed period of time and does not attempt to vary — becoming longer if the inventor has not accrued enough return or shortening if the investor has accrued too much return. The access holiday, like a patent, would start from the time the investment 'comes on line' rather than from the time it generates 'positive revenues'.

These features, in our opinion, are strengths, not weaknesses of an access holiday. The length of the holiday is determined before investment and fixed. It

is not altered after the investment is made and is not revisited by the regulators when market information on project success or failure becomes available. In other words, it solves the regulatory commitment problem. If the access holiday was variable *ex post* at the discretion of the regulator then it would simply suffer from the same commitment problems that face access prices.

Allowing a regulator to *ex post* revisit the length of the holiday would not only raise the spectre of regulatory commitment, it would also create severe moral hazard problems of the sort that exist in all price regulation. For example, suppose the holiday included all time up until an investment became 'cash-flow positive' then lasted for a fixed time after this date. This would provide strong incentives for investors to delay the date at which their investment becomes cash-flow positive. They would have an incentive to artificially inflate the variable costs associated with the project, particularly in the short-term. This could be achieved in a variety of ways: substituting variable factors for fixed factors, bringing expenses forward in the project life to keep down cash flows, structuring debt so that it is biased towards short-term payments and so on. A fixed length access holiday that applies from a specified date (for example, a certain time after construction commences, or the date on which the project comes 'on line') will minimise the incentives for costly manipulation of cash flows and information by the investors.

In summary, like a patent, a fixed term access holiday provides a relatively simple regulatory procedure that avoids the problem of regulatory commitment and is relatively free of manipulation.

### **Criteria and Length of Access Holidays**

In its report, the PC noted that there are different types of infrastructure investment. For example, some investment can only be undertaken by one firm because it involves existing infrastructure or easements associated with that firm. Other infrastructure is contestable in the sense that a number of firms might all find the investment feasible and may even race to be the first to build. Similarly, different investment faces different levels of *ex ante* risk. While a new cable network in a previously unserved country town may be highly risky, an augmentation to an existing cable network in a metropolitan area with a long history of service demand might have little risk. When judging both whether an access holiday should be granted to a facility and the length of any holiday, these differences need to be considered.

### **Risky Monopoly Investment**

We begin with the example that has implicitly motivated our discussion so far. Suppose that the relevant infrastructure investment can only be undertaken by one group of investors and that this investment is highly risky *ex ante*. In fact, the risk associated with this investment is such that, in the absence of any access holiday, the investment will not go ahead despite it being socially desirable.



In this situation, an access holiday can overcome the truncation problem, but how long should the holiday be? At a minimum, the holiday must be long enough to overcome the truncation problem. In other words, the access holiday must be of sufficient length so that, given the post-holiday regulatory regime, investors still find it profitable to invest in socially desirable infrastructure *ex ante*. For example, with infrastructure with a life-span of 30 to 50 years, the relevant minimum holiday might be between 10 and 20 years for high-risk projects.

At the same time, additional factors come into play when determining the optimal length of the access holiday. The underlying idea of a holiday is to improve investment incentives. But this has two parts. First, the access holiday overcomes the truncation problem and leads to privately profitable investment in situations where investment would otherwise not be privately profitable at any time. This can be achieved by the minimum holiday that makes investment profitable. However, once private investment is going to take place, it is desirable to align the timing of this investment with socially optimal investment timing. In general, a monopoly investor will tend to delay investment relative to socially optimal timing. The reason for this is simple. When a firm invests, it bears all the costs of that investment but does not reap all the social returns. Even a monopoly leaves some consumers' surplus that cannot be seized by the firm so that social benefits tend to exceed private benefits from investment. Thus, when a private firm considers whether to invest today or wait until tomorrow, it considers all the benefits of waiting, in terms of reduced investment costs and improved technology, but it does not consider the full social cost of the surplus foregone by waiting. The firm only considers the loss of profits due to delayed investment, and ignores the lost consumers' surplus. The result is that the firm will tend to wait too long to invest from a social perspective.

The tendency for a monopoly investor to delay means that it may be optimal to extend the access holiday beyond the time that just makes the investment profitable, so long as this will lead to investment being brought closer to its socially optimal time. This in turn depends crucially on the project's life-cycle.

Increasing the length of an access holiday allows investors to earn monopoly profits for a longer period of time. If an access holiday ends when per period monopoly profits are relatively high, then increasing the length of the holiday will encourage earlier investment. In such a situation, increasing the access holiday beyond its minimum length will tend to accelerate investment in a socially desirable way.

On the other hand, if the profits at the end of an access holiday are relatively low, the benefits of increasing the length of the holiday may also be low. In that situation, it could be the case that the minimal length access holiday that just avoids the truncation problem is also socially desirable.

Two lessons flow from this. First, while at a minimum the access holiday must be long enough to overcome the truncation problem, so that investment is just privately profitable, it might be socially desirable to go further and create an access holiday long enough to leave some *ex ante* expected economic rents with the investor. Put simply, socially optimal access holidays need not, in general,

force private expected profits to the minimum that is just needed to allow investment. Leaving rents with the investors can be socially beneficial.

Second, it is only beneficial to leave rents with the investors if this promotes earlier investment. This in turn depends on the project life-cycle of the investment. The regulator needs to be cognisant of the likely timing of the flows of returns to the project when setting the length of the access holiday.

### Low-Risk Monopoly Investment

Our discussion so far has centred on the truncation problem facing high risk investment. Because regulators cannot commit to access prices, private investment might be rendered unprofitable. At the heart of this problem is the *ex ante* project risk and the failure of regulators *ex post* to fully compensate investors for the *ex ante* risks of project failure.

This suggests that there would be no role for access holidays in situations where infrastructure investment is low risk, so that even with *ex post* regulated access pricing the investment is privately profitable. However, this is false. Access holidays can play an important role even when infrastructure investment is privately profitable, because it can alter the timing of this investment.

Consider the case where only a single group of investors can make the infrastructure investment. The investment is low risk, for example, an augmentation or an extension to an existing network. As noted earlier, in this situation investment will tend to be undesirably delayed from a social perspective. While the monopoly investors will invest, they delay the investment relative to the social optimum. An access holiday can help to deal with this problem.

To see this, suppose the monopoly investors have to decide when to invest in a project that will generate a social surplus of \$53 per year for ever or \$530 in present value terms where the discount rate is 10 per cent. Also, imagine that due to technological change, real investment costs are declining. These are depicted in the second line of Table 2. The third line of the table presents the net present value of the social value of the investment if it occurs in a particular year. Thus, if the investment occurs in year 1, total social benefits are \$530 and the cost is \$350. So the net social benefit is \$180 in year 1, or \$164 today at a 10 per cent discount rate. Notice from the table that the net present social value of the investment is maximised if the investment takes place in year 2.

Now consider a scheme (first derived by Sappington and Sibley, 1988) — which can be termed the ‘incremental surplus subsidy’ scheme or ISS — in which the investors are paid the social surplus created by the investment for one year. During that year, the investors must also pay the costs of the investment. Once the year is over, the firm is then compensated for the actual costs of its investment in each successive year.

This regulatory scheme perfectly aligns social and private incentives over investment timing. When the monopolist invests it receives the social benefit of \$53 for one year. But they must pay the capital cost of the investment for that year. Given the 10 per cent discount rate this capital cost is just one tenth of the

total investment cost; \$54 if they invest today, \$35 if they invest in year 1, and so on. The monopolist will not invest immediately because the annual capital cost of \$54 for one year outweighs the payment of \$53. The monopolist will also not invest after one year. While they make a profit of \$18 if they invest after one year (that is, \$53 minus \$35), they make a profit of \$25 (that is, \$53 minus \$28) if they invest after two years. Discounting these profits back so that they are expressed in 'current' dollars, the monopolist makes approximately \$21 by investing after two years rather than \$16 by investing after one year. The monopolist's present value of profits under the ISS scheme are presented in the final row of table 2.

The monopolist, however, will not wait for three years. Investing after three years makes the monopolist a once off profit of \$26. But in terms of current dollars this is just under \$20. The provider prefers to invest after two years as this maximises its profit. This is also the socially optimal investment timing.

**Table 2: Relevant Investment Outcomes**

<b>Year</b>	<b>Today</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
<i>Social Benefit</i>	530	530	530	530	530
<i>Investment Cost (Current)</i>	540	350	280	270	260
<i>Net Present Value at 10% discount rate</i>	-10	164	206	195	184
<i>ISS</i>	-1	16	<b>21</b>	20	18

Source: Gans and King (2000)

The intuition behind this optimal regulatory scheme is straight forward. The pricing formula for investment gives the monopoly one period of social surplus but requires them to bear the investment costs for one period. The difference between the one-period social surplus generated by the investment and its one-period cost provides a rent for the investors. If this rent is increasing (in present value terms) over time, then the investors will prefer to wait and defer the investment. The investors will maximise this rent by investing when the difference between social surplus and investment cost is at its greatest (taking into account impatience). Hence, the investors' incentives over investment timing are perfectly aligned with the social incentives.

A problem in applying the ISS scheme is that it will often be difficult, if not impossible, to measure the social surplus. Sappington and Sibley suggest, however, that by letting the monopolist have free-reign and be unregulated for a year, then at least the monopoly incentives to invest can be generated without having to grant the provider a life-long monopoly. This suggestion, of course, is simply a short access holiday.

A short access holiday may help to align private and social investment incentives even when there is little project risk and the truncation problem is absent so that private investment will occur at some time. In such a situation, the short access holiday will tend to promote earlier private investment in a way that is socially desirable. However, the access holiday also creates a period of monopoly pricing with its associated loss of social surplus. The length of the holiday needs to trade-off the investment timing benefits of a short holiday with the monopoly pricing distortion that arises during the holiday. The access holiday needs to be sufficient to allow the investors to reap some rents from the project. But (given that these are for low-risk projects) this need not be long; perhaps five years or less.

### **Contestable Investments**

The discussion above has concentrated on the situation where only one group of investors can make the relevant infrastructure investment. But, as the PC points out, in many situations there are multiple potential investors. If investments are contestable, then investment incentives change substantially.

Suppose an investment is low risk, but contestable. If there is no access regime but the investment is profitable then there will tend to be a 'race' to invest. The first firm to invest in the facility will win the race and own the essential facility with all related monopoly profits. The result is investment too early from a social perspective. Firms will invest as soon as there are any profits to be seized, because they know if they wait then another firm will invest and leave them empty handed.

This situation is modified by access pricing. Suppose that after the essential facility is built it will be subject to an access regime. Then firms have a choice between investing and being the access provider and waiting. If they wait until another firm invests then they become the access seeker. The outcome is a 'game of patience' that depends on the expected profits from being an access seeker relative to the expected profits from being the access provider. In the short term, while investment costs are high, all firms will want to wait. As investment costs fall and technology improves over time, the benefits of being the access provider rather than the access seeker rise. At some time, the profits from being an investor just equal those of being an access seeker and at this point the game ends with one firm investing.

The game of patience will depend on the access price. The higher the access price, the more profits will flow to the access provider and the earlier the investment. The reverse holds for low access prices. This means that well designed access prices can shift the timing of investment, potentially aligning private investment timing with socially optimal timing. A general approach to this problem is presented in Gans (2001) with Gans and Williams (1999a, b) providing a regulatory solution to a specialised case.

A problem with this approach, however, is the ability of the regulator to commit to the desirable access pricing rule. If this is impossible, then an access

holiday may help to provide socially desirable incentives *so long as investment is delayed relative to the social optimum*.

Thus, again, an access holiday has two potential benefits for contestable investments. First, it can overcome the extreme truncation problem, where in the absence of a holiday no private investment would be forthcoming. Further, for low risk projects it may also help to align privately profitable and socially desirable investment timing. But its role here is muted relative to a direct approach that uses investment pricing. Further, an access holiday is less likely to align timing than in the case of a monopoly investment. This is because the holiday, in general, cannot overcome an investment race — it can only help to speed up delayed investment.

The reduced efficacy of an access holiday for contestable investments runs counter to the PC's conclusion. The PC argued that an access holiday was most useful for contestable investments. In fact, our analysis suggests the opposite. While an access holiday can overcome the truncation problem for either monopoly or contestable investment, it can also help offset the delay associated with monopoly investment. But it cannot in general offset the race that is associated with contestable investment and may, in fact, undesirably exacerbate this race if a holiday were misapplied.

### Adjustments for Investment Value

Thus far, we have considered access holidays that are of a fixed length and perhaps industry specific. It is conceivable, however, that the holiday length could be related to information about the value of the investment taking place.

The PC received several submissions arguing that the length of a holiday should alter along with the evolving realised value of the investment. The most notable was from the Network Economics Consulting Group (NECG) who recommended an endogenous time for the access holiday: that is, it would last until the net present value (NPV) of the investment becomes positive. What this appears to mean is that the holiday will end as soon as the investment costs have been recovered according to some *ex ante* agreed rate of return. Regulation after the holiday will then involve some form of profit sharing modelled on a resource rent tax.

Before evaluating this NPV approach it is useful to reflect on what a 'state-contingent' access holiday would look like. The principle that drives the length of an access holiday is simple: extend it only if the social loss from monopoly is low relative to the improvement in investment incentives over the extension period. So an optimal access holiday regime would shift profits to those states where monopoly behaviour had a relatively low social cost and away from those states where it is relatively costly.

In practice, this counters the NPV proposal. That approach would tend to equalise investors' profits regardless of the state realised. This runs counter to the socially optimal state-contingent holidays that involve higher profits in states where monopoly behaviour creates less social deadweight loss.

The problem with an endogenous approach to the length of an access holiday is that it creates exactly the type of issues that the access holiday was designed to overcome. We noted this above. A regulator rarely has the necessary information to accurately judge NPV for the unregulated investment. That information usually lies with the investors. However, this is where private and social incentives conflict. When NPV is high, the investors have a greater incentive to conceal information about the desirability of the investment. The investors will expend effort in arguing and establishing that an NPV is low when it is high. As a result, the regulator will not be able to rely on any information from investors when determining whether a holiday should continue or be terminated. All of the information will have to come from its own information and the information it can extract from other market participants (Caillaud and Tirole, 2000).

Further, a state-contingent approach like the NPV proposal reintroduces the issues of regulatory judgment and commitment that an access holiday is designed to eliminate. It simply moves the commitment problem from one based on *ex post* access prices to one based on *ex post* regulatory judgment about NPVs or other industry variables. As such the NPV proposal would appear to be a step backwards.

## Conclusion

Access holidays are a second-best solution to the investment problem created by regulatory pricing. An access holiday can improve the timing of infrastructure investment from a social perspective, albeit at the cost of creating a deadweight loss of surplus due to temporary monopoly pricing.

Access holidays are clearly not the only way to improve infrastructure investment incentives in the presence of regulated access pricing. For example, a preferred outcome would involve regulators using optimal state-contingent access pricing. But this outcome is unlikely due to the legislative rules that bind regulators, political constraints that limit regulator discretion and the practical problems associated with optimal regulatory pricing. Alternative price-based solutions also face the problem of regulatory commitment. For example, Evans and Guthrie (2002) consider how alternative approaches to constructing price caps can alter investment incentives. However, such caps (or at least the key values to be used *ex post* to construct the caps) still need to be set in advance and may be more open to regulatory opportunism than a simple access holiday.

In our opinion, access holidays are a regulatory concept whose 'time has come' in Australia. Access holidays provide a way to overcome regulatory risk and to improve investment incentives for essential infrastructure. They are relatively simple, with the relevant authorities determining the length of the holiday in advance of any investment and then leaving the investors to act in their own interest. Further, access holidays simplify the task facing regulators after the holiday ends. Because the holiday compensates the investors for *ex ante* project risk, regulators do not need to try and evaluate this risk after the holiday ends. As such, a holiday removes one of the most contentious problems of access pricing.

This said, access holidays will not be the same for all investments. In particular, the appropriate length of the holiday depends on the risk facing the project and the contestable nature of the project. The lower the risk and the more contestable the investment, the shorter should be the relevant access holiday. However, as our analysis shows, even for low risk monopoly investment, a short access holiday can provide benefits, accelerating socially desirable investment.

It is sometimes argued (by regulators or access seekers) that access holidays are addressing a non-existent problem. In particular, it is claimed that there is no evidence that investment is deterred by post-investment regulatory opportunism. Of course, it is difficult to show instances of something that does not occur, and the experience with access regulation in Australia is relatively short. But overseas experience provides ample evidence of the potential for regulation to deter investment. One well-known example was the development by AT&T of the Picturephone in the 1970s. Support for this highly risky (but potentially highly socially desirable) was withdrawn by AT&T, in part because of regulatory opportunism.

[T]he regulatory regime provided strong disincentives for AT&T to incur initial losses to solve the start-up problem. If the service had been successful, AT&T's profits would have been limited by rate-of-return regulation. ... On the downside, it is clear that Picturephone costs would be disallowed (as indeed 'Heads, I gain a little; tails, I lose big time'.) Given these odds it is understandable that AT&T made only a half-hearted attempt to solve the start-up problem and fairly quickly cut its losses. (Rohlf, 2001:86).

Access holidays provide a way of avoiding the problem of regulatory commitment, but they do not solve all issues relating to investment timing. For example, access holidays cannot retard investment in the case of contestable investment and an investment race. In such a situation, alternative solutions need to be considered. This is clearly an area for future research.

## References

- Caillaud, B. and J. Tirole (2000), 'Essential Facility Financing and Market Structure,' mimeo., IDEI, France.
- Commonwealth of Australia (1993), *National Competition Policy: Report by the Independent Committee of Inquiry*, August, AGPS, Canberra.
- Evans, L. and G. Guthrie (2002), 'Efficient Price Regulation When Networks Have Sunk Costs,' ISCR Working Paper, August, Victoria University, Wellington.
- Gans, J. (2001), 'Regulating Private Infrastructure Investment: Optimal Pricing for Access to Essential Facilities,' *Journal of Regulatory Economics* 20:167-189.
- Gans, J. and S. King (2000), 'Options for Electricity Transmission Regulation in Australia,' *Australian Economic Review* 33:145-160.

Gans, J. and S. King (2002), 'Access Holidays and the Timing of Infrastructure Investment,' mimeo., Melbourne Business School, The University of Melbourne.

Gans, J. and P. Williams (1999a), 'Access Regulation and the Timing of Infrastructure Investment,' *Economic Record* 79:127-138.

Gans, J. and P. Williams (1999b), 'Efficient Investment Pricing Rules and Access Regulation,' *Australian Business Law Review* 27:267-279.

King, S. and R. Maddock (1999), 'Evaluation of the Proposed Competitive Imputation Rule for Third Party Access Regime for the AustralAsia Railway Project', Submission to the National Competition Council, March 1.

King, S. and R. Maddock (1996), *Unlocking the Infrastructure*, Allen & Unwin, Sydney.

National Competition Council (2000), *AustralAsia Railway Access Regime: Final Recommendation*, February, Melbourne.

Newbery, D. (1999) *Privatization, Restructuring and Regulation of Network Utilities*, MIT Press, Cambridge, MA.

Productivity Commission (2001), *Review of the National Access Regime*, Report No. 17, AusInfo, Canberra.

Rohlfs, J. (2001), *Bandwagon Effects in High-Technology Industries*, MIT Press, Cambridge, MA.

Sappington, D. and D. Sibly (1988), 'Regulating Without Cost Information: The Incremental Surplus Subsidy Scheme', *International Economic Review* 29:297-306.

Sidak, J. and D. Spulber (1997), *Deregulatory Takings and the Regulatory Contract*, Cambridge University Press, Cambridge, UK.

*We would like to thank Warwick Tudehope, Darryl Biggar and two anonymous referees for their helpful comments on an earlier version of this paper. Responsibility for all views expressed lies with the authors. A more technical presentation of many of the issues considered in this paper is provided in Gans and King (2002).*



---

## *REVIEWS*

---

### **Bridging the Disconnect in Australian Politics**

*Peter Dawkins and Paul Kelly (eds), **Hard Heads and Soft Hearts: A New Reform Agenda for Australia**, Allen & Unwin, Crows Nest, 2003.*

*Reviewed by **Ron Duncan***

*Hard Heads and Soft Hearts* is a compilation of opinions from some of the 80 papers presented to the *Towards Opportunity and Prosperity* Conference held in Melbourne in April 2002 and in articles written for *The Australian* newspaper in a lead up to the Conference. The Conference was organised by the Melbourne Institute of Applied Economic and Social Research of the University of Melbourne and *The Australian*. The publication was edited by Peter Dawkins, Director of the Melbourne Institute of Applied Economic and Social Research and Paul Kelly, Editor-at-Large, *The Australian*.

According to its editors, the aim of the conference discussion and the lead-up articles was ‘... to end the disconnect in Australian politics ... between our political dialogue and our real needs ... and to identify broadly agreed principles for a new and different reform agenda.’ (p. 3) One might ask whether the Conference did end the disconnect; whether a broadly agreed set of principles resulted; and what are they? Moreover, what are the major areas of disagreement remaining?

It cannot be seen from the book whether there was any meeting of minds of those attending the Conference. The book only presents the views expressed in the papers. There is no reporting of any discussion and its outcome. Nor was there a broadly agreed set of principles for future reform. The editors present their views of what these principles should be. These follow what is described as the ‘hard heads, soft hearts’ approach, which is essentially the continuation of reforms to improve resource allocation, productivity, and economic growth, while providing the opportunities for all in society to share in economic growth. While this may be a new concept in discussion of economic policy in Australia, these ideas have been at the forefront of thinking about how to approach development in poor countries for the past several years. Strategies for overcoming poverty in developing countries are now largely concerned with identifying and overcoming the obstacles to poor people participating in economic growth; which essentially means removing the obstacles to them gaining employment and control over income-earning assets. These constraints usually take the form of poor access to education, health, and markets, and insecure title to land. Dawkins and Kelly’s conclusions are basically the same: continue with economic reform while

improving policies in education, health, work and family, and ageing and retirement, to promote employment, especially for jobless households.

The book is divided into two main parts: the first presents views on the performance of the Australian economy in recent decades, largely involving discussion of past economic reforms and their impact on income distribution. The second part presents discussion of various important policy issues such as unemployment, education, health, population, work and family, and the environment.

In the first part there is general agreement among the discussants that productivity has increased since 1990 and that this increase has been largely due to the economic reforms of the 1980s. There is some concern that income inequality has increased, although this finding is disputed and there is no discussion of whether any increase in inequality may be due to faster increase in incomes at the top end, rather than no increase or a decline in incomes toward the bottom. There is widely agreed and valid concern about joblessness within low-income households, especially households with children. Ross Garnaut of The Australian National University makes the point that Australian society will not accept growth without equity and says that the emphasis must be on employment expansion. He and others push the case for tax credits and restraint of minimum wages as a means of overcoming the effective marginal tax rate 'trap' that has bedevilled welfare policy in Australia for three decades or so. I do not see any conflict between economic growth and equity viewed in this way. In fact, international evidence shows that economic growth that is not distorted by the creation of privileged positions — such as through import substitution policies, or through policies associated with the exploitation of natural resources or control over fixed resources such as radio, TV or telecommunications that privilege the few — is the best means of reducing poverty and increases incomes in much the same proportion from the top to the bottom of the income distribution.

Ken Henry, Secretary, Department of the Treasury, is another contributor to the first part of the book. He makes the case for concentrating on absolute poverty (defined according to the individual's 'capability of community participation', which will differ by country) rather than on income inequality. I agree wholeheartedly with this sentiment. Indeed, there is almost an obsession with income inequality among economists. Unfortunately, the 'politics of envy' is a game that some politicians are eager to play, taking advantage of the fact that individuals are too prone to judge their standard of living in relative terms. Economists would be much better employed arguing that income inequality is not something that should be of great concern if it comes about through individual effort rather than through government favours. Australia's measure of income inequality would worsen if Bill Gates moved his business here, but it would be hard to argue that we would be worse off. We should worry about absolute poverty, and that means worrying about how to create jobs.

Henry also argues for more flexible labour markets and effective education and training programs. However, there is no consideration of tax policy — a Treasury brief — and whether enough is being done to make sure that Australia

has an internationally-competitive taxation policy. The Treasury's 'closed market' philosophy towards taxation policy should be challenged. Later in the book John Freebairn of The University of Melbourne makes the case for an internationally-competitive tax regime in a world of highly mobile capital and skilled labour. He suggests this would mean lower marginal income tax rates and lower taxes on capital and more emphasis on the GST and payroll taxes. He points out that more capital and skilled labour would combine with unskilled labour to give higher productivity and, by implication, more jobs for the unemployed. The editors pick up the point in their final 'Way Forward' chapter, but only mention lowering the top marginal income tax rate and suggest that it may be something worth doing in the longer term. But surely the focus should be on taxation of capital.

Missing from the book is discussion of Australia's welfare programs, particularly the composition of welfare recipients. In order to discuss the importance of job creation to poor households, or the importance of pre-school and primary education for poor households with children, or the importance of reforms to the health system for the disadvantaged and the elderly, it is desirable to have a breakdown of the whole welfare distribution system by household characteristics and by form of entitlement. This would also help in discussing the balance between helping the disadvantaged through removing disincentives to the demand for and supply of labour in order to reduce joblessness and helping people directly through income redistribution policies. Unfortunately, Australia seems to have gone mainly down the latter path. Public crowding out of family welfare responsibilities appears to have been the curse of the high-income countries in the past century. Poor countries cannot do this. They have to concentrate on economic growth to overcome poverty.

The chapter on education is largely taken up with discussion of policy towards universities from Bruce Chapman and Peter Karmel from The Australian National University and Alan Gilbert from The University of Melbourne, whose views are well known. While there should be concern about equality of opportunity for students from lower-income families to enter universities, surely the bigger problems that have to be tackled are the difficulties with early education that lead to children from disadvantaged homes not having developed the skills that would enable them to enter and flourish in universities or TAFEs. The point is made by the editors that childbearing is increasingly being assumed by poorer households. Thus policies dealing with pre-primary and early primary education, so critically important to the development of literacy and numeracy skills, have to come to grips with this growing problem area.

Chapters 7 and 9 focus on Population Policy and Ageing and Retirement. Peter McDonald, of the Demography Program at The Australian National University, provides some projections of Australia's population under various assumptions about the fertility rate, life expectancy and immigration numbers. His favoured projection is a population peaking at 25-28 million by around 2050, assuming a fertility rate of 1.65 and migration averaging 80,000 per year. McDonald sees it as unlikely that Australia's fertility rate will fall below 1.4

because 'policy support' would kick in. Based on the tendency for demographers to under-forecast the decline in the fertility, the experience of other Western countries, and the unlikelihood that policy can affect the fertility rate, my guess is that McDonald's favoured projection is too high. The fertility rate could fall even further than McDonald's 'Low' scenario rate of 1.4, resulting in a lower peak in population around 2030.

Ross Guest of Griffith University and Ian McDonald of The University of Melbourne provided results of a modelling exercise that projects that future living standards would be largely unaffected by changes in fertility rates and immigration. They conclude that population policy should not be concerned with this issue. Ross Garnaut and Glenn Withers of The Australian National University in separate papers argue for a population policy that encourages population growth, particularly through immigration, because of its positive economic growth effects. The differences between their judgements and those of Guest and McDonald appear to rest on differences in assumptions about productivity growth, investment, and ability to retain young, skilled people. In the first part of the book Garnaut had also made the point that the rapid ageing of Australia's population will reduce the disadvantages of low-skilled labour and, by implication, increase job opportunities among the lower-income households. In what must have been one of the more entertaining papers at the Conference, Withers demolishes the arguments that NSW Premier Bob Carr has made for restricting immigration to avoid adversely affecting the environment of the Sydney region.

The book has an unusual format. Each chapter has an introduction by the editors outlining the views expressed by the various presenters whose papers are summarised in the chapter; followed by the summaries and several short boxes highlighting points made by others who had made presentations or written articles for *The Australian*. Then the editors draw conclusions from the foregoing material. The format is repetitive and very disjointed, although one becomes used to the style after a while. The use of large numbers of boxes allows the highlighting of many points of view and is probably an effective device for avoiding the adverse visual effect of large blocks of quotations.

*Ron Duncan is Executive Director, Pacific Institute of Advanced Studies in Development and Governance, The University of the South Pacific and Visiting Professor, Faculty of Economics and Commerce, The Australian National University.*

## Passion for Liberal Economic Reform

*John Hyde, Dry: In Defence of Economic Freedom, Institute of Public Affairs, Melbourne, November 2002.*

*Reviewed by Rob Albon*

This is a very important and timely book, that will be of interest to any student (in the broadest sense) of Australian economic policy making, past or present. It is particularly apposite that this book comes at a time when many aspects of real economic reform have stalled, amidst a fair amount of real or imagined 'reform fatigue' and a balance of political forces not conducive to radical change. In some ways, the situation is not dissimilar to that in which John Hyde and the other Liberal 'Dries' in the Federal Parliament found themselves some twenty five years ago. For those wanting liberal reform, this book will be an inspiration and a source of ideas about how to get things happening. Unwittingly, it also contains a few good ideas for those seeking to resist that reform.

The book begins with some background chapters setting the scene. These cover the role of dry philosophy; post-war Australian economic and political history from Chifley to Whitlam (and featuring the part played by another liberal farmer, Bert Kelly); the role of the think-tanks (with a big emphasis on the Institute of Public Affairs and the Centre for Independent Studies); and what was happening outside Australia (for example, through GATT/WTO; Thatcherism and in New Zealand). There is a very good chapter on the decline of industry protection ('From Tariff Board to Productivity Commission').

Then we get the accounts of the different governments from Hyde's time in Parliament and beyond. These cover Hyde's frustration at the lost opportunities in the Fraser years; his delight in the (somewhat unexpected) achievements of Hawke's governments; the near success of *Fightback!* (and its contrast with Keating's *One Nation*); a cameo on the 'Kennett Revolution' in Victoria, and a not particularly glowing account of the 'Howard Years' (and the Opposition). The book ends with a chapter on backlash and a very short one on the importance of what we say and do now for the future.

Different readers will take different things from this book, but there are three things that stand out for me:

Most important, is the inspiration and hope that comes from this book. Back in the early eighties the current circumstances and the outlook were rather bleak — a serious recession; a drought; heavily-protected industry; inefficient infrastructure; costly work practices; etc. and no reform in sight. To everyone's surprise, this dark period was followed by a long period of intense and broadly well-considered economic reform.

Second, the book stands as a reminder of where Australia would probably be if that reform program had not eventuated. While some people ask when we are going to see the gain from the pain, it is fairly clear that the productivity surge of the 1990s (and those forty odd quarters of continual economic growth) would not have occurred without the enabling reforms that preceded it. And — in spite of some claims to the contrary — this was achieved without any widening of the gap between rich and poor.

Third, there is the man, John Hyde, a farmer with a passion for ideas and the humanitarian liberal cause, and basically a ‘good bloke’. Because he was prepared to think outside the square John Hyde was criticised — and vilified — both from within and without the Liberal Party. As former Labor Finance Minister Peter Walsh reminds us in the Foreword, Hyde was even labelled a ‘fascist’. While Walsh earnestly sets the record straight on this, I have my own recollections of Hyde’s moderation. For example, he was invited to The Australian National University *circa* 1980 to address economics students, and he took us through the main arguments for government intervention with a clear focus on justification resting on the standard market failures and demonstration of the government being able to do better. The lecture went very well, but later John was somewhat bemused by the experience after gaining the impression that many of his audience found him too interventionist!

The Dries are an important part of Australia’s post-war economic experience, and it is great to see their role so skilfully and knowingly explicated.

*Rob Albon is an Associate Editor of Agenda.*

*The Editors congratulate Rob Albon on his award of a Centenary Medal for ‘services to Australian Society through the advancement of economic, social and political issues’.*

---

## ***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)

---

### **Why Ethics Committees Are Unethical**

**Paul Oslington**

Recently UNSW academics at the Australian Defence Force Academy were asked to attend a presentation by members of the UNSW Human Research Ethics Committee, who appeared to be somewhat disturbed by our previous lack of interest in their activities. Their posture was one of missionaries bringing ethics to us poor savages for the first time, and they began with a short history of ethics committees and guidelines. They explained that Nazi medical experimentation in the concentration camps is the reason we have human ethics committees in the universities today, and that they stand between us and these repugnant practices. When challenged with an alternative account of ethics committees arising to protect the university administration against litigation and scandal, and their expanding through the internal dynamics of bureaucracy, the ethics committee representatives became somewhat defensive. When persuasion failed they resorted to threat (that academics would not be covered by the university insurance policy if their research does not go through the ethics committee — untrue as it turns out) and coercion (no ethics approval, no research funds — unfortunately true). Ironically, the presentation and subsequent debate was being filmed without the slightest hint of the informed consent the committee representatives were so earnestly exhorting us savages to seek for almost everything.

Furthermore, when it was suggested that most of their exhortations were irrelevant as few in the audience are involved medical research, we received the quite incredible response that an activity which affects humans in any way, or an

---

*Paul Oslington is a Senior Lecturer in Economics at The University of New South Wales/Australian Defence Force Academy.*

environment in which humans live, must go before the committee. This assertion stretches — to say the least — the NHMRC guidelines. If this interpretation of the NHMRC guidelines is followed almost all economics research would require ethics approval. This would be a serious issue indeed.

Ethics committees have been part of the life of medical researchers for some decades, based on guidelines which flow from the World Medical Association's 1964 *Declaration of Helsinki*. This draws heavily on the Nuremberg Code developed during the trials of Nazi doctors after WWII. In Australia the NHMRC issued its first *Statement on Human Experimentation* in 1966, and both this and the World Medical Association's declaration have been amended many times. The current set of NHMRC guidelines is the *National Statement on Ethical Conduct in Research Involving Humans* issued in 1999. Further details of the history of ethics committees and rules may be found in NHMRC (1999) and Annas and Grodin (1992).

What is new is not so much the rules, as the attempt to expand the applicability of the rules to non-medical researchers. The argument seems to be that since non-medical research also involves humans or at least impacts humans it should be subject to the same set of procedures as medical research. Examples of problems with anthropological and sociological fieldwork are used to support such an extension (see for example Burgess, 1984:chapter 9). The extension to economics does not seem well founded, nor is it convincing that the burden of proof should be on those who resist the extension of the medical rules to economics.

It would be easy to write off our unfortunate encounter with the Ethics Committee representatives as an aberration, but there are reasons to be deeply concerned about the current approach to regulating research ethics in the Australian Universities, especially when extended beyond the original medical context. Discussions with colleagues at other institutions reveals similar incidents, and a widespread hostility to the activities of ethics committees.

Some concerns about the current approach to regulating ethics in Australian Universities are now discussed in turn.

### **A Particular Approach to Moral Philosophy**

The quasi-legal procedures of ethics committees privilege a particular rule-based approach to moral philosophy. The procedures are quasi-legal because the committee evaluates proposals based on the 'law' of the ethics guidelines and relies heavily on precedent. It is difficult for a committee to operate in any other way, but quasi-legal and legal procedures miss the point. The guidelines ethics committees follow present themselves as universal rules derived from reason. This rule based approach is one among many within moral philosophy and certainly not uncontroversial (see, for instance, MacIntyre, 1981; Williams, 1985). For instance, it is sharply inconsistent with the approach of virtue ethics (that is, ethics is about encouraging institutions that nourish virtue). It does not sit easily with 'consequentialism' (that is, activities are acceptable if good consequences



outweigh the bad) of which economists' favourite moral philosophy — utilitarianism — is a special case.

A committee must have some substantive ethical position, even if not written down or explicitly invoked by the committee. Ethical neutrality is self contradictory, and a dangerous delusion. The concern is not that the committee has a substantive position, but that a particular substantive position has been chosen unreflectively, as a consequence of the way the ethics approval process has been established. A quasi legal process implies rules. This rule based approach to ethics is not described, let alone defended in the various ethics guidelines. The committees seem mostly unaware of their substantive ethical position, believing themselves to be operating on an obvious and uncontroversial basis. Their focus (perhaps rightly, given the way the system has been set up) has been on the application of the rules to particular cases.

### **Perverse Incentives in Ethics Committees**

Committees are not very good at ethical deliberation, especially those operating in a quasi-legal way. It is even worse when committee members are subject to the reward structures of contemporary Australian Universities. Philip Pettit has discussed at some length the problems in an Academy of Social Sciences Lecture (Pettit, 1992).

One problem identified by Pettit is *asymmetric rewards* for wrongly approving and wrongly rejecting proposals. A committee which accepts a proposal which turns out badly (for example, a scandal reported in the press) will suffer serious consequences. However being over cautious and rejecting a proposal that should have been accepted is likely to have minimal consequences for the committee, and occasional complaints from researchers may even (as Pettit suggests) reassure senior university administrators that the committee is doing a good job. The reward structure suggests committees systematically will get it wrong.

Another problem is *ratchet effects*. In a rule based system it is far easier to invent new rules — scandals as Pettit notes lead to a flurry of this sort of activity — than to delete or ignore bad rules. The dynamics of the system will tighten ethical requirements on researchers over time, regardless of the appropriateness.

*Self-righteousness* is another danger suggested by Pettit — the ethics committee member enjoys the warm inner glow that comes from adherence to a set of ethical rules while costs of this are imposed on researchers and the general public. Still another is assertiveness — the understandable desire of the committee to avoid being seen as a rubber stamp. This, like the other tendencies, will lead to excessively tight rules and over-zealous committees.

It is sometimes suggested that ethics committees attract self-righteous busybodies, but Pettit's arguments are more that the incentive structures will push otherwise reasonable people to behave this way.

## **Maintenance of Ethical Responsibility**

Ethics committees undermine the ethical culture. The current procedures encourage researchers to treat ethics as someone else's problem. A division of labour will develop where the researcher does the work and a distant (often anonymous) committee makes sure certain ethical rules are adhered to. Ethics is trivialised. Researchers will get out of the habit of ethical reflection, except of the most minimal sort which considers what will get past the committee. Encouraging them to pass the ethical buck undermines the ethical capacities of researchers, and universities and disciplines as ethical communities.

A related issue is trust. It is not prudent to trust people in all circumstances, but the strong message to researchers in the current procedures that the University does not trust their ethical judgement could very well be self fulfilling.

## **Cost**

The current ethics approval system is very costly, and these costs will rise rapidly as ethics committees seek to expand their activities. It would be very interesting to have an estimate of the cost of the current ethical approvals process in Australian Universities. At the unfortunate presentation described above, a committee member pointed with satisfaction to graphs showing strong growth in applications to the ethics committee and staffing of the ethics secretariat. Apparently an increase in costs was a good thing — but it is unlikely that many academics, who have been doing more with less resources in recent years, would have been so easily persuaded.

Administrative costs of the ethics secretariat are probably just the tip of the iceberg. Academics serving on the committees speak of large numbers of applications to be read. Researchers at UNSW must read the *Rules Governing Human Research Projects at UNSW*, then decipher the obtuse and internally contradictory twenty page *Guidelines for Applications to the Institutional Ethics Committee at UNSW and its Affiliated Teaching Hospitals to Undertake Research Involving Humans* plus the sixty six page NHMRC (1999) *National Statement on Ethical Conduct in Research Involving Humans*, the several hundred page *Commentary on the National Statement on Ethical Conduct in Research Involving Humans* (NHMRC, 1999), plus about ten other lengthy documents referred to within the Guidelines. At UNSW the blank application form is twenty pages long, and applicants must submit nineteen copies to the Ethics Secretariat for review.

Another cost of the current system is the exposure of ethics committee members to legal liability as a result of the activities of the committees. Restricting their role or abolishing the committees would reduce this exposure, and perhaps even reduce the risk of litigation for the university as a whole.

## **Delay**

The need for ethical approvals before commencing research delays projects. Delays can run to many months while preparing applications and waiting for

ethics approval, even if the application goes through without the need for clarification or modification. Once approval is obtained that is not the end of the matter. If any change is made to the research as described in the application, approval must be sought and obtained in advance of this change. This is incredibly inhibiting of research projects, especially in rapidly developing fields, where new methods are being experimented with, and of research on contemporary issues.

Talking with colleagues in other institutions the activities of the ethics committee are having major effects on the topics on which academics undertake research and the methods they use. The ethics committee would probably see this as a gratifying sign that they are not wasting their time. However, the impression one gets is not that academics are avoiding unethical research, but that they are avoiding anything they perceive the ethics committee is likely to make trouble over or which is time consuming to justify. The burden of proof is on the researcher. It is likely that novel or unconventional research is being inhibited, rather than unethical research. This is no small matter given the large amounts of government and private money spent on research. It would be unfortunate if the excesses of human ethics committees drove this type of research out of the university system, or overseas.

### **Inappropriate Medical Model**

An inappropriate medical model is being applied to the social sciences and humanities. A recent report to the Executive Committee of the Academy of Social Sciences in Australia, summarised in Gregson (2002), raised concerns about the applicability of the medical rules — especially consent forms and procedures — to the psychological sciences. How much greater are the applicability problems for the social sciences and humanities.

In my own discipline of economics the most recent Nobel Prize was awarded to Vernon Smith for his work in experimental economics. The NHMRC rules prohibit offering money to subjects participating in experiments — for reasons one can understand in a medical context. However experimental economics is about seeing how subjects respond to monetary incentives of various kinds in different settings, so strict application of the NHMRC rules would cut Australian researchers out of this increasingly important area of economics. Common sense may prevail here, but this is just one example of the problems created when medical rules are applied outside their original context.

Another risk is that rules about harm will impinge more widely on economic research. One can readily imagine a zealous ethics committee member objecting to anything that has a whiff of ‘economic rationalism’ about it, as economic rationalism is known by all right thinking people to be damaging to society. Research involving organisations such as the World Bank, IMF and WTO could also come under scrutiny from our zealous ethics committee member. It is worth noting that under the current NHMRC rules there is no specific requirement for

the committees sitting in judgement of economic research having members with expertise in economics.

### **Summarising the Concerns**

These arguments can be drawn together by taking up the suggestion at the unfortunate ethics committee presentation that it is the ethics committee that stands between us and the horrors of Nazi medical experimentation. Suppose the Nazis had a Human Research Ethics Committee, and that all research had to be approved. It is unlikely that there would be much difficulty in Nazi Germany in finding sufficient people of the required types (that is, seven including a lawyer, doctor, and minister of religion) to sit on a properly constituted Human Research Ethics Committee under the NHMRC guidelines and approve the abhorrent Nazi experiments. A culture that generates these acts would have no difficulty generating committee members prepared to endorse the acts and construct justifications for them under almost any set of general principles, including the current NHMRC guidelines. 'Informed consent' will be obtained, harm of course is minimal as the experiments are not dealing with people, benefits for the Nazi state are great, and so on. In fact Mengele's notorious twin experiments at Auschwitz were approved by the Research Review Committee of the Reich Research Council (see Annas and Gordon, 1992:35). Ethics approval procedures and guidelines would seem to provide little real protection if the ethical culture is compromised or suppressed by power, as it was in Nazi Germany.

The Nazi era, as well as providing evidence that ethics committees and approvals in themselves mean little, suggests that principles devised by the well intentioned and a bureaucratic state are a potentially dangerous mix. Annas and Grodin (1992:17-31) document the rise of the racial hygiene movement in pre-WWII Germany, suggesting that it was developed by well meaning doctors and became 'orthodox' in the German medical community before the Nazis came to power. Such an orthodoxy in the medical community however could only become deadly when it gained the support of the Nazi state and was implemented with the help of the state's efficient bureaucracy and coercive power. Annas and Grodin's book also discusses the way 'mercy killing' evolved out of the original medical context to a social doctrine that justified the gassing Gypsies, Jews and other undesirables. It is particularly interesting the way ethical debate and individual dissent was crushed as these principles were appropriated, twisted and absolutised by a bureaucratic state. Committees functioned to co-opt doctors and spread Nazi ideology rather than restrain evil.

In conclusion, there can be no running away from ethical issues in research, but in my view the cause of ethical research in Australian Universities would be better served by scaling back the activities of ethics committees to ensuring compliance with legal requirements. A committee and approval process like the current one is probably not the most efficient way of ensuring compliance with legal requirements, so perhaps the Universities would be better served by some other structure or no structure at all.

Ethics committees and approval processes are a sign of and contributor to an unhealthy ethical culture. It would be good to see more discussion of the University as an ethical community and a greater level of ethical leadership from senior academics and administrators. Ethics education is also important, provided it is serious and not just about how to get through the committee or expanding people's capacity to justify dubious actions. Ethics should be more integrated with our day to day professional practice — and reclaimed from the bureaucrats.

In my own discipline of economics, as argued persuasively by Hausman and McPherson (1996) and others, ethical problems are inseparable from theoretical and empirical economics. As well as the ethical content of their models, economists face ethical issues in selecting topics, honesty and care. At the very least, if we are going to extend ethics approval procedures beyond the medical disciplines to economics we need to think beyond an inappropriate medical model.

## References

Annas, G. and M. Grodin (eds) (1992), *The Nazi Doctors and the Nuremberg Code-Human Rights in Human Experimentation*, Oxford University Press, Oxford.

Burgess, R. (1984), *In The Field: An Introduction to Field Research*, Allen and Unwin, London.

Gregson, R. (2002), 'Ethics of Non-Medical Human Experimentation', *Dialogue* 21(3):59-64.

University of New South Wales (undated), 'Guidelines for Applications to the Institutional Ethics Committee at UNSW and its Affiliated Teaching Hospitals to Undertake Research Involving Humans', available at <http://www.ro.unsw.edu.au/ethics/humanrules2000.rtf>.

National Health and Medical Research Council (1999), 'National Statement on Ethical Conduct in Research Involving Humans', available at <http://www.health.gov.au/nhmrc>.

National Health and Medical Research Council (1999), 'Commentary on the National Statement on Ethical Conduct in Research Involving Humans', available at <http://www.health.gov.au/nhmrc>.

Hausman, D. and M. McPherson (1996), *Economic Analysis and Moral Philosophy*, Cambridge, Cambridge University Press.

MacIntyre, A. (1981), *After Virtue — A Study in Moral Theory*, Duckworth, London.

Pettit, P. (1992), 'Instituting a Research Ethic: Chilling and Cautionary Tales.' *Bioethics* 6(2):89-112.

Williams, B. (1985), *Ethics and the Limits of Philosophy*, Fontana, New York.

*Thanks to Keiran Sharpe, Frances Miley, Cecelia Andrews, Damian Grace, Hugh Smith and other colleagues for helpful comments on an earlier draft. The views expressed are my own. I also thank the editor and anonymous referees for their comments which improved the paper.*