The effects of changes in family composition and employment patterns on the distribution of income in Australia

David Johnson and Roger Wilkins

Melbourne Institute of Applied Economic and Social Research

Paper presented to the Towards Opportunity and Prosperity conference, The University of Melbourne, April 2002

Trends in the distribution of family disposable income in Australia over the last two decades are examined using the unit record files of the ABS Income Distribution and Income and Housing Costs surveys. The effects on income inequality of government intervention, in the form of provision of transfers such as pensions, benefits and allowances, and through personal income tax, are also investigated. We find government intervention has acted to reduce growth in inequality in disposable incomes over the period, to a significant extent offsetting the substantial growth in private (before income taxes and transfers) income inequality. Decomposition techniques are applied to identify the roles played by potential sources of this growth in private income inequality, including the contribution of changes in family composition, changes in labour force status, changes in demographic factors and changes in the wage distribution.

1. Introduction and motivation

A number of researchers have studied changes in the distribution of income, and components of income (such as wages and salaries), over the last twenty years or so. The findings of this research may be summarized as follows. Inequality in the distribution of labour market earnings has grown significantly over the period, particularly in the last part of the eighties, but continuing through the nineties. Market income, largely reflecting what is happening with wages, has also become much more dispersed. Government intervention has, however, served to constrain the increases in income inequality. Inequality in the distribution of gross income, which comprises market income plus transfer payments, has increased only mildly, while the increase in inequality in disposable income, which additionally takes into account the effects of income taxes, has been even more muted.

The findings have suggested that the increases in inequality primarily derive from a greater rate of growth of incomes at the top end of the distribution than at both the middle and lower end of the spectrum. The evidence on expenditure inequality is somewhat different, with lower levels of inequality than for the income distribution, but the trend increase in inequality is nonetheless evident for the expenditure distribution. The research has also found that, despite significant changes in the composition of the household and family units in recent decades, the above trends are true of both individual incomes and household or family incomes (although the increases at the individual level are more marked than at the family or household level).

What is responsible for these changes? In particular, what is the source of the increases in market income inequality? There are several key possibilities: First, the widening disparity at the family level may to a large extent derive from increased dispersion in market incomes, particularly in wages. Second, the composition of families has changed, and it may be that a consequence of these compositional changes is increased inequality in incomes across families. Third, a number of changes have taken place in the labour market in recent decades, including increases in female participation, falls in male participation, delays in (full-time) labour market entry (as students prolong their studies), and possibly increases in both jobless families and two-earner families. These changes in the labour market are likely to have had significant effects on the distribution of income across families. Other possible sources of changes to the income distribution include changes in demographic characteristics such as the age structure, educational attainment and place of birth.

The objective of this project is to confirm the developments in the distribution of income in Australia over the past 20 years noted above and explore potential sources of the changes that have occurred. In particular, we decompose changes in the distribution of disposable income, considering the role of government intervention via transfer payments and income taxation, and examining the effects of such factors as labour force status, family structure and other demographic and economic variables. The project therefore provides important insights into the causes of trends in the distribution of income in Australia. These insights have wide-ranging policy relevance, including the evaluation of the effects on income inequality of policy in relation to income taxation, social security support, unemployment, and family and labour market policies more generally.

Consistent with previous research, we find that market income inequality grew substantially between 1982 and 1997-8, but that transfer payments and income taxes have exerted a strong equalising force over the period, thereby substantially reducing the extent of growth in disposable income inequality. Decomposition of the changes in the distribution of private income suggests that approximately half the growth in inequality between 1982 and 1997-8 derived from changes in the distribution in the population of labour force outcomes, family types and other

demographic and economic variables. Changes in labour force outcomes were a particularly important source of the increases in income inequality, with changes in the composition of family types in the population also acting to marginally increase inequality. Other changes in the characteristics of the Australian population, including increases in educational attainment and changes in the age structure, have, by contrast, worked to decrease income inequality over the period.

1.1 Discussion of trends and levels

Trends in earnings and wage rates

Table 1.1 shows trends in the distribution of full time earnings of Australians at different points of the earnings distribution between 1975 and 1999. The table describes the trends using the ratio of earnings at different points of the distribution. Thus P10/50 refers to the earnings of a person at the tenth percentile relative to the earnings of a person at the fiftieth percentile. Table 1.1 is reported by Borland, Gregory and Sheehan (2001). The table suggests that there has been substantial widening of earnings differentials between workers at different points of the distribution, with the earnings of higher paid persons increasing at a faster rate right through the distribution. For both males and females earnings for workers with below median earnings have decreased relative to workers at the median point of the distribution though the difference is more pronounced for males than for females. Similarly in the top half of the distribution the earnings of high earning men and women have increased relative to those of men and women at median earnings.

Table 1.1: Earnings dispersion –weekly earnings of full-time employees in main job 1975-99, Australia

	P10/50	*P25/50*	*P75/50*	*P90/50*	*P90/10*
Males					
1975	0.683	0.834	1.266	1.654	2.422
1980	0.625	0.816	1.316	1.714	2.742
1985	0.619	0.803	0.313	1.621	2.619
1990	0.593	0.777	0.309	1.616	2.725
1995	0.594	0.765	0.360	1.750	2.946
1999	0.590	0.76	1.401	1.878	3.183
Females					
1975	0.633	0.834	1.192	1.44	2.275
1980	0.604	0.802	1.225	1.538	2.546
1985	0.599	0.811	1.240	Na	Na
1990	0.604	0.804	1.281	1.604	2.656
1995	0.631	0.797	1.289	1.598	2.532
1999	0.620	0.793	1.323	1.661	2.679

Notes: from Borland, Gregory and Sheehan (2001) Table 1.1, p5

Trends in income unit composition

The top panel of Table 1.2 shows the proportion of all households comprised of each different household type in each of five years over the period 1976 to 2001.[1] The table shows a steady growth in the proportion of lone person households and a concomitant fall in the proportion of households comprising families. The data indicate that the proportion of people living in group households has stabilised at around 4.5 percent since the mid eighties.

The lower panel of the table shows, among households comprising families, the proportion with no dependents and with dependents. It shows there have been important changes in the composition of families between 1976 and 1996. In particular, there has been steady growth in the proportion of families without dependents, rising from 45 percent in 1976 to nearly 50 percent in 1996. Most of the growth has been in families comprising couples only, with growth in this family type partially offset by a decline in the proportion of families comprising couples plus non-dependents. The decline in the proportion of families with dependents is the net outcome of a very large decline in two parent families, from 48.4 to 40.6 percent of all families, and a partially offsetting increase in one-parent families, from 6.5 to 9.9 percent of all families.

Table 1.2: Household and family types, Australia 1976 to 2001 (%)

	1976	*1981*	*1986*	*1991*	*1996*	*2001*
Household type						
One person	15.7	18.0	18.5	19.6	23.1	25.2
Group			4.1	4.5	4.4	4.5
Family	84.3	82.0	77.3	75.7	72.5	70.3
Family type						
Families without dependents	45.0	44.7	47.4	46.8	49.5	
Couple only	28.0	28.7	30.3	31.4	34.1	
Couple plus non-dependents	11.1	10.0	10.9	9.5	9.0	
Other	5.9	6.0	6.2	5.9	6.4	
Families with dependents	54.9	55.2	52.6	53.2	50.5	
One parent	6.5	8.6	7.8	8.8	9.9	
Two parents	48.4	46.6	44.8	44.4	40.6	

Sources: ABS (2001a)

In summary, there have been important changes in the structure of household types. Couples with dependents now constitute less than a third of all households[2] and no longer represent the predominant family type, which is now couples without dependents.

[1] In 1976 and 1981, group households were not distinguished from other families.

[2] Specifically, 70.3 percent of households comprise families, and 40.6 percent of families comprise couples with dependents, implying 28.5 percent of all households are couples with dependents

Trends in the labour force and in work patterns of families

Table 1.3 presents employment-population ratios among males and females over the last two decades. The proportion of males in employment has fallen from 74 percent in 1981 to 67 percent in 2001, while over the same period female employment has risen from 41 percent to 52 percent of the female population.

Table 1.3: Employment-population ratios in Australia, persons aged 15 years and over, 1981 to 2001 (November)

	1981	*1986*	*1991*	*1996*	*2001*
Males	73.6	69.5	66.8	67.2	67.2
Females	41.3	44.5	46.9	49.4	51.6
Persons	57.2	56.8	56.7	58.1	59.3

Sources: ABS (2001b)

Associated with these changes in the labour force behaviour of males and females have been important changes in the relationship between families and work. Table 1.4 reports trends in the numbers of adults working in families with dependent children over the last ten years. There has been growth in families in which there are two workers and a decrease in families in which there are no workers or just one worker.

Table 1.4: Families and work, Australia 1991 to 2001

Household type	*1991*	*1996*	*2000*
Couple families with children under 15 and			
Two adults working	51.8	54.5	56.3
One adult working	40.1	37.6	36.2
No adult working	8.1	7.9	7.5
One parent families with			
One adult working	43.2	42.7	47.3
No adult working	56.8	57.3	52.7

Sources: ABS (2001c)

Not evident in Table 1.4, because of the limited time-frame and exclusion of households without dependents, is the finding of several commentators of an emerging polarisation of families into the "work rich" and the "work poor", meaning there has been growth in both the number of families with both adults in work and the number of families with no adults in work (see Dawkins, Gregg and Scutella, 2001 and Burbidge and Sheehan, 2001).

Another important trend in family work patterns has been the increase in average hours of work of some families. Wooden (2001a) and Wooden and Loundes (2001) find that the proportion of the employed workforce working 45 or more hours increased from around 20 percent in 1975 to 28 percent in 1995, and has remained stable at around this level since.

1.2 In this paper

In the next section we review recent findings about trends in Australian income inequality. In Section 3 we describe the data sources and the main limitations of the data. Section 3 also contains discussion of the alternative approaches to the study of income inequality and justification for the approach taken in this paper.

Section 4 explores recent trends in the income distribution utilising unit record data from ABS income surveys spanning the period 1981-2 to 1997-8, and examines the impact of government policy in the form of the provision of transfer payments and levying of income taxation. In particular, the impact of the tax and transfer system on changes to the income distribution is achieved by comparing the changes for disposable income with those for gross (before taxes) income and those for private (before taxes and transfers) income. We report results for the distributions of private, gross and disposable income using graphical representations of the income distributions as well as a variety of statistical measures. We also report trends in the income distribution separately for each of four income unit types: single persons, couples with no dependents, sole parents and couples with dependent children.

Based on the results obtained in Section 4, our focus in Section 5 turns to identification of the sources of changes in the distribution of private (or market) income. Specifically, we attempt to decompose changes in the distribution of private income into those attributable to changes in the income unit composition, the distribution of labour force status and the distribution of demographic characteristics. This is undertaken by adapting a semiparametric method for decomposing distributional changes developed by DiNardo, Fortin and Lemieux (DFL) (1996). Section 6 concludes.

2. Recent findings on inequality in Australia

General comments

The most recent literature on inequality in Australia has tended to extend investigation to developments in expenditure/consumption inequality as well as income inequality. Arguments in support of using consumption expenditure rather than income as a proxy for well being are based on: i) the fact that expenditure is less subject to short term fluctuations, i.e. most households are capable of borrowing/saving to smooth out movements in transitory income over time; ii) that utility is typically defined over consumption rather than income and that resources consumed over a given period are not necessarily equal to received income. Furthermore, income data is sometimes considered unreliable for use in welfare based distributional comparisons because of apparent income concealment for the purpose of tax evasion etc. Nevertheless income data is much easier to gather and policy analysts are interested in the distribution of both income and total household expenditure.

The trend in income inequality in Australia is reasonably well established. Economists broadly agree that income inequality rose through the 1980s and up until the mid 1990s, with some evidence (e.g. Harding and Greenwell (2001)) suggesting that the trend has continued since then. The path of expenditure inequality meanwhile has been found to contrast with that of income inequality, though findings on developments in consumption/expenditure have conflicted somewhat. Overall, however, reports have agreed that income inequality has generally risen at a faster rate than consumption inequality, if the latter has in fact risen at all. Indeed, the somewhat limited studies to date essentially concur that current expenditure inequality fell between 1984 and 1993-94. These findings appear to suggest that government tax and transfers as well as income smoothing have helped to mitigate the impact of rising earnings inequality.

Other elements highlighted by the various findings on inequality include:

- Within-group inequality dominates differences in incomes between groups as the source of income inequality (where groups are defined by socio-economic characteristics); and
- Incomes at the top end of the distribution have grown at a significantly greater rate than at both the middle and lower end of the spectrum.

Disparities in results between studies are largely explained by methodological decisions made prior to analysis that have inevitably influenced the overall results. For example: the study by Barrett, Crossley and Worswick (2000) excludes the top and bottom 3% of observations, plus all households with a head younger than 25 or older than 49 years of age; and Blacklow and Ray (2001) use different equivalence scales and rank households rather than individuals, claiming that it cannot be assumed that 'resources are equally shared within the household'. Additionally, results are somewhat sensitive to the data source examined, with Harding and Greenwell (2001) finding that the ABS Income Distribution (IDS) and Household Expenditure (HES) surveys offered somewhat contrasting pictures of income inequality over time.

Recent studies in Australia

Borland and Kennedy (1998) report evidence of growing earnings inequality in Australia over the 1980s and 1990s using the ABS IDS surveys between 1982 and 1994-5. Decomposition of the sources of changes in overall earnings inequality suggested that the growth in earnings inequality stemmed largely from increases in within group inequality rather than between group inequality (where groups are defined by socio-economic characteristics).

A specific finding of their analysis is that the increase in earnings inequality between 1982 and 1995 for a sample composed of individuals aged 15-64 years was substantially lower than that for a sample of individuals aged 25-59 years. They conclude that the differing trends suggest significant changes in the age composition of the workforce due to an increase in school retention rates. Additionally, their research highlighted differences in income inequality between workforce groups, finding increases in earnings inequality for employees in the private sector, but not the public sector, and that inequality was confined largely to particular areas of industry.

Barrett, Crossley and Worswick (2000) consider income and expenditure inequality in Australia, studying trends in the HES between 1975 and 1993. They find that consumption was far more equal than income, although their research identified significant increases in both over the period examined. The authors also illustrate the role that taxes and transfers played in redistributing income. They look at three measures of inequality: private income (gross income minus government transfers and benefits); net income (more usually known as disposable income and equal to private income plus government transfers and benefits and minus income taxes) and consumption inequality. Comparing the Lorenz curves for these three measures, they find net income to be much more equal than private income, but consumption to be still more equal. They conclude that government transfers and taxes helped mitigate the incidence of rising inequality, and income smoothing by households further reduced inequality.

Analysis of the data also showed real incomes rose at the top of the income distribution, remained stable in the middle and fell at the bottom with real income losses particularly concentrated between the 10^{th} and 25^{th} percentile suggesting a possible growing incidence of working poor. Meanwhile, the very bottom of the distribution showed rising real consumption levels over the data period, implying growing dissaving.

Blacklow and Ray (2000) extended Barrett, Crossley and Worswick (2000) to include multiple family households consisting of unrelated young adults and others. In their analysis, they also

include durables expenditure and examine the impact of changing equivalence scale specifications on inequality magnitudes and on their movements over time. Overall, Blacklow and Ray (2000) agree that income inequality increased over the period. However, they find that, over the same period, expenditure inequality either fell sharply or retained a comparatively flat trajectory. Sensitivity analysis nevertheless revealed that the results were quite sensitive to the equivalence scale used.

Decomposition analysis undertaken by Blacklow and Ray (2000) found that the picture of rising income inequality and decreasing expenditure inequality held across most households types – old age pensioners and single parent families excepted. Again within-group inequality dominated differences in incomes between groups.

A study by Harding and Greenwell (2001) extended the period of study, making use of the HES for 1998-99. Findings in the study agree with those from previous papers. Results suggest that income inequality rose through the 1990s, while expenditure inequality remained stable – the latter reinforcing the findings of Blacklow and Ray (2000). The HES show (most notably through the latter half of the nineties):

- Very marked increases in incomes at the top end of the distribution;
- Marked increases in incomes at the middle; and
- Stable incomes at the 10th percentile of income distribution, but falling incomes at the 5th percentile.

However, these findings are not entirely backed up by the IDS, which suggest that income inequality remained unchanged through 1994-5 to 1997-8. There is some consistency between the surveys in that the relative income share of both the middle and bottom has fallen, while the income share of the top 10% has increased.

Overall, the study by Harding and Greenwell (2001) found strong evidence of a rise in income inequality between the late 1980s and mid-1990s, with some evidence to suggest that the increase has continued since then. This increase was driven by a decline in the income share of the bottom 10% of Australians, a marginal decline in income of the middle quintile and an increase in the income share of the top 10%. In broad agreement with findings by Blacklow and Ray (2000), the study also found that expenditure on current goods and services has not altered significantly over the past 10 to 15 years. However, when investigating inequality of all expenditures (including 'savings' - expenditure on investment property, superannuation, etc.), results suggest an increase between 1984 and 1988-9, perhaps followed by a fall between 1988-9 and 1998-9.

The results of Harding and Greenwell (2001) suggest a consistent relationship between spending and income. The only notable area of change in the last 15 years has been in the spending to income ratio of the bottom decile. The authors' analysis of this suggests that there has been a significant shift in the composition of the bottom decile, with retired households and working poor without dependents households moving in, and income support-recipient families with dependents moving out. The study notes a possibility that these new entrant groups may have greater accessibility to credit/savings, and as such are more able to smooth income over time, which could help explain the dramatic shift in the spending/income ratio.[3]

Using IDS surveys for 1986, 1990 and 1996, Pappas (2001) found that market income became less equal across families, with income units in the top half of the income distribution receiving

[3] In the bottom decile reported spending is 2.3 times reported income. There are strong doubts about the validity of the income data with suggestions that there may be considerable under-reporting of income.

significantly larger proportional increases in market income between 1986 and 1996 – concurring with previous findings. As with previous researchers, he concludes that the tax/transfer system has offset the growth in income inequality and appears to be targeting those families most in need.

Pappas (2001) decomposed inequality in an attempt to cast light on the factors driving the increase. He found a significant increase in the contribution of wages and a decrease in the contribution of investment to market income inequality over time. This trend was particularly pronounced in single and couple (without dependents) income units. Further decomposition showed that these changes were influenced by education, and that wage inequality was decreasingly correlated with age and sex.

3. Data

As the previous discussion has suggested, the study of trends in income distributions require from the researcher choices concerning a wide range of data-related issues. These include the most appropriate choice of the target population, the sample to be examined, the unit of observation (for example, person, income unit, family or household) and the definition of income to be used. In addition, the researcher must be mindful of the availability of explanatory variables, the extent of their comparability over time, and the degree of confidence in the individual records. The range of choices includes the following:

- Whether to examine income or expenditure. Related to this is the choice of data source, with the income surveys containing information about income, and expenditure surveys containing information about both income and expenditure (but often attenuated);
- The definition of income (or expenditure) to be applied. Several different definitions of income are possible: gross, private or disposable income, actual or equivalent income (and if equivalent, the equivalence scale to use), before or after housing costs, etc.;
- The "observational unit" for the analysis (individual, income unit, family, household);
- The criteria for selection into the sample (i.e. who we examine, who we exclude);
- The distributional measures to be examined (e.g. Gini coefficient, Theil coefficient, coefficient of variation, percentile log differences, whole densities etc);
- The sub-groups of the population to be examined (e.g. examine separately groups defined by gender, age, family type, etc.);
- For decomposition analysis, the characteristics on which we condition as sources of distributional change (i.e. we can isolate the effect of changes to many separate sources); and
- The decomposition method to be used. There are a number of alternative decomposition methods available.

3.1 The data source

Several data sources are available for Australia that are potentially suited to the study of income or expenditure inequality. Two sources commonly used by researchers for such studies are the ABS income surveys and expenditure surveys (which we refer to as the IDS and the HES,

respectively). The ABS has made available "confidentialised" unit record files for seven of the eight IDS, spanning the period 1981-2 to 1997-8, and all five of the HES, spanning the period 1975-6 to 1998-9. Unfortunately, for both series of surveys there have been significant changes over time in the sampling frame, in the questions asked and in the variables recorded for respondents in the surveys, creating substantial problems for comparability of the surveys over time. For the most part, however, these problems are not insurmountable, although they do impose significant constraints on the decomposition analysis that is feasible in Section 5.

The HES contain information about both income (at the personal and household level) and expenditure (at the household level), while the IDS contain information only about income (at the personal and income unit level, with the ability to also infer family and household income). The IDS therefore appear preferable to the HES if we are to examine income, since we can identify income for a wider range of units (personal, income unit, family and household) than is possible with the HES (personal and household only). However, the IDS are only a viable option if we are to focus on income, such that the HES need to be used if we are to study expenditure inequality. It follows that, in deciding on the data source to be used, the relative merits of income-based versus expenditure-based studies of inequality need to be considered.

The fundamental motivation for the study of income or expenditure distributions is interest in the distribution of *consumption*, in turn motivated by the view that an individual's level of consumption is an important contributor to the individual's welfare. A focus on income is justified on the grounds that, at least in the long run, an individual's income places an upper bound on consumption possibilities. The qualifier *in the long run* is, however, an important one, since the ability for individuals to intertemporally smooth consumption implies income over a limited time frame may provide a poor measure of the consumption possibilities of the individual over that time period. For example, it has been argued that an increase in weekly income inequality over time may not translate into an increase in consumption inequality, but rather reflect an increase in variability of weekly income for each individual. That is, it may reflect an increase in transitory income inequality, with there being no increase in permanent income inequality. This has in part motivated the study of expenditure inequality, on the basis that expenditure is likely to have a closer relationship to consumption than income (for example, Barrett, Crossley and Worswick (2000)).

However, expenditure-based measures do suffer from failings that income-based measures do not. In particular, increases in income inequality may not translate to increases in expenditure inequality for reasons that are unrelated to greater inequality in transitory (as opposed to permanent) income. The essence of the argument is that low income persons may borrow more/save less if income falls, and high income persons may save more if income rises, for example to facilitate earlier retirement, higher consumption in retirement or increased bequests. These effects on lifetime consumption inequality (including consumption of leisure) are not captured by expenditure-based measures, yet seem relevant to the social welfare implications from which stems our interest in income inequality. Thus, the potential for consumption and income in a given period to differ, which gave rise to criticism of income measures, in fact also causes expenditure based measures to be inadequate.[4]

It therefore doesn't follow, on the criteria of correspondence to (lifetime) consumption, that expenditure-based measures of inequality are necessarily better than income based ones. Indeed, the best compromise would seem to be to use income measured over a reasonable long time

[4] Of course, an increase in dispersion in wage rates may induce contemporaneous labour supply responses which attenuate increases in both income and expenditure inequality, but this is a failure of *both* measures to fail to account for leisure consumption/income.

frame. For example, changes in the extent of transitory income fluctuations are likely to be relatively unimportant for annual income. Expenditure-based measures, it should be noted, do retain some appeal, however, because of the apparent unreliability of income information in most data sources. Non-reporting of income is a major problem for all income surveys, and even among those reporting income, individuals may misreport income for reasons such as concerns for privacy and perceptions that reported income information may be used by government authorities to determine tax obligations and welfare entitlements.

Notwithstanding this concern, in this study we focus on income-based measures, and analysis has been conducted on all seven IDS. The ABS describe the seven surveys by a reporting year, but information gathered may be for a recent financial year and for current information at the time the survey is undertaken. The 1982 survey was undertaken over a two-month period in the fourth quarter of the year and gathered current information for 1982 and annual information for the 1981-2 financial year. The 1986 and 1990 surveys were also undertaken over two months in the fourth quarter in 1986 and 1990, respectively, and gathered current information for 1986 and 1990 and annual information for 1985-6 and 1989-90. The 1994-5 survey was undertaken over the financial year and gathered current information for 1994-5, but annual information for 1993-4. Similarly the 1995-6, 1996-7 and 1997-8 surveys were undertaken for the whole of a financial year and gathered current information for 1995-6, 1996-7 and 1997-8 respectively, and annual information for 1994-5, 1995-6 and 1996-7.

Harding and Greenwell (2001) argue that only the 1990, 1994-5, 1995-6 and 1997-8 surveys are usable (the 1982 and 1996-7 surveys appear "unreliable", and the 1986 survey does not report imputed weekly income tax payable). Harding and Greenwell (2001) also reweight observations in the 1990 survey. Not having access to the reweighting scheme used by Harding and Greenwell, we adopt the ABS population weights for all surveys, while we have decided to persist with all the surveys despite the reservations expressed by Harding and Greenwell (2001).

3.2 The target population

The target population comprises all persons over the age of 15. Most studies of income in Australia have, however, found it necessary or desirable to impose various restrictions on the sample examined. Studies have variously excluded those whose income unit, family or household:

- contains a self employed person;
- contains only one person and that person is under 21 years of age;
- has no income, negative income, and/or income outside some range defined by either dollar amount thresholds (for example, Hyslop and Mare (2001) censor at $2400 and $268,000 in their study of actual (as opposed to per-member equivalent) household income in New Zealand[5]) or percentile rank in the income distributions (for example, Barrett, Crossley and Worswick (2000) exclude observations outside the 3-97 percentile range);

[5] Hyslop and Mare (2001) present an analysis of changes in the distribution of gross household income in New Zealand over the period 1983 to 1998, decomposing trends in inequality into effects of changes in pension rates, household socio-demographic attributes and employment outcomes. They find that changes in household structure (particularly the declining proportion of two-parent families), attributes, and employment outcomes each contribute to the observed increase in inequality, while the changes in returns are estimated to reduce the level of inequality. Collectively these factors account for about fifty percent of the observed increase, depending on the measure of inequality used. The results confirm other research findings that the changes were concentrated during the late 1980s.

- is not headed by a person in a specified age range (for example, this range is 25-59 years in Barrett, Crossley and Worswick (2000)); and
- is not a household that comprises only one family (Barrett, Crossley and Worswick (2000)).

In this study, we exclude only those for whom income information is missing and those with incomes outside a specified percentile range.[6]

Outliers

Decisions on sample restrictions *based on reported income* are part of a more general decision process regarding what to do with observations with extreme values for income, or that are missing income information altogether. Those with missing income are generally dropped in all studies, but the treatment of those with very low or very high reported incomes varies a great deal across studies. Extreme observations are likely to reflect measurement error, and it is generally desirable to minimise the effects of measurement error on inequality estimates. One option is to recode very low incomes to some arbitrary level so as not to drop observations, motivated by evidence that individuals reporting non-credible low incomes do indeed have very low incomes (although not as low as reported). This approach averts the information loss associated with dropping observations outside a pre-specified income range, but involves "making up" results to some extent. A better compromise would seem to be to drop observations with incomes outside a pre-specified percentile range. We then know that we are examining the middle x percent of the income distribution, and therefore how to interpret the results. While not informing us about the entire distribution, this approach does not suffer from the problem of "making up" any of the results, nor the problem of changing the distribution under study across different samples (e.g. time periods) that may occur with the dollar thresholds approach. That is, excluding observations on the basis that income is below some lower threshold or above some upper threshold may result in variation across samples in the proportion of observations that are excluded. For example, in some years, we may be examining 95% of the distribution, and in other years, 90% of the distribution.

This point is being laboured because inequality estimates are in fact quite sensitive to the approach adopted. It is therefore important to establish a sound basis for the approach taken. Consistent with the above reasoning, individuals with income unit income outside the 3-97 percentile range are excluded from the analysis. We should note, therefore, that the true extent of income inequality will most likely be understated in this study.[7]

3.3 The unit of observation

An issue concerns the appropriate "observational unit" and the appropriate associated income variable. There is no consensus in the literature on this issue. Studies have variously employed as the observational unit the individual, the income unit, the family or the household, and as the income variable the disposable, gross or private income of the individual, income unit, family or household. The choice of observational unit determines the population for which the distribution of income is examined. For example, taking the household as the observational unit means the distribution of income across households is being examined (although if each household is

[6] However, sensitivity tests (not reported) have in fact been conducted with alternative sample exclusions.

[7] Note that this sample restriction applies to total income only, so we are looking at the same sample for all analysis. For example, when looking at private income, sample restrictions are still on the basis of total income.

weighted according to the number of members of the household, then the observational unit is in fact the individual). The choice of income variable does not need to match the observational unit, in the sense that it is valid to examine the income for a larger unit than is the unit of observation. For example, it is common to examine the income of the income unit to which an individual belongs. This is in fact the approach taken in this paper: we treat the individual as the observational unit and examine the income of the income unit to which that individual belongs. Consequently, we are examining the distribution of income over the population of individuals aged over 15 years.[8] This approach accords equal weight to each individual in the population over the age of 15 years, while ascribing to the individual the total income to which that individual is likely to have (at least partial) access.[9] This overcomes the problem of finding a large number of individuals have no income, which would occur if personal income was the income variable, while it still gives equal weight to each person in the population over the age of 15 years. The reason for not including individuals aged under 15 years is not from a lack of concern for the welfare of children, but rather because our primary interest is in the population who could potentially work or take other actions to influence the income unit income distribution (such as choosing who to live with). This is not an uncontroversial sample selection restriction.

3.4 Income

Annual or weekly income

Descriptive statistics are presented for both annual and weekly income. As discussed earlier, annual income is probably preferable to weekly income, in order to reduce the impact of changes in the extent and distribution of transitory fluctuations in income. However, an issue that arises for the decomposition analysis in Section 5 is that it is more difficult to decompose changes in annual income measures, because most characteristics of interest (for their effects on the income distribution) are only known for the current week. This includes labour force status, employment status, hours worked and income unit type. It is possible to use current week attributes as conditioning variables for annual income, but it is not clear how to interpret results based on annual income and current week attributes. Consequently, all decomposition analysis is done for weekly income. However, descriptive statistics are nonetheless produced for annual income to aid interpretation of the results for weekly income. In particular, changes to transitory income fluctuations will be revealed by changes in the disparity between annual and weekly estimates, which we can factor into the interpretation of weekly estimates.

"Equivalising"

An issue arising from the choice of *income unit* income as the income variable is that of whether adjustments should be made for the number (and ages) of persons dependent on that income (i.e. comprising the income unit). That is, it needs to be decided whether to use what is termed in the literature an "equivalence scale", and if so, what scale to use. A common rule of thumb scale is to divide income unit income by the square root of the number of members of the income unit, the motivation being there are economies of scale in family or income unit production. Another equivalence scale is calculated by dividing income by one plus 0.6 for each person over the age of

[8] Note that this approach is not possible for the HES, which reports expenditure at the household level, and income at both the household and individual (but not income unit) level. We may be able to construct income units from the person-level information available, and hence income unit income, but we will not be able to identify income-unit expenditure (because it is not recorded at the person level).

[9] Note therefore, that the *person* weights (supplied by the ABS) are used in all the analysis.

15 years and 0.3 for every person under 15 years. Ultimately, the choice of equivalence scale is arbitrary, and the problem arises that the choice of scale is likely to alter inferences on changes to the distribution of income. In this study, it was decided to not use any equivalence scale, and instead allow changes in the income unit composition over time be an explanatory factor in decomposing sources of change in the income distribution. This approach is to some extent consistent with the approach of Hyslop and Mare (2001) (although their study is of household gross income, with no adjustments for household size and, more importantly, no weighting by household size, so that each household, rather than each individual, receives equal weight). However, although the primary focus is on actual income unit income, a limited number of results are presented, as a sensitivity test, using the equivalence scale in which income is divided by the square root of the number of members of the income unit.[10]

The income variable

As discussed, to ascertain the relative roles of income taxes, transfer payments and changes to market income in producing changes in the distribution of income, we report results for three different income concepts: disposable (after taxes and transfers), gross (before taxes and after transfers) and private (before taxes and transfers) income.

Several issues associated with the construction of the income variables used warrant mention. First, the income unit income information has been derived from person record information, with the income unit income assumed equal to the sum of the incomes of the individuals who comprise that income unit. This approach has been taken for two reasons. First, the 1982 survey does not report income unit income, and to be consistent across surveys the same approach should be adopted for each survey. Second, income unit income is missing for a significant number of individuals. It is unclear why this is the case, but it may be related to changes in income unit composition over time. This implies that the income variable used in this study is "the income received in the relevant period by the income unit to which the individual *currently* belongs, irrespective of whether the individual belonged to the income unit in the period over which income is being measured." This is particularly important to be aware of when interpreting results for annual income, since some individuals will not have been in the current income unit in the previous financial year.

Second, the 1982 survey does not allow for negative business income, investment income or other income. To ensure consistency across all surveys, business, investment and other income that is negative has been set equal to zero in all the surveys. This provides a further reason (in addition to the restriction to those with incomes in the 3-97 percentile range) why actual income inequality will be greater than measured.

Third, for disposable income, a problem is that the 1982 survey does not report imputed income tax payable for either annual or weekly income, while the 1986 survey does not report imputed weekly income tax payable. We have imputed annual and weekly income tax for the 1982 survey based on estimated net income and the income tax rates, deductions and rebates in place in the relevant period (1981-2 for annual income, 1982-3 for weekly income).[11]

Finally, all income variables have been converted to fourth quarter, 2000 prices using the ABS Consumer Price Index.

---

[10] Income inequality estimates were also obtained using several alternative equivalence scales, with no significant differences in results.

[11] At this stage, no attempt has been made to impute weekly income tax in 1986, since the primary focus is on the changes in the income distribution over the entire sample period (1982 to 1997-8).

4. Trends in income inequality in Australia

Examination of the income distribution is undertaken using both kernel density estimates of log income and statistical measures of central tendency and dispersion of the income distribution. Kernel density estimation techniques permit examination of the entire income distribution, and are particularly useful for providing an overview of the distribution and the nature of changes over time. They essentially permit us to draw a histogram of the income distribution.[12] This provides a useful visual representation of the distribution of income and how it has changed over the sample period.

For the statistical measures presented, the mean and median are employed as measures of central tendency, while the Gini coefficient, Theil coefficient, coefficient of variation and three percentile log differences are used to describe the extent of dispersion or inequality in the distribution of income. Percentile log differences also provide information about the extent of inequality at different locations in the distribution. For example, the difference in the log of the income at the 90th percentile from the log of median income, denoted P90/50, provides information about the extent of dispersion in the upper tail of the distribution.[13]

In the first part of this section we present results for the income distribution of all persons. We start with a focus on private, gross and disposable *annual* income. Differences between private and gross income distributions reflect the effects of government transfers (pensions, benefits and allowances). Differences between gross and disposable income distributions reflect the effects of income taxation. Means and medians of each distribution are shown first. Densities for three years, 1981-2, 1989-90 and 1996-7, for private, gross and disposable income are presented to indicate trends, while to compare distributions across all survey years we use summary measures. The robustness of the results to alternative income measures are then assessed by examining densities and summary measures for both weekly income and "equivalised" annual income. The weekly income results are also important because the decomposition analysis in Section 5 is conducted on weekly income.

The second part of this section examines trends in the distribution of income within and between income unit types. Specifically, we identify four types of income units – couples, couples with dependents, singles and singles with dependents – and measure trends in inequality between the groups and within the groups.[14]

[12] Appendix 1 contains details on the kernel density estimation methods used.

[13] Note, however, that percentile log differences are not defined for private income in some cases because of the presence of persons with zero income unit private income. For example, the P50/10 percentile log difference is not defined when more than 10 percent of individuals have a zero income.

[14] In the decomposition analysis in Section 5, 12 groups of income units are distinguished; couples and singles with dependents are further disaggregated according to the number of dependents; and singles are separated according to gender. The restriction to four groups for this part of the analysis is to render interpretation of the roles of changes between and within income unit types more straightforward.

4.1 Trends in the aggregate distribution

Mean and median income

Table 4.1 shows trends in real "unequivalised" (actual) income unit annual income for seven years from 1981-2 to 1996-7. The table shows private, gross and disposable income.

A number of trends can be identified from the table:

- Mean private, gross and disposable incomes rise over the period;
- Median income is always lower than mean income, indicating a skewed distribution;
- Median private and gross incomes fall over the period;
- Transfers have become more important over time, evidenced by growth in the mean level of transfers; and
- Mean income taxes are fairly constant over the period.

The divergence between the mean and the median for all three measures of income suggests increasing inequality. This is particularly so for private income, and least so for disposable income, suggesting that the effect of taxes and transfers has been to mitigate the increase in inequality due to increased inequality in private income.

Table 4.1 Trends in mean and median income unit income 1981-2 to 1996-7, annual income (Dec 2000 prices)

	1981-2	1985-6	1989-90	1993-4	1994-5	1995-6	1996-7
Private income							
Mean	36002	36769	38897	36804	37517	36752	37980
Median	34612	34543	35661	33160	32617	32735	32776
Gross income							
Mean	39603	40604	42426	41021	41835	41226	42616
Median	36085	36101	37023	35809	35411	35075	35811
Disposable income							
Mean	32256	32864	33978	33489	34016	33615	34556
Median	30217	29751	30807	30469	30014	29936	30761
Differences in means							
Transfers	3601	3835	3529	4217	4318	4474	4636
Tax	7347	7740	8448	7532	7819	7611	8060

Distributions of annual income

Figures 4.1, 4.2 and 4.3 present the kernel density estimates of the distributions of private, gross and disposable annual income for the 1982, 1990 and 1997-8 surveys. The densities presented for private income, although estimated over the entire population, do not provide a complete picture

of the income distribution. This is because of a large "spike" at zero, which cannot be displayed in a graph of the distribution of *log* income. Consequently, the proportion of the population with no income unit private income is also reported to provide an indication of changes in the magnitude of the spike at zero.

The shape of the curve indicates the level of inequality. In simple terms, a narrow, high density indicates a more equal society, and a broad, low density indicates a less equal society. Thus a change in the shape of the curve, basically a broadening and flattening of the density, indicates an increase in inequality.

Figure 4.1 Distribution of log annual income unit private income - December 2000 prices

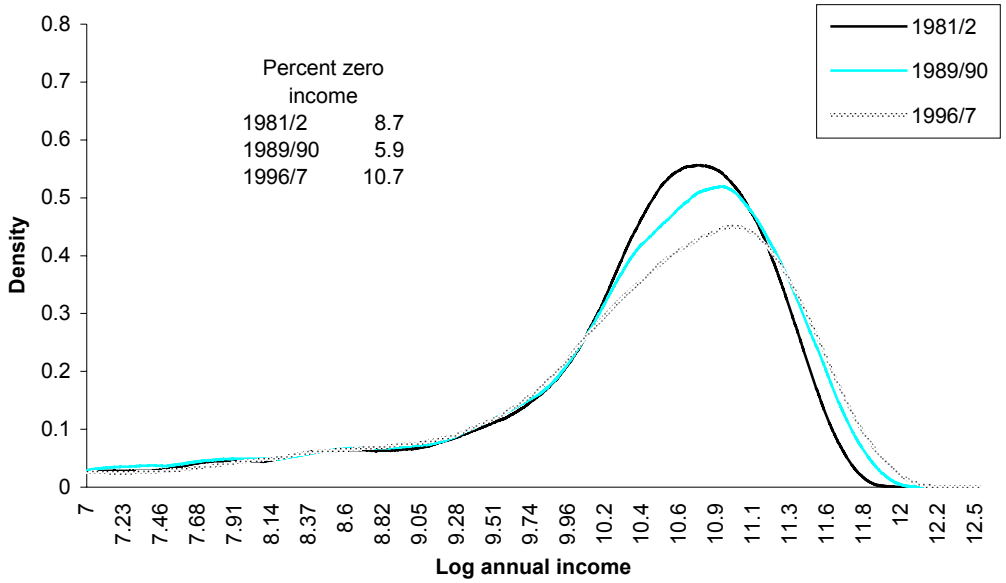


Figure 4.1 shows the distribution of the log of annual *private* income at December 2000 prices for 1981-2, 1989-90 and 1996-7. In 1981-2 there is a modal peak at about log(10.7) or $44,355 per year. The vertical axis indicates the density or the proportion of individuals at any value of log annual income (per unit of log annual income). The area under a section of the curve measures the proportion of individuals with income unit income in the range of incomes covered by that area. For instance, in 1981-2 the area under the curve between log(10.4) (an income unit private income of $33,000) and log(11.05) (an income unit private income of about $63,000) includes about 34 percent of all individuals ((11.05-10.4)*0.53).

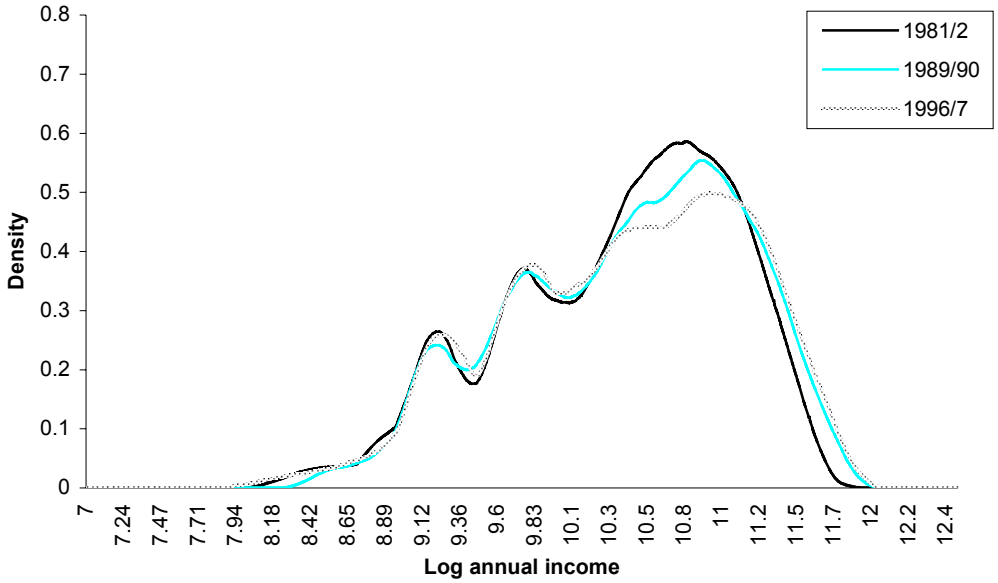
By 1989-90 the modal peak had fallen and shifted to the right. The peak is now at about 10.9 or $54,000 and there are increased numbers of recipients receiving this and higher incomes. By 1996-7 the peak had fallen further and shifted further to the right so that modal income is now about 11.0 or an income of $60,000. Again, the curve is to the right of the previous two, indicating increased numbers of individuals in income units receiving higher incomes.

As mentioned above, the densities for private income cannot display the proportion of individuals in income units with zero private income, and so this information is displayed separately in

Figure 4.1. It shows that, between 1981-2 and 1996-7, the proportion of individuals with no private income increased from 8.7 to 10.7 percent.[15] The figure implies two effects on the distribution of private income between 1981-2 and 1996-7: an increase in inequality brought about by higher income people receiving even higher incomes, shown by the 1996-7 curve being to the right of the 1981-2 curve; and an increase in the size of the spike at zero private income, indicating an increase in inequality most likely caused by a falling rate of employment.

Figure 4.2 shows the distribution of the log of annual *gross* income at December 2000 prices for 1981-2, 1989-90 and 1996-7. There are now three peaks. The modal income for 1981-2 is at log(10.8) or $49,020 per year. There is a secondary peak at log(9.81) or $18,214 per year. This is around the level of social security payments for couples dependent on government benefits and pensions. The third peak at log(9.2) is at a disposable income of around $10,000 per year, or about the income of a single income support recipient. As with the figure for private income, going from 1981-2 to 1989-90 and then to 1996-7, the modal peak shifts to the right and the proportion of people at the peak falls. On the high income side (to the right of the modal point) the 1989-90 curve is to the right of the 1981-2 curve and the 1996-7 curve is to the right of the 1989-90 curve. To the left of the modal point the curves seem to be coincident and the shape of the distribution does not indicate a clear increase in inequality.

Figure 4.2 Distribution of log annual income unit gross income - December 2000 prices

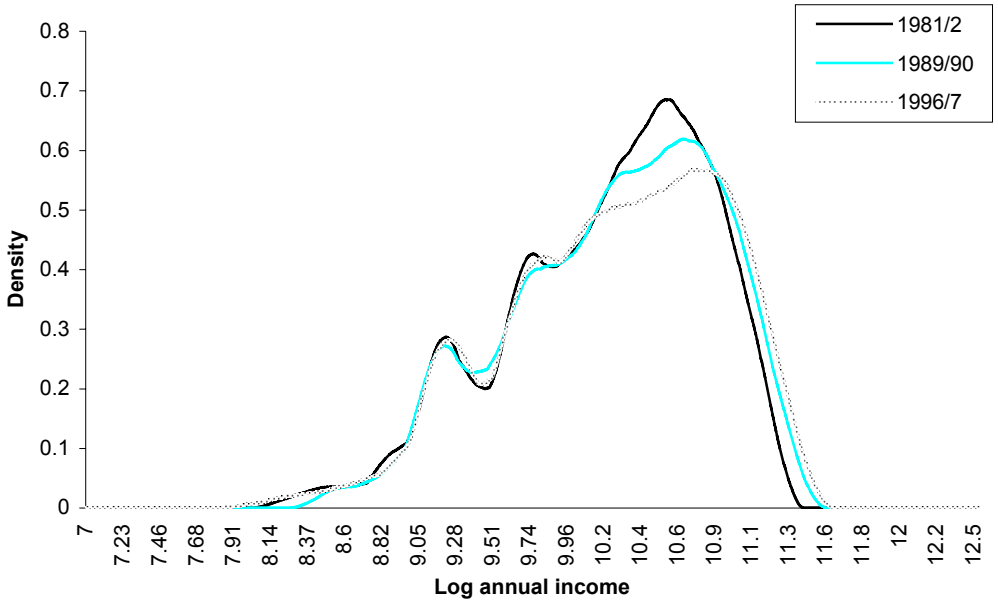


[15] Figure 4.1 also shows that the proportion of income units with no private income in 1989-90 was 5.9 percent. This apparently anomalous result is potentially explained by the effects of the business cycle, which had economic conditions peaking and unemployment in a trough in 1989-90. Thus, the proportion of income units with no private income fell to a low, even though the longer term trend may have been upward.

Figure 4.3 shows the distribution of the log of annual *disposable* income at December 2000 prices for 1981-2, 1989-90 and 1996-7. The three peaks identified in the gross income distribution remain, but the effect of income taxation is to reduce the width of the peaks. The modal peak has been shifted to the left in comparison with gross income. In 1981-2 it is at log(10.6) or $40,135 per year. The second and third peaks remain at their values for gross income (log(9.81) or $18,214 per year, and log(9.2) or around $10,000 per year). As for gross incomes, these peaks roughly correspond with the level of income support payments for single and couple income units.

As with the previous figures, going from 1981-2 to 1989-90 and then to 1996-7, the modal peak shifts to the right and the number of people at the peak falls. On the high income side (to the right of the modal point) the 1989-90 curve is to the right of the 1981-2 curve and the 1996-7 curve is to the right of the 1989-90 curve. However, left of the modal point the curves cross at a number of points and no clear pattern emerges.[16] In summary, the shape of the distribution does not indicate a clear increase in inequality.

Figure 4.3 Distribution of log annual income unit disposable income - December 2000 prices



Summary indexes

The graphs shown above are useful ways of depicting trends in a small number of distributions but are cumbersome when we wish to compare many distributions, and do not permit us to quantify distributional changes. For these purposes, we use a number of summary indexes. To

[16] Perhaps what is happening here is a shift of the whole curve to the right indicating an increase in all incomes but not necessarily an increase in dispersion of those incomes.

illustrate their use we first consider a single year, 1996-7. Table 4.2 reports the mean and median income unit annual income as well as six summary measures of inequality in the distribution.

The first three summary measures are the Gini index, the Theil index, and the coefficient of variation. The Gini index is the most widely used summary index, with an upper limit of one, indicating maximum inequality, and a lower limit of zero, indicating a perfectly equal distribution. The Theil index also has a lower bound of zero at maximum equality, but no upper bound, although in all practical circumstances it is unlikely to exceed one. The coefficient of variation also is bounded at zero where there is perfect equality, but has no upper bound.

The other three measures compare the incomes of income units at the ninetieth and fiftieth percentiles (P90/50), at the fiftieth and tenth percentiles (P50/10) and at the seventy-fifth and twenty-fifth percentiles (P75/25). The P90/50 index measures the difference in the log of income at the ninetieth percentile and the log of income at the fiftieth percentile. A value of one for this index indicates that income at the ninetieth percentile is 100 percent higher than income at the fiftieth percentile, when evaluated at the midpoint of the two incomes. These three indexes provide information about the extent of inequality at different points in the income distribution, namely the upper tail (P90/50), the lower tail (P50/10) and the middle (P72/25) of the distribution. In all cases, higher values correspond to greater inequality.

The table shows all summary measures of inequality are reduced by both government transfers and personal income tax. With the provision of government pensions, benefits and allowances, the Gini index falls from 0.474 to 0.373. When income is taxed the Gini falls further to 0.329.

Table 4.2 Distribution of income unit income 1996-7, annual income (Dec 2000 prices)

	Private	Gross	Disposable
Mean	37,980	42,616	34,556
Median	32,776	35,811	30,761
Gini index	0.474	0.373	0.329
Theil index	0.454	0.221	0.171
Coefficient of variation	0.843	0.672	0.584
P90/50	0.934	0.854	0.729
P50/10	-	1.229	1.087
P75/25	2.013	1.213	1.015

Trends in annual income inequality

Table 4.3 presents summary indexes for all three measures of income for all seven survey years. The top panel shows the results using the Gini index and the lower three panels for the three percentile log difference indexes.

The Gini index for private income increased steadily through the eighties and early nineties but slowly in the latter part of the nineties. The Gini for gross income also increased until the mid-nineties and remained flat over the last four years. The Gini for disposable income increased over the first four years but was virtually constant thereafter.

The P90/50 index provides a measure of inequality in the upper part of the distribution. For private income, it increased until 1994-5, dipped in 1995-6 and returned to its 1994-5 level in 1996-7. The P90/50 index for gross and disposable income follow a similar pattern. The similarity between the pattern for private and gross income is understandable, since this index is unlikely to be much effected by transfer payments.

The P50/10 index provides information on the lower part of the distribution, but is of limited value for the analysis of private income because of the low number of income units with private income in the lowest decile. Indeed, more than 10 percent of individuals had no private income in the last two survey years, such that the P50/10 index is not defined in these two years. The P50/10 index for gross income oscillated over time, but trended down. It also oscillated for disposable income, but was fairly constant through the period. The similarity between the pattern for gross and disposable income is perhaps to be expected, since income taxation is not likely to have a great deal of effect on the bottom half of incomes.

The P75/25 index measures the trend in inequality for the middle part of the distribution. The index for private income rose to 1993-4 with a downward blip in 1989-90, and then remained fairly constant. The trend for gross and disposable income was similar.

Table 4.3 Trends in income unit income inequality, 1981-2 to 1996-7, annual income (Dec 2000 prices)

	1981-2	1985-6	1989-90	1993-4	1994-5	1995-6	1996-7
Gini index							
Private income	0.427	0.445	0.439	0.468	0.472	0.470	0.474
Gross income	0.346	0.358	0.360	0.373	0.377	0.369	0.373
Disposable income	0.306	0.324	0.316	0.331	0.336	0.327	0.329
P90/50							
Private income	0.755	0.786	0.830	0.872	0.928	0.894	0.934
Gross income	0.729	0.755	0.798	0.809	0.854	0.835	0.854
Disposable income	0.636	0.703	0.692	0.704	0.754	0.729	0.729
P50/10							
Private income	5.349	5.254	4.059	10.232	7.938	-	-
Gross income	1.246	1.244	1.201	1.267	1.229	1.195	1.229
Disposable income	1.069	1.083	1.037	1.106	1.071	1.061	1.087
P75/25							
Private income	1.548	1.805	1.530	2.113	1.975	1.958	2.013
Gross income	1.1	1.151	1.137	1.205	1.202	1.171	1.213
Disposable income	0.922	0.968	0.954	1.008	1.019	0.989	1.015

Weekly income

In this section we explore the sensitivity of results to the income period examined by examining *weekly* income, which is particularly important for the interpretation of the decomposition results

obtained in Section 5. Figures 4.4 and 4.5 present densities for weekly private and disposable income in 1982, 1990 and 1997-8, while Table 4.4 reports means and Gini indexes for private, gross and disposable weekly income.

Figure 4.4 Distribution of log weekly income unit private income - December 2000 prices

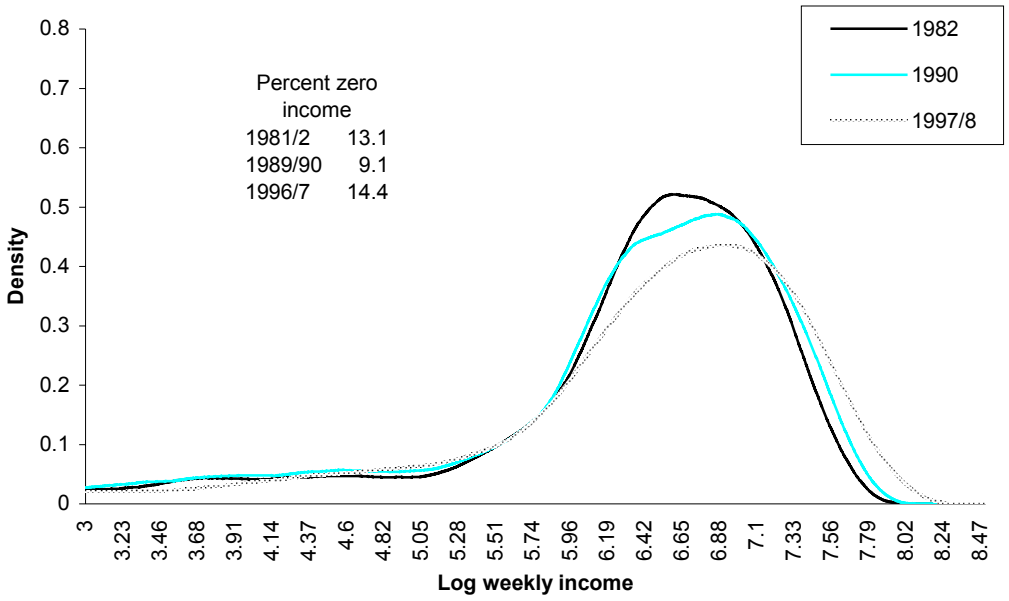
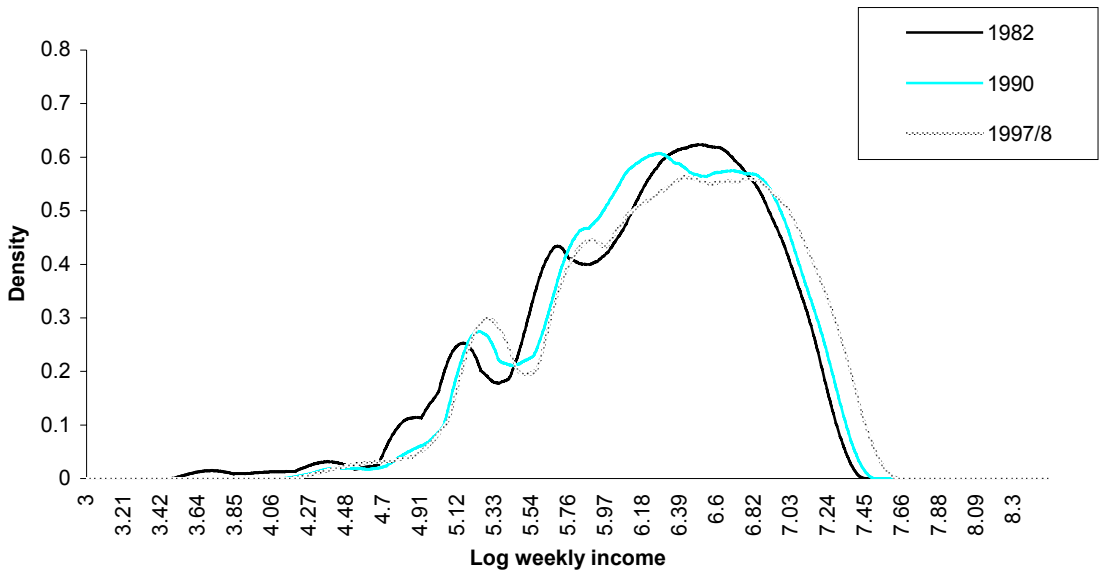


Figure 4.4 shows the same pattern as for Figure 4.1, with the modal peak for private income in 1981-2 falling in 1989-90 and falling further in 1996-7. As for Figure 4.1, the distribution flattens out, with a longer tail at the high end and an increased proportion of income units with zero private income.[17] The distribution shown by Figure 4.5 is likewise similar to Figure 4.3, although a move of the *whole* curve to the right is apparent, as opposed to increases in incomes only at the top end of the distribution. This suggests taxes and transfers have been more important for weekly income than was apparent for annual income in counteracting the increased inequality of private incomes.

[17] As with annual income, the percent of units with zero private income is at its lowest in 1989-90. We attribute this result to business cycle conditions.

Figure 4.5 Distribution of log weekly income unit disposable income - December 2000 prices



This picture is confirmed by the results in Table 4.4, which compares means and Gini indexes for private, gross and disposable weekly income in all survey years. The table shows:

- Real mean private, gross and disposable incomes all increased over the 15 years;
- Growth in mean gross and disposable incomes exceeded that of mean private income;
- Inequality in private weekly income increased over time;
- Inequality in the distribution of both gross and disposable weekly income increased until the mid 1990s, thereafter remaining relatively constant;
- The greatest increase in dispersion over the whole period occurred with private income, followed by gross income and disposable income, in that order.

All of these findings are consistent with those for annual income reported earlier. This is despite the potential for changes in transitory income fluctuations and the fact that there have been changes in the sampling method of the surveys over the period. In the surveys up until 1990, weekly income was constructed from a sample gathered over a two months period, whereas in the surveys since it has been based on a continual (year-round) survey. In undertaking further decompositions in the next section, we report results for weekly income only, since most of the information available on characteristics applies only to the current week. The consistency of the inequality results with those for annual income is therefore important for the greater confidence it permits us in the validity of inferences based on the decomposition of changes in weekly income inequality.

Table 4.4 Trends in the distribution of income unit income, 1982 to 1997-8, weekly income (Dec 2000 prices)

	1982	1986	1990	1994-5	1995-6	1996-7	1997-8
Mean							
Private income	638.03	646.04	680.29	677.9	665.29	678.57	711.75
Gross income	706.34	719.47	754.67	771.28	757.05	775.2	810.02
Disposable income	579.33	-	614.09	631.1	621.6	635.75	662.38
Gini index							
Private income	0.445	0.448	0.447	0.469	0.478	0.477	0.479
Gross income	0.356	0.350	0.352	0.357	0.365	0.360	0.366
Disposable income	0.316	-	0.306	0.315	0.321	0.315	0.321

Equivalised income

Figure 4.6 reports the distribution of equivalised annual disposable income for 1981-2, 1989-90 and 1996-7. The same general trends described above for annual disposable income are noted for this distribution. However the effect of equivalising is to eliminate one of the peaks – in Figure 4.6 there are just two peaks, a modal peak and a second peak corresponding to the equivalised value of social security pensions and benefits.

Figure 4.6 Distribution of log annual equivalent disposable income - Dec 2000 prices

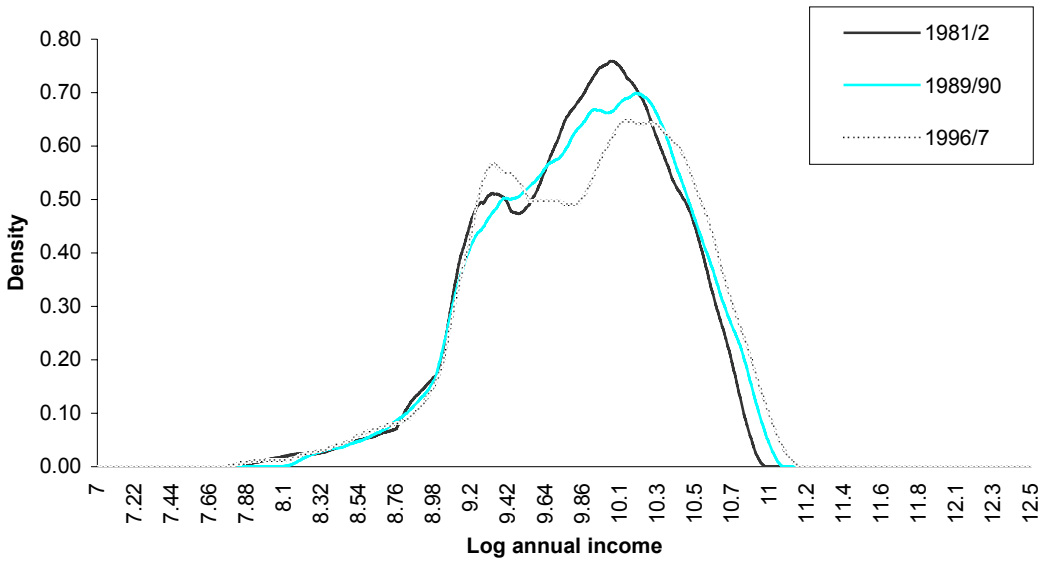


Table 4.5 reports values of the Gini for all seven years and for private and gross income as well as disposable income when income is equivalised. Again we see the same general results previously

noted. For mean equivalised private incomes there is a rise until 1989-90 then a plateauing through to 1996-7. The trend is similar for equivalised gross income, with a faster rise in the period to 1989-90, followed by a slight rise in the period to 1996-7. Mean equivalent disposable income also rises more steeply in the first two periods, but continues to rise thereafter, although at a reduced rate.

Inequality of equivalised annual income also follows the same trends noted for inequality of actual annual income. Inequality of private income rises most, and inequality of disposable income rises least. The consistent results here suggest that findings are not likely to be sensitive to the equivalence scale adopted.[18]

Table 4.5 Trends in the distribution of income unit income, 1982 to 1997-8, equivalised annual income (Dec 2000 prices)

	1981-2	1985-6	1989-90	1993-4	1994-5	1995-6	1996-7
Mean							
Private income	23131	23756	24972	24037	24488	24036	24838
Gross income	25710	26481	27480	26973	27498	27160	28033
Disposable income	21014	21534	22106	22080	22420	22219	22789
Gini index							
Private income	0.412	0.431	0.420	0.453	0.455	0.451	0.458
Gross income	0.315	0.328	0.328	0.345	0.346	0.337	0.344
Disposable income	0.270	0.292	0.279	0.299	0.300	0.290	0.296

4.2 Trends in private and disposable income of income unit types

In this section we trace the fortunes of each of four income unit types, considering changes in both inequality between the income unit types and inequality within each type. Results in Section 4.1 confirmed the similarity in trends for annual and weekly income; consequently, consistent with the decomposition analysis in Section 5, we focus here on weekly incomes. The first part of this section examines between-group inequality by comparing mean incomes. We present the trends in mean private and disposable weekly income separately for couples, couples with dependents, singles and singles with dependents. The second part examines within-group inequality using the Gini index, again for both private and disposable income.

Between-group inequality

Table 4.6 shows the mean real weekly private income of the four income unit types. There are enduring large differences between them. However note that in this table, we take no account of income unit size or composition (income is not equivalised). Couples with dependents earn the

[18] Estimates have in fact been produced employing several alternative equivalence scales. Changes in inequality are remarkably similar to those for actual income irrespective of the equivalence scale used, despite significant changes in the income unit type composition of the population over the sample period.

highest average private income, followed by couples with no dependents, then singles, and singles with dependents earn the lowest average private income.

The table shows large increases in real private incomes for all groups over the 15 year period. However the increases were greatest for couples and couples with dependents. The increases were fairly evenly spread over the period, except for single persons, for whom most improvement occurred during the nineties.

As a proportion of initial income, the greatest gains were made by singles with dependents, who increased their mean real private income by about 19 per cent, followed by couples and couples with dependents, with respective gains of 14.5 and 14 percent. The smallest real increases were earned by singles who registered a 10.5 percent increase.

Table 4.6 Trends in mean income unit private income 1982 to 1997-8, weekly income (Dec 2000 prices), by income unit type

Year	Couples	Couples with dependents	Singles	Singles with dependents
1982, $	654.51	874.74	363.62	271.14
1986, $	649.9	890.03	375.75	285.31
1990, $	701.45	935.65	372.79	299.09
1994-5, $	720.68	937.26	386	302.12
1995-6, $	709.06	927.58	372.53	324.22
1996-7, $	714.07	954.54	393	276.8
1997-8, $	749.6	997.34	401.7	323.5
Changes				
82/90, $	46.94	60.91	9.17	27.95
90/97-8, $	48.15	61.69	28.91	24.41
82/97-8, $	95.09	122.6	38.08	52.36
82/97-8, %	14.5	14.0	10.5	19.3

Table 4.7 shows trends in mean disposable income for the income unit groups. All income unit types make real gains in income on average. The largest gains occur in couples with dependents followed by singles with dependents, couples and singles. The largest proportionate gains are made by singles with dependents followed by couples with dependents, couples and singles.

Comparing Tables 4.6 and 4.7 shows the effects of the tax and transfer system on the mean income of each income unit type. Their effect is to considerably reduce the income of couples with dependents, moderately reduce the income of couples without dependents and singles and to increase significantly the income of singles with dependents. The effect of government intervention, through transfer payments and personal income tax was to moderate the gains of couples, improve the gains of couples with dependents and singles and markedly improve the gains of singles with dependents. Since singles with dependents comprise the second-poorest income unit type, the large increase in their average incomes would suggest a decline in inequality. However, such income units make up a very small proportion of the population. The poorest group of all, singles, and one of the largest groups in terms of size, managed only a 12 percent increase in mean income. This would have acted to increase aggregate income inequality. Overall, the trends in mean incomes suggest an increase in between group inequality.

Table 4.7 Trends in mean income unit disposable income 1982 to 1997-8, weekly income (Dec 2000 prices), by income unit type

Year	Couples	Couples with dependents	Singles	Singles with dependents
1982, $	622.45	746.41	348.68	380.37
1986, $				
1990, $	655.5	791.06	360.39	423.13
1994-5, $	684.7	822.57	376.06	464.19
1995-6, $	676.01	818.18	365.47	463.65
1996-7, $	685.3	841.11	382.63	455.68
1997-8, $	710.95	878.13	389.62	482.03
Changes				
82/90, $	33.05	44.65	11.71	42.76
90/97-8, $	55.45	87.07	29.23	58.9
82/97-8, $	88.5	131.72	40.94	101.66
82/97-8, %	14.1	17.6	11.8	26.5

Within-group inequality

Table 4.8 reports values of the Gini index for private income for individuals in each of the income unit groups, while Table 4.9 reports values for disposable income.

Table 4.8 Trends in within-group weekly private income inequality 1982 to 1997-8, Gini indexes by income unit type

Year	Couples	Couples with dependents	Singles	Singles with dependents
1982	0.472	0.294	0.515	0.663
1986	0.484	0.295	0.503	0.666
1990	0.463	0.303	0.505	0.624
1994-5	0.474	0.331	0.523	0.611
1995-6	0.482	0.345	0.520	0.609
1996-7	0.479	0.339	0.526	0.668
1997-8	0.482	0.343	0.532	0.641

Table 4.8 indicates that the greatest private income inequality within groups is among single parents, followed by singles, couples and couples with dependents. The high level of inequality among singles with dependents to a significant extent reflects the large magnitude of the proportion of this group with no private income. Inequality within groups increased for all income unit types, with the greatest increases among couples with dependents.

Table 4.9 reports estimates of inequality for disposable weekly income. Because the 1986 survey did not report tax payments for weekly income we are not able to calculate disposable income on a comparable basis to other years. In comparison with the results for private income, Table 4.9 shows the marked reduction of inequality when the effect of transfer payments and income taxation are taken into account. Inequality among income unit types is remarkably similar when disposable income is the measure. Although the Gini values are erratic from year to year, overall they indicate little change in within group inequality. The contrast between the results for private income and disposable income again emphasise the important role played by transfers and taxes.

Table 4.9 Trends in within-group weekly disposable income inequality 1982 to 1997-8, Gini indexes by income unit type

Year	Couples	Couples with dependents	Singles	Singles with dependents
1982	0.294	0.222	0.320	0.278
1986	na	na	na	na
1990	0.283	0.217	0.292	0.245
1994-5	0.292	0.221	0.305	0.243
1995-6	0.299	0.229	0.300	0.251
1996-7	0.290	0.221	0.302	0.238
1997-8	0.299	0.222	0.312	0.253

4.3 Summary

In this section we have described trends in income inequality over the period 1981-2 to 1996-7 using information from seven income surveys. Despite the limitations of the surveys, we believe they provide a reasonable picture of developments in the income distribution. We note that many of the limitations are more important to estimates of the *level* of inequality at a point in time than they are to estimates of *changes* in inequality over time. Specifically, to the extent that the limitations apply to all seven surveys, measures of change are not likely to be significantly compromised by the limitations, and it is change in inequality that is the focus of this paper.

The section has indicated that at the level of disposable income, there was a modest increase in inequality over the fifteen years, and most occurred in the eighties and early nineties. At the same time, the real mean income unit incomes of all persons, and of each income unit type, increased over the fifteen year period. Decomposition of the relative roles played by market income, government transfers and income taxation in the changes in the disposable income distribution has shown that there were large increases in inequality at the level of private income, but this was mitigated by the provision of government transfers and by income taxation. The increased income inequality at the private level appears to derive from both greater increases in incomes of high income units and falls in employment in low income units.

Analysis of trends in the distribution by income unit type confirm the aggregate findings. Dispersion has increased for private income, but not for disposable income. In the next section, we aim to uncover the sources of the increased dispersion at the private income level.

5. Decomposing Distributional Changes

In Section 4, we decomposed changes in disposable income into those due to income taxes, those due to transfers, and those due to private income changes. It was found that private income inequality has increased significantly over the sample period, thereby acting to increase disposable income inequality, but to a significant extent offset by the tax and transfer system. We now focus on what has driven these changes in private income.

The primary interest for the decomposition analysis is in the effect on the income distribution of changes to the income unit structure, distribution of employment across income units and the wage distribution, although the effects of other factors are also explored. The specific variables constructed to some extent depend on the data available, and are discussed further below.

The DFL method

The decomposition methods employed are based on the semiparametric conditional density estimation approach developed by DiNardo, Fortin and Lemieux (1996).19 The intuition for this approach is that a density (of any distribution, including income) may be viewed as a weighted average of densities conditional on characteristics, where the weight assigned to each conditional density is proportionate to the share of the population with that set of characteristics. Distributional changes over time can then be viewed as the outcome of two distinct forces: changes in the group weights brought about by changes to the "characteristics composition" of the population; and changes in the conditional densities (i.e. the incomes accruing to members of each group).

Adopting this framework, the separate impact of the two sources of change to the income distribution may therefore be identified by holding constant the weight assigned to each group. For example, the population weights in the 1997-8 survey could be altered such that the distribution of characteristics in this reweighted sample is the same as in the 1982 survey, allowing us to identify the distribution of income that would have prevailed in 1997-8 had the distribution of characteristics remained as they were in 1982. It then follows that the differences between this counterfactual distribution and the distribution actually prevailing in 1997-8 derive solely from changes to the distribution of characteristics between 1982 and 1997-8, evaluated at the incomes (or "prices") associated with each set of characteristics in 1997-8. Alternatively, we could reweight the 1982 sample to have the same distribution of characteristics as in 1997-8, and interpret the differences between the actual distribution in 1982 and this counterfactual distribution as deriving from changes to characteristics over the sample period, but evaluated at 1982 "prices".

The approach can also be adapted to isolating the effects of changes to particular characteristics by varying the characteristics used to reweight the samples. For example, the effects of changes to labour force status may be identified by comparing the results when labour force status is included in the reweighting process with the results obtained when it is excluded.

In practice, the reweighting of the samples is undertaken via estimation (on all observations in both the base and end years) of binary choice models of the probability an observation with given characteristics comes from the base year (see Appendix 2 for details).

[19] Formal description of the DFL method and its implementation is provided in Appendix 2,

Implementation

In order to ascertain the role of changes to characteristics in changes to the income distribution we need to construct variables capturing characteristics. The information available in all seven of the IDS permit us to potentially construct variables based on:

- Income unit type (single or couple, number and ages of dependent children);
- Years of age;
- Highest educational attainment (bachelor's degree or higher, other post-school qualification, no post-school qualifications);
- Labour force status (employed full-time, employed part-time, unemployed, not in the labour force);
- Whether self-employed, an employer or an employee;
- Region of birth (including whether foreign-born);
- State of residence;
- Current usual hours worked per week (although this is categorical information, with the categories varying over surveys);
- Income by source (welfare payments, wages/salary, own business, dividends/rent, superannuation, other);
- Income tax paid (annual and weekly, imputed); and
- Industry, occupation and sector of employment.

An issue that arises in determining the appropriate approach to defining variables to reflect characteristics is that it is the distribution of *income unit* income that is under study. This implies that it is not appropriate to define variables simply in terms of the personal characteristics of the individual; rather, the characteristics of the income unit to which the person belongs should be used. The reason why it is not appropriate to use individual-specific characteristics is that if we hold a set of personal characteristics constant, this will have implications for the income unit income of persons for whom we are not holding characteristics constant. For example, if we accord less weight in 1997-8 to those married females who are working (so as to hold this characteristic of the population constant at the 1982 level), we necessarily should give less weight to males who are married to those females, since the income unit income of these males in part derives from the earnings of the females.

It is not, however, entirely straightforward how to define variables for income unit characteristics. One possibility is to adopt the approach of Hyslop and Mare (2001), who employ variables such as "the proportion of household adults in employment". We take a slightly different approach, however, and define many income unit characteristics in terms of the characteristics of the reference person and the characteristics of the partner of the reference person. For example, variables capturing the educational attainment of both the reference person and the partner of the reference person are used.[20]

The variables created fall into three broad groups of types of income unit characteristics that we attempt to capture:

[20] Note that the partner is always female in all surveys, and the reference person is only female if the income unit is a single woman (with or without dependent children).

1. Income unit type.

Dummy variables for each of 12 income unit types are employed: single male, single female, couple with no dependents, single male with one dependent child, single male with two dependent children, single male with three or more dependent children, single female with one dependent child, single female with two dependent children, single female with three or more dependent children, couple with one dependent child, couple with two dependent children, couple with three or more dependent children.

2. Labour force status:

Dummy variables are created for: employed full-time, employed part-time, unemployed and not in the labour force. For couples, separate dummies are employed for each possible combination of labour force status of the reference person and the partner of the reference person (16 dummies).

3. Demographic characteristics, the variables for which comprise:

Age: 6 dummy variables for the age in years of the reference person in the income unit: 15-24, 25-34, 35-44, 45-54, 55-64, 65+;

Country of birth dummy variables: Foreign-born, Australian-born. For couples, separate dummy variables are employed for each possible combination (4 dummies); and

Educational attainment dummy variables: Degree, other post-school qualification, no post-school qualifications. For couples, separate dummies are employed for each possible combination (9 dummies).[21]

In estimating a binary choice model to derive the reweighting function, we employ a large number of interaction terms in order to allow as much flexibility as possible in the changes in the distribution of characteristics in the population. For example, the effects of an increase in the polarisation of work across income units may not be fully identified without interactions between income unit type and labour force status variables, particularly if there is little change in labour force status in the aggregate. As such, the labour force status, country of birth and highest educational attainment dummies are all fully interacted with the income unit type dummies.

Identification of the separate effects of each of the three groups of characteristics is achieved by sequential elimination from the logit equations of the variables for each group of characteristics. Specifically, the effects of labour force status are identified by comparing results when all variables are included with results when the labour force status variables are excluded. The effects of demographic characteristics are then isolated by additionally eliminating the variables for these characteristics from the logit equation. The effects of income unit compositional changes are derived from the logit equation with only the income unit type variables included.[22]

---

[21] A significant omission for demographic characteristics is a variable for *gender*. This is because the variables used for income unit types differentiate between males and females, such that this characteristic is accounted for by the income unit type variables rather than the variables for demographic characteristics.

[22] Results are potentially sensitive to the order in which variables are removed from the logit equation because of potential correlations between variables measuring different characteristics. For example, an increase in educational attainment may be associated with an increase in labour force participation. If we eliminate variables capturing labour force status first, some of the change in the income distribution attributed to changes in educational attainment may in fact reflect the effects of labour force status changes

To illustrate, consider decomposition of changes in some moment M of the income distribution (e.g. the mean) between the base year and end year. If M_B is the actual value of the moment in the base year, M_E is the value in the end year, M_A is the value in the end year when all observable characteristics are held at base year levels, M_{A-L} is the value in the end year when all observable characteristics except labour force status are held at base year levels and M_I is the value in the end year when only the distribution of income unit types is held at the base year distribution, then we have the following:

Total change: $M_E - M_B$

Effect of changes in the distribution of all characteristics: $M_E - M_A$

Effect of changes in the distribution of labour force status:
$(M_E - M_A) - (M_E - M_{A-L}) = M_{A-L} - M_A$

Effect of changes in the distribution of demographic characteristics:
$(M_E - M_{A-L}) - (M_E - M_I) = M_I - M_{A-L}$

Effect of changes in the distribution of income unit types: $M_E - M_I$

The change in the moment M which is not explicable by changes to observable characteristics $(M_A - M_B)$ is that due to changes in the distributions of income among each combination of characteristics; that is, it is a "within-group" effect. This may be interpreted as the effects of changes in the "prices" of observable characteristics, but it will in fact also reflect the effects of changes in unobservable characteristics.

Decomposition results

Although there are numerous decompositions possible, we focus only on changes over the entire sample period 1982 to 1997-8 and on the effects of the above three groups of characteristics. The results of the decomposition analysis are presented in Figures 5.1 and 5.2 and Table 5.1.

Figure 5.1 displays the actual densities for weekly private income unit income in 1982 and 1997-8 along with the counterfactual density derived by reweighting observations in the 1997-8 sample to have the same distribution of observed characteristics (as measured by the included variables discussed above). Figure 5.2 displays the same information, but with the counterfactual density constructed by reweighting the 1982 sample to have the same distribution of characteristics as in 1997-8. The effects of changes in characteristics are given by the move from the counterfactual density to the 1997-8 density in Figure 5.1, and by the move from the 1982 density to the counterfactual density in Figure 5.2. The residual change between 1982 and 1997-8 can be interpreted as the change in the distribution of incomes associated with each set of characteristics, or more simply as changes in the "price" of each bundle of characteristics. Any differences

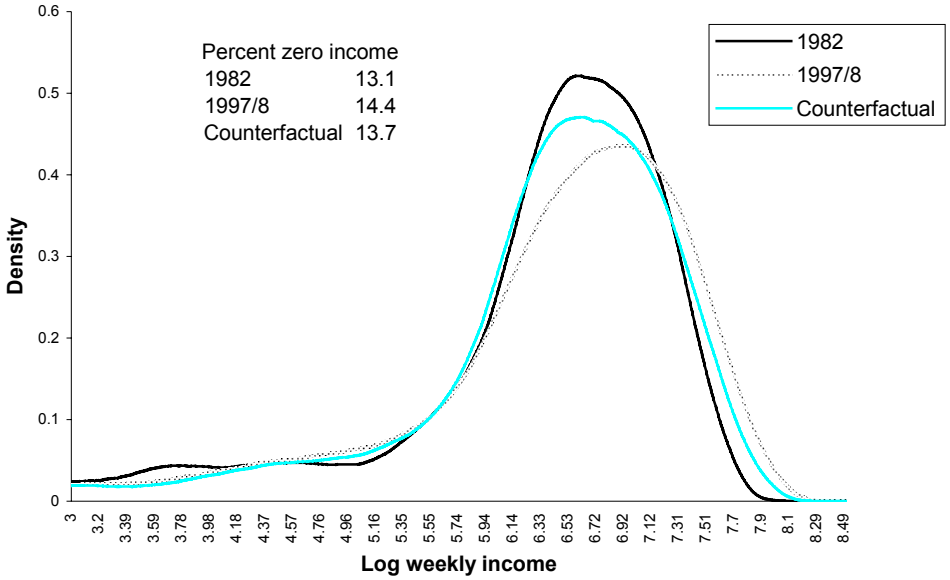
---

(although note that in this case it might be argued that the change in labour force status is caused by the educational attainment change, in which case it is correct to attribute all of the effect to educational attainment, both directly and via its impact on labour force status). In general, we should expect effects to be greater the later a characteristic is eliminated from the logit equation. Consequently, as a sensitivity test, Appendix 3 reports results when demographic characteristics are eliminated first. Results are very similar. Note that it is not possible to completely eliminate income unit type variables before the variables for the other characteristics because of the nature of the variables constructed for these other characteristics. For example, the labour force status variables for couples are not defined for single persons.

between Figures 5.1 and 5.2 in the implied effects of changes in characteristics are the outcome of evaluating at different "prices" (1997-8 prices in the former, 1982 prices in the latter).

Both figures show that changes in the distribution of observed characteristics have acted to decrease the proportion of the population with income unit private income at middle levels and increase both the proportion of the population with high private income and the proportion with no private income. As a rough approximation, about half the decrease in the density at middle income levels and increase in the density at high income levels and at zero income between 1982 and 1997-8 is explained by changes to observed characteristics. This suggests that approximately half the growth in private income inequality over the sample period is due to changes to the distribution of income unit types, labour force status and demographic characteristics.

Figure 5.1 Distribution of log weekly income unit private income - December 2000 prices
Effects of changes to characteristics - Evaluated at 1997 "prices"



The peak of the counterfactual distribution is, however, at approximately the same income level as the 1982 distribution in Figure 5.1, and at approximately the same income as the 1997-8 distribution in Figure 5.2. This implies the rightward shift between 1982 and 1997-8 of the peak of the income distribution derives from changes in the "prices" of characteristics rather than from changes in the distribution of characteristics.

Figure 5.2 Distribution of log weekly income unit private income - December 2000 prices
Effects of changes to characteristics - Evaluated at 1982 "prices"

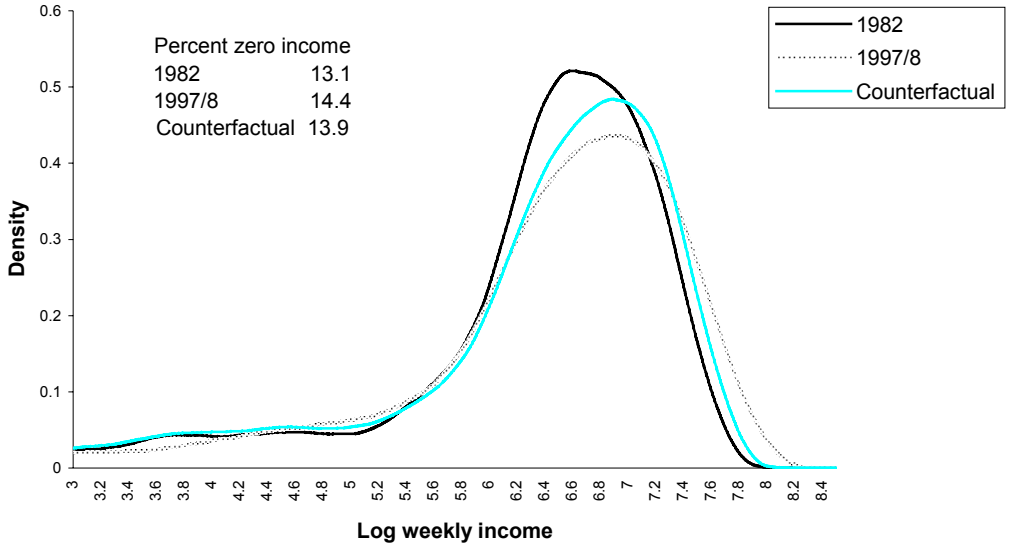


Table 5.1 permits us to quantify the distributional effects of changes to characteristics. It shows the actual change in selected distributional measures between 1982 and 1997-8, along with the effects of changes in characteristics, and with the characteristics effects themselves further decomposed into three groups of characteristics. As mentioned, the separate identification of the effects of each group was achieved by sequentially eliminating from the logit equation the variables capturing each group of characteristics.[23] Results are presented evaluated at both 1997-8 prices (reweighting the 1997-8 sample) and 1982 prices (reweighting the 1982 sample).

Consistent with the impression generated by the densities in Figures 5.1 and 5.2, approximately half the increase in inequality, as measured by the Gini coefficient, Theil coefficient and coefficient of variation, is accounted for by changes in characteristics. For example, whether evaluated at 1982 or 1997-8 prices, the increase in the Gini coefficient associated with changes to characteristics is 0.018, compared with a total increase of 0.034. Not evident from the densities is that characteristics also explain up to half the increase in the mean private income, and acted to increase the median income by as much as 30 percent more than the actual increase in the median. Evaluated at 1997-8 prices, the increase in the mean due to characteristics was $39, compared with a total increase of $74, while the increase in the median was $25, compared with an actual increase of $19.

Interestingly, only about one third of the increase in the 90-50 percentile log difference is explained by changes to characteristics. Although the role of wage changes are not explicitly explored in this paper, a possible explanation, consistent with the findings of Borland (1998), is that an important contributor to the increase in income dispersion in the top half of the distribution has been strong growth in wages at the top end of the wage distribution.

[23] See Appendix 3 for effects derived from an alternative sequence of elimination.

Table 5.1 Change in the income distribution 1982 to 1997-8 – Income unit weekly private income

	Actual change	Effect of changes in the distribution of							
		All characteristics		Labour force status		Demographic characteristics		Income unit types	
		1997-8 prices	1982 prices	1997-8 prices	1982 prices	1997-8 prices	1982 prices	1997-8 prices	1982 prices
Mean	73.72	38.96	28.11	-3.64	-3.44	54.18	42.11	-11.57	-10.56
Median	19.45	25.17	13.06	-17.51	-17.42	50.33	45.73	-7.66	-15.24
Gini	0.034	0.018	0.018	0.019	0.015	-0.007	-0.005	0.005	0.007
Theil	0.064	0.034	0.033	0.037	0.028	-0.014	-0.008	0.011	0.013
CV	0.069	0.029	0.029	0.033	0.026	-0.016	-0.011	0.012	0.014
P90-50	0.141	0.043	0.058	0.050	0.046	-0.013	-0.006	0.005	0.019
P50-10	-	-	-	-	-	-	-	-	-
P75-25	0.378	0.404	0.488	0.384	0.302	-0.118	-0.043	0.138	0.229
Zero	0.013	0.007	0.008	0.014	0.010	-0.010	-0.006	0.003	0.004

Note: Effects evaluated at 1997 "prices" are derived from the counterfactual income distribution that is obtained by reweighting observations in the 1997 sample to produce the same distribution of characteristics in the reweighted sample as in the 1982 sample. Effects evaluated at 1982 "prices" are derived from a reweighting of the 1982 sample. 'Zero' is the proportion of the population with no income unit private income.

Identification of the separate effects of changes to the distribution of income unit types, labour force status and demographic characteristics reveals that, for the most part, it is changes in the distribution of labour force status that have been responsible for characteristics-induced increases in income inequality. For example, evaluated at 1997-8 prices, the Gini coefficient increases by 0.019 due to labour force status changes (and by 0.015 evaluated at 1982 prices). This implies that changes to the distribution of work in the population have been an important source of growth in private income inequality between 1982 and 1997-8. This would include the effects of the growth in both the number of workless households and dual-earner households that has been documented by Dawkins, Gregg and Scutella (2002). These changes have also had a negative effect on the average level of real private incomes, with labour force status changes decreasing the mean income by approximately $3.50 and the median income by $17.50.

Changes to the income unit composition of the population have also acted to both increase income inequality and decrease the average real level of private incomes, reflecting the relative growth in single person and sole parent income units. Evaluated at 1997-8 prices, the Gini coefficient increases by 0.005 (0.007 at 1982 prices) and the mean decreases by $11.57 ($10.56 at 1982 prices). Changes to demographic characteristics, by contrast, have acted to decrease income inequality and increase average real private incomes. The Gini coefficient decreases by 0.007 (evaluated at 1982 prices) due to demographic changes, which also act to increase the mean by $54 and the median by $50. The mean and median income increases most likely reflect the increase in educational attainment, and to a lesser extent the increase in the average age in the population, between 1982 and 1997-8.

6. Concluding comments

The results reported in Section 4 confirm the findings of other recent studies of trends in inequality in Australia. At an aggregate level, inequality of private income increased, particularly during the eighties and less so during the nineties. The effects on disposable income inequality of the increases at the private level were, however, significantly offset by the effects of transfer payments and income taxation. As a consequence, the increase in inequality of disposable income was much more muted than it might have been but for government intervention.

Among income unit types, the story is mixed. Average real disposable incomes of all main groups increased, but while the greatest increases were for one of the poorest groups, single persons with dependents, the smallest increases were for the second poorest group, single persons without dependents. Overall, between-group inequality appears to have increased. As was the case for the aggregate income distribution, increases at the private level were moderated by government action, with the increases in disposable income inequality much smaller. Within-group change at the disposable income level was erratic, with increases for some groups and decreases for others, and overall not much change over the period. Again, the changes in disposable income were the outcome of greater change at the private level moderated by government action through taxes and transfers.

The main new findings of the paper concern the decomposition of dispersion in private incomes at the aggregate level. Here we find:

- About half the increase in income inequality is explicable by changes to observable characteristics, comprising changes in income unit composition, changes in demographic factors and changes in labour force status;
- Changes in labour force status alone have acted to increase inequality by the full amount attributed to characteristics;
- Changes to the income unit composition of the population have also had a positive (but smaller) effect income unit income inequality, while changes in the distribution of the observed demographic characteristics have had a negative effect on inequality (of a similar magnitude to the effect of changes to the income unit composition); and
- Much of the increase in inequality not attributable to characteristics – that due to changes in the "prices" of characteristics – is likely to reflect changes in wage rates. For example, it is consistent with our decomposition analysis that there has been a relative increase in wage rates at the top end of the wage distribution.[24]

How do we interpret these findings? The decomposition approach adopted does not involve the imposition of any assumptions, nor does it deliver any information, on the causal relationships between incomes and characteristics, or the reasons for changes over time in the characteristics composition of the population and the relationships between incomes and characteristics. For example, although the DFL approach identifies the important role of changes in the distribution of work in increasing income inequality, it provides no explanation for this change. In particular,

[24] However, note that we do not explicitly identify the role played by changes in wage rates. Thus, for example, we cannot exclude the possibility that changes in wage rates have acted to increase income inequality by more than is implied by our analysis, but are offset by some other unobserved effect that acts to decrease income inequality.

the relative roles of supply and demand factors, and the specific nature of these factors, are not identified by the decomposition method. Rather, the value of the decomposition is that it informs us that the changes in labour force participation have been very important for their effects on the income distribution. Therefore, from an income inequality perspective, the distribution of labour market activity is an important issue for policy and for further research.

Although the possibilities for further work are numerous, several avenues of research stand out as particularly worthy of inquiry. First is more detailed study of the effects of changes in specific economic and social characteristics of the Australian population. For example, this might include examination of the effects of increased female labour force participation, the rise in the number of workless income units and the growth in single parent income units. The DFL method is quite readily adapted to such analysis, although issues may arise over the most appropriate approach to constructing variables to measure the effects of specific changes.

Second, it would be informative to identify of the effect of changes to the distribution of wages on changes to the private income distribution. This is complicated by the fact that, in the same manner as income, wages changes can derive from changes to characteristics and from changes to the wages attached to each bundle of characteristics. It may therefore be necessary to adopt an alternative perspective to the DFL approach in order to identify the role of wage changes. For example, it may be appropriate to attempt to decompose distributional changes into those due to changes in the wage distribution, those due to changes in the distribution of other income and those due to changes in the relationship between wages and other income. It is not clear how this would be implemented. One possibility, however, is to adopt the approach of Burtless (1999), who investigates the effects of growing wage disparities and changing family composition on the extent of dispersion of the US income distribution. The approach involves first normalising mean incomes and earnings to be the same in base- and end-years, and then examining the extent of dispersion in the income distribution in the end-year holding constant the extent of earnings dispersion at the base-year level. This approach does, however, only identify the effect of changes in earnings inequality on the extent of income *dispersion*, and not the effects on the income distribution more generally.[25]

Third, the use of alternative data sources permitting us to examine expenditure rather than income, different time frames, and perhaps decompose distributional changes using other characteristics would be an important addition.

References

Australian Bureau of Statistics (ABS) 2001a, *Australia Now: Population, households and families*, downloaded from Ausstats website www.abs.gov.au/ausstats Canberra

— 2001b, *Labour Force statistics*, downloaded from Ausstats website www.abs.gov.au/ausstats Canberra

— 2001c, *Labour Force and Other Characteristics of Families, Australia*, Catalogue No. 6224.0, various issues, Canberra.

[25] Burtless finds that, while US income inequality rose sharply after 1979, the direct contribution of increased earnings inequality was surprisingly modest. Even if male and female earnings inequality remained unchanged at 1979 levels, about two-thirds of the observed increase in overall US inequality would still have occurred. Other factors contributing to higher overall inequality include the growing correlation between husband and wife earned incomes and the increased percentage of Americans who live in single adult families. These families have much more unequal incomes than couple families.

Barrett, G.F, Crossley, T F and Worswick, C. 2000, 'Consumption and Income Inequality in Australia', *Economic Record* Vol 76. No 233, June 2000

Blacklow P and Ray R. 2000, 'A comparison of Income and Expenditure Inequality Estimates: The Australian Evidence, 1975-76 and 1993-94', *The Australian Economic Review, 33, 317*

Borland, J. (1998) 'Earnings Inequality in Australia: Changes, causes and consequences,' Centre for Economic Policy Research Discussion Paper No. 390, ANU.

Borland J and Kennedy S. 1998, 'Earnings Inequality in Australia in the 1980s and 1990s', Centre for Economic Policy Research Discussion Paper No. 389, ANU

Borland, J, Gregory, R and Sheehan, P. 2001, 'Inequality and economic change', Chapter 1 in J Borland, B Gregory and P Sheehan (eds) *Work Rich, Work Poor – Inequality and economic change in Australia*, Centre for Strategic Economic Studies, Melbourne

Burbidge, A and Sheehan, P. 2001, 'The polarisation of families', Chapter 7 in J Borland, B Gregory and P Sheehan (eds) *Work Rich, Work Poor – Inequality and economic change in Australia*, Centre for Strategic Economic Studies, Melbourne

Burtless, G. 1999, 'Effects of growing wage disparities and changing family composition on the US income distribution', *European Economic Review* 43, 853-65

Dawkins, P, Gregg, P and Scutella, R. 2001, *The Growth of Jobless Households in Australia*, Melbourne Institute Working Paper No 3/01, Melbourne Institute of Applied Economic and Social Research, Melbourne

Dawkins, P, Gregg, P and Scutella, R. 2002, *The Growth of Jobless Households and the Polarisation of Employment in Australia*, paper presented at the 2002 towards Opportunity and Prosperity Conference, Melbourne.

Di Nardo, J, N. Fortin and Lemieux, T. 1996, 'Labour Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach', *Econometrica*, 64(5), 1001-44

Harding, A and Greenwell, H. 2001, Trends in Income and Expenditure Inequality in the 1980s and 1990s, paper presented to 30th Annual Conference of Economists, 24 September

Hyslop, D and Mare, D. 2001, Understanding changes in the distribution of household incomes in New Zealand between 1983-6 and 1995-8, Treasury Working Paper 1/21

Pappas N. 2001, 'Family Income Inequality', Chapter 2 in J Borland, B Gregory and P Sheehan (eds) *Work Rich, Work Poor – Inequality and economic change in Australia*, Centre for Strategic Economic Studies, Melbourne

Silverman, B. (1986) *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Wooden, M and Loundes, J. 2002, *How unreasonable are long working hours?* Melbourne Institute Working Paper No 1, Melbourne

Wooden, M. 2001, The Growth in Unpaid Working Time', *Economic Papers*, 20 (1) March, 29-44

Appendices

Appendix 1: Kernel density estimates

The weighted kernel density estimate $\hat{f}_h$ of a univariate density f based on a random sample $y_1,...,y_n$ with weights $\theta_1,...,\theta_n$, $\sum_i \theta_i = 1$, is:

$$\hat{f}_h(y) = \sum_{i=1}^{n} \frac{\theta_i}{h} K\left(\frac{y - y_i}{h}\right) \tag{0.1}$$

where h is the bandwidth and $K(.)$ is the kernel function, and which together regulate the relationship between the distance of y_i from y and the weight given to observation i in the estimation of the density at y. The kernel function chosen is not critical, but it should be positive, integrate to unity, be symmetric around zero and decreasing in the absolute value of its argument. We adopt an Epanechnikov kernel:

$$K(x) = \begin{cases} \frac{3}{4}\left(1 - \left(\frac{1}{5}\right)x^2\right) \Big/ \sqrt{5} & \text{if } |x| < \sqrt{5} \\ 0 & \text{otherwise} \end{cases} \tag{0.2}$$

More critical is the choice of bandwidth, with smaller bandwidths resulting in less bias but more variance than larger bandwidths. We adopt the rule of thumb bandwidth (due to Silverman (1986):

$$h = 0.9 \min\left(\sigma_y, IQR/1.34\right) n^{-1/5} \tag{0.3}$$

where σ_y is the estimated standard deviation of the density and IQR is the inter-quartile range. This results in bandwidths of approximately 0.05 to 0.15 for the distributions examined in this paper.

Appendix 2: The DFL decomposition method

The DFL approach views each observation i as a vector (y_i, x_i, t_i) consisting of an income y, a vector of characteristics x and a date t, that belongs to a joint distribution $F(y,x,t)$ of incomes, characteristics and dates. The density of incomes at one point in time $f_t(y)$ can then be written as the integral of the density of incomes conditional on a set of individual characteristics and on a date t_p, over the distribution of individual characteristics $F(x \mid t_x)$ at date t_x:

$$\begin{aligned} f_t(y) &= \int_{x \in \Omega_x} f\left(y \mid x, t_p = t\right) dF\left(x \mid t_x = t\right) \\ &\equiv f\left(y; t_p = t, t_x = t\right) \end{aligned} \tag{0.4}$$

where Ω_x is the domain of definition of the individual characteristics.

The notation in the second line of equation (0.4) allows us to express equations for counterfactual densities, with t_p denoting the date from which the incomes associated with each set of characteristics are drawn (i.e the characteristic "prices"), and t_x denoting the date from which the distribution of characteristics is drawn. For example, while $f\left(y;t_p=97,t_x=97\right)$ represents the actual density of incomes in 1997-8, $f\left(y;t_p=97,t_x=82\right)$ represents the density that would have resulted in 1997-8 had the distribution of characteristics remained as it was in 1982. This hypothetical density is identified as follows:

$$\begin{aligned} f\left(y;t_p=97,t_x=82\right) &= \int f\left(y\mid x,t_p=97\right)dF\left(x\mid t_x=82\right) \\ &= \int f\left(y\mid x,t_p=97\right)\psi_x\left(x\right)dF\left(x\mid t_x=97\right) \end{aligned} \tag{0.5}$$

where ψ_x is a "reweighting"' function:

$$\psi_x\left(x\right)\equiv\frac{dF\left(x\mid t_x=82\right)}{dF\left(x\mid t_x=97\right)} \tag{0.6}$$

The counterfactual density is identical to the 1997-8 density except for the function $\psi_x\left(x\right)$, so that once an estimate of this function, $\hat{\psi}_x\left(x\right)$, is obtained, the counterfactual density can be estimated by weighted kernel methods as:

$$\hat{f}\left(y;t_p=97,t_x=82\right)=\sum_{i=1}^{n_{97}}\frac{\theta_i}{h}\hat{\psi}_x\left(x_i\right)K\left(\frac{y-y_i}{h}\right) \tag{0.7}$$

where n_{97} is the 1997-8 sample size and the summation is over observations in the 1997-8 sample.

Essentially, each observation in the 1997-8 sample is reweighted so as to give the same distribution of individual characteristics as in the 1982 sample. Applying Bayes' rule to the ratio $\frac{dF\left(x\mid t_x=82\right)}{dF\left(x\mid t_x=97\right)}$ gives the following equation for the reweighting function:

$$\psi_x\left(x\right)=\frac{\Pr\left(t_x=97\right)}{\Pr\left(t_x=82\right)}.\frac{\Pr\left(t_x=82\mid x\right)}{\Pr\left(t_x=97\mid x\right)} \tag{0.8}$$

Since the weights θ_i sum to one for each date we can set $\Pr\left(t_x=97\right)=\Pr\left(t_x=82\right)$ and ignore these unconditional probabilities. The probability of being in period t given characteristics x $\left(\Pr\left(t_x=t\mid x\right)\right)$ can be estimated non-parametrically, by identifying the proportion of individuals with each characteristic combination at each date, or by a discrete choice model like the logit. We take the latter approach, since implementation is straightforward and the equivalent of the non-parametric model can be achieved by specifying a full set of dummy variables indicating each possible value of x:

$$\Pr\left(t_x=t\mid x\right)=\Lambda\left(H\left(x\right)\beta\right) \tag{0.9}$$

where Λ is the cumulative logistic distribution and $H(x)$ is a vector of covariates that is a function of x.

The weights resulting from estimation of this model are normalised to sum to one and can then be used to estimate the counterfactual density as per equation (0.7) or indeed any summary measure of the counterfactual distribution, such as the mean or Gini coefficient.

Using the logit equation estimates, the 1997-8 observations have therefore been reweighted to produce a counterfactual earnings distribution: that which would have prevailed in 1997-8 if the characteristic composition, *as measured by the included variables*, had remained as it was in 1982, and the "prices" of the characteristics had remained at their 1997-8 levels. The effects of changes in particular characteristics can also be isolated by the selective exclusion of variables from the logit equation estimated. This does require additional assumptions, however, since if the excluded variables are correlated with included variables, the distribution of these variables in the reweighted sample will be altered from their actual distribution, and so we will also pick up this effect in the counterfactual density.

By comparing the counterfactual density with the actual densities we can get a visual representation of the effects of changes in characteristics versus changes to the incomes associated with each set of characteristics. For example, comparison of the above counterfactual density with the actual 1997-8 density shows the effect of changes to characteristics between 1982 and 1997-8, if both 1982 and 1997-8 individuals were to face 1997-8 "prices". Alternatively, the effects of characteristics changes can be evaluated at 1982 "prices" by estimating the density that would have prevailed in 1982 if the distribution of characteristics was the same as in 1997-8.

Appendix 3: Alternative decomposition results

Table 3A Change in the income distribution 1982 to 1997-8 – Income unit current weekly private income – Alternative Decomposition

	Actual change	Effect of changes in the distribution of:							
		All characteristics		Labour force status		Demographic characteristics		Income unit types	
		1997-8 prices	1982 prices	1997-8 prices	1982 prices	1997-8 prices	1982 prices	1997-8 prices	1982 prices
Mean	73.72	38.96	28.11	7.43	2.37	43.11	36.30	-11.57	-10.56
Median	19.45	25.17	13.06	-1.09	-10.89	33.92	39.19	-7.66	-15.24
Gini	0.034	0.018	0.018	0.014	0.015	-0.002	-0.005	0.005	0.007
Theil	0.064	0.034	0.033	0.027	0.027	-0.004	-0.007	0.011	0.013
CV	0.069	0.029	0.029	0.022	0.025	-0.005	-0.010	0.012	0.014
P90-50	0.141	0.043	0.058	0.027	0.044	0.010	-0.005	0.005	0.019
P50-10	-	-	-	-	-	-	-	-	-
P75-25	0.378	0.404	0.488	0.311	0.291	-0.046	-0.031	0.138	0.229
Zero	0.013	0.007	0.008	0.007	0.006	-0.003	-0.002	0.003	0.004

Note: Demographic variables are eliminated from the logit equation first, followed by labour force status variables.